FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

POCH
QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL
PORTUGAL 2007.2013
Governo da República Portuguesa
UNIÃO EUROPEIA
Fundo Social Europeu

Cofinanciado por:

COMPETE 2020
PORTUGAL 2020
UNIÃO EUROPEIA
Fundo Social Europeu

Probing the effects of emotion and encoding task on source memory attributions: Behavioral and electrophysiological evidence
Diana Cristina Rodrigues Pereira
UMinho | 2020

Universidade do Minho
Escola de Psicologia

Diana Cristina Rodrigues Pereira

Probing the effects of emotion and encoding task on source memory attributions: Behavioral and electrophysiological evidence

junho de 2020

Universidade do Minho
Escola de Psicologia

Diana Cristina Rodrigues Pereira

Probing the effects of emotion and encoding task on source memory attributions: Behavioral and electrophysiological evidence

Tese de Doutoramento em Psicologia Básica

Trabalho efetuado sob a orientação da
Professora Doutora Ana Patrícia Teixeira Pinheiro
e da
Professora Doutora Adriana da Conceição Soares Sampaio

junho de 2020

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição
CC BY

https://creativecommons.org/licenses/by/4.0/

AGRADECIMENTOS[1]

À Fundação para a Ciência e a Tecnologia (FCT), bem como ao Centro de Investigação em Psicologia (CIPsi) da Universidade do Minho por providenciarem as condições necessárias para concretizar os vários estudos e actividades associadas a este projecto de investigação.[2]

Às orientadoras, professoras Ana Pinheiro e Adriana Sampaio, por terem abraçado este projecto e por me apoiarem nas diferentes etapas desta jornada. Pelo vosso tempo, disponibilidade e pelas oportunidades de aprendizagem.

À comissão de acompanhamento do programa doutoral em psicologia básica (professora Ana Paula Soares; professor Marco Vasconcelos; professor Alberto Crego; professora Joana Coutinho), bem como aos professores do programa doutoral pela dedicação e pelo empenho. Ao Grupo de Investigação em Memória Humana e, em especial, ao professor Pedro Albuquerque, à professora Flávia Santos e à Patrícia Matos. Ao professor Mikael Johansson e à professora Inês Bramão, bem como ao Departamento de Psicologia da Universidade de Lund (Suécia) por me receberem durante o estágio de investigação (tusen tack!). Agradeço também aos meus colegas de doutoramento, pelas experiências que partilhamos (ainda que fugazes) e pela entreajuda.

A todos os participantes que colaboraram nos vários estudos deste doutoramento, tanto presencialmente como em plataformas online. O vosso interesse, empenho e tempo são uma verdadeira dádiva tanto para a investigação, como para todos nós enquanto sociedade. O meu genuíno obrigada!

Aos que me prestaram apoio durante as recolhas de ressonância magnética, nomeadamente aos técnicos do SMIC da Boavista, ao Moisés Silva, ao Miguel Soares e ao Alberto González-Villar. À Sara (Snow) e à Anabela pelo apoio nas mais variadas etapas destes (exigentes) processos de recolha.

A todos os membros do Laboratório de Neurociência Psicológica, os que tiveram de passagem, aqueles cujos percursos se cruzaram com o meu em diferentes momentos (Ana G., Andreia, Eugénia, Helga, Sílvia) e, claro, aos que actualmente fazem parte deste (incrível) grupo. Em especial, à Carlota, à Joana, à Margarida e à Paulinha pelos momentos de partilha e pelo apoio. Ao Augusto, ao Diogo e à Marlene por estarem sempre receptivos a ajudar, especialmente em situações que envolviam a variável improviso. Ao Alberto e à Carol por serem um exemplo de auto-determinação e resiliência, pelo companheirismo e por me inspirarem continua e diariamente.

Por último, agradeço à Joana C. e à Andreia S. pelos passeios e pelos momentos de reencontro, ainda que (demasiado) breves. À minha irmã e aos meus pais por estarem sempre presentes. À Leela, à Poe e à Mi por animarem os meus dias. Ao José por ser um pilar e por me (re)lembrar que mesmo à deriva (em momentos de serenidade ou intempérie), há sempre um porto seguro à nossa espera.

Mais uma vez, muito obrigada por tudo e a todos!

To this enriching journey we call life:

"(...) For still there are so many things/that I have never seen:/

in every wood in every spring /there is a different green. (...)" - J. R. R. Tolkien

[1] To future readers of this dissertation: we would like to challenge you to search for a song entitled "**Remembering**" by Avishai Cohen. Although this section is written in Portuguese, we hope this song conveys to you how grateful we are for the journey that led up to this moment. This acknowledgment section is all about "**remembering**" the people and institutions that supported us throughout this journey and, as you may foresee, this dissertation is also about "**remembering**"!

[2] This work was funded by the Portuguese Foundation for Science and Technology (FCT, Portugal), with the doctoral grant with the reference PD/BD/105964/2014. The scientific research reported here was conducted at the at the Psychology Research Centre (PSI/01662), University of Minho, supported by FCT and the Portuguese Ministry of Science, Technology and Higher Education (UID/PSI/01662/2019) through national funds (PIDDAC) and co-funded by FEDER through COMPETE2020 under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653).

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Probing the effects of emotion and encoding task on source memory attributions: Behavioral and electrophysiological evidence

The effects of emotional information on source memory, representing the ability to retrieve and decide about features that constitute a given event, have yielded mixed evidence. Accordingly, the studies of this dissertation probed whether the way information is encoded contributes to clarify such discrepant findings. The broad goal of this work was to explore the effects of emotion and encoding task (e.g., self-referential *vs*. semantic processing) on internal source memory attributions and prospective metamemory judgments of source, including not only behavioral but also electrophysiological indices. In the first study, a meta-analytic overview on the effects of emotional information (valence: negative/neutral/positive; arousal: low/medium/high) on source memory was conducted. Valence-based studies revealed null to very small effects, while arousal based-studies showed beneficial effects of medium/high (*vs*. low) arousal on source memory accuracy. However, due to the heterogeneity verified across studies, moderator and sub-group analyses were computed, showing that factors such as the type of source memory tasks and of source memory features influenced the size and direction of the effects. In the second study, two distinct experiments revealed that whereas the source memory for production mode (read silently *vs*. aloud) did not differ between emotional and neutral words, the source memory for self-referentially encoded positive/neutral words was better than for negative ones, and it was better for neutral than for emotional words in the semantic condition. In both experiments, even though judgments of source did not predict later source memory performance, they were modulated by valence and encoding tasks. In the third study, instead of manipulating the affective properties of the stimuli, emotional encoding contexts together with neutral stimuli were examined in two experiments, using a mixed and a blocked design. No reliable emotion-related modulations on source memory were observed in this case. In the fourth study, examining electrophysiological correlates of source memory, words studied in self-referential conditions were associated with enhanced mean amplitudes during encoding (700 – 1200 ms) and with an early old/new effect (300 – 500 ms) during recognition, which was not detected for the semantic condition. Similarly, in the 500 – 800 ms latency window, no old/new effect emerged for negative words studied in the semantic condition. Additionally, negative self-referential and correct rejections mobilized greater post-retrieval monitoring processes, as indicated by an increased mean amplitude in the 800 – 1200 ms window. Overall, the findings of this dissertation corroborated that both source memory and judgments of source are modulated as a function of emotion and encoding task. Thus, the way information is encoded is a relevant factor to consider when probing the interplay between source memory and emotion.

Keywords: emotion; judgments of source; self-reference effect; source memory; valence.

Os efeitos da emoção e da tarefa de codificação em atribuições de memória de fonte: Evidência comportamental e eletrofisiológica

Os efeitos da informação emocional na memória de fonte, que é a capacidade de recuperar e de decidir sobre características que compõe um dado evento, têm produzido evidências contraditórias. Neste sentido, os estudos desta dissertação indagaram se a maneira como a informação é codificada pode contribuir para clarificar estes achados discrepantes. O objetivo geral deste trabalho foi assim explorar os efeitos da emoção e tarefa de codificação (processamento auto-referencial *vs*. semântico) em atribuições internas de memória de fonte e em julgamentos prospetivos de metamemória de fonte, incorporando não só indicadores comportamentais, mas também eletrofisiológicos. Num primeiro estudo, uma meta-análise sobre os efeitos da informação emocional (valência: negativa/neutra/positiva; ativação: baixa/média/alta) na memória de fonte foi conduzida. Estudos baseados na valência revelaram efeitos nulos a muito pequenos, enquanto estudos baseados na ativação mostraram efeitos benéficos na precisão da memória de fonte no caso da ativação média/alta (*vs*. baixa). Contudo, dada a heterogeneidade verificada nos estudos, análises de moderação e de subgrupo foram concretizadas, revelando que fatores como o tipo de tarefas de memória de fonte e o tipo de características testadas influenciaram a direção e o tamanho dos efeitos. Num segundo estudo, duas experiências distintas revelaram que, enquanto a memória de fonte para o modo de produção (lido silenciosamente *vs*. em voz alta) não diferiu entre palavras emocionais e neutras, a memória de fonte para palavras positivas/neutras auto-referenciadas foi melhor do que para palavras negativas, sendo também melhor para palavras neutras do que para palavras emocionais no caso da condição semântica. Em ambas as experiências, embora os julgamentos de fonte não foram preditivos da performance de memória de fonte, estes foram influenciados pela valência e pelas tarefas de codificação. Num terceiro estudo, em vez da manipulação da emoção ser introduzida ao nível das propriedades afetivas dos estímulos, esta foi introduzida como elemento externo durante a codificação de estímulos neutros. Foram concretizadas duas experiências, uma com um desenho experimental misto e outra com um desenho em bloco. Não foram observados efeitos confiáveis de emoção neste caso. Num quarto estudo, que examinou os correlatos eletrofisiológicos da memória de fonte, palavras estudadas em condições auto-referenciais foram associadas a amplitudes médias aumentadas durante a codificação (700 – 1200 ms), bem como ao aparecimento de um efeito precoce antigo/novo durante o reconhecimento que não foi detetado na condição semântica. Da mesma forma, na janela de latência 500 – 800 ms, o efeito antigo/novo esperado não emergiu no caso de palavras negativas estudadas na condição semântica. Adicionalmente, processos de monitorização pós-recuperação foram mais mobilizados para estímulos negativos que foram auto-referenciados ou novos, tal como evidenciado por um aumento da amplitude média na janela de 800 – 1200 ms. No geral, os achados desta dissertação corroboraram que tanto a memória de fonte como os julgamentos de fonte são influenciados em função da emoção e da tarefa de codificação. Portanto, a forma como a informação é codificada é um fator relevante a considerar no estudo das sinergias entre memória de fonte e emoção.

Palavras-chave: efeito de auto-referência; emoção; julgamentos de fonte; memória de fonte; valência.

TABLE OF CONTENTS

CHAPTER I – INTRODUCTION AND THEORETICAL BACKGROUND 1

Source memory: The nuts and bolts..... 6

When source memory is tinged with emotion: How emotion and self-referential encoding affect our memory for source features?..... 8

Delving into prospective metamemory judgments of source: How does one perceive the effects of emotion on future source memory performance?..... 15

Neural correlates of source memory..... 17

Temporal dynamics of source memory: Event-related potential signatures..... 21

Neural basis of source memory: fMRI evidence..... 23

A brief note on methodological choices, main objectives, and outline of this dissertation..... 29

References..... 31

CHAPTER II – HOW DOES EMOTION AFFECT SOURCE MEMORY? A META-ANALYSIS OF BEHAVIORAL EVIDENCE..... 66

Abstract..... 67

Emotion effects on source memory..... 68

Theoretical approaches to the effects of emotion on source memory..... 70

Candidate factors to moderate the effects of emotion on source memory..... 73

The current meta-analysis..... 76

Method..... 77

Literature search strategy..... 77

Eligibility criteria..... 77

Study selection..... 78

Coding procedure (main and moderator variables)..... 78

Quality assessment..... 81

Meta-analytic procedure..... 81

Results..... 83

Characteristics of the studies included..... 83

Quality assessment..... 84

Publication bias..... 85

Main findings on source memory accuracy..... 85

Valence-based effects - Negative *vs.* neutral valence..... 85
Positive *vs.* neutral valence..... 88
Negative *vs.* positive valence..... 89
Arousal-based effects - High *vs.* low arousal..... 90
Medium *vs.* low arousal..... 90
High *vs.* medium arousal..... 91
Discussion..... 91
Limitations..... 103
Conclusion..... 104
References..... 105

CHAPTER III – DIFFERENTIAL EFFECTS OF VALENCE AND ENCODING STRATEGY ON INTERNAL SOURCE MEMORY AND JUDGMENTS OF SOURCE: EXPLORING THE PRODUCTION AND THE SELF-REFERENCE EFFECT..... 121
Abstract..... 122
Experiment 1..... 127
Material and methods..... 128
Participants..... 128
Materials..... 128
Procedure..... 129
Data analysis..... 130
Results..... 132
Recognition accuracy..... 132
Item memory..... 132
Source memory..... 133
Judgments of source (JOSs)..... 133
Discussion: Experiment 1..... 136
Experiment 2..... 137
Material and methods..... 137
Participants..... 137
Materials..... 138
Procedure..... 138

Data analysis..... 138
Results..... 138
Recognition accuracy..... 139
Item memory..... 139
Source memory..... 139
Judgments of source (JOSs)..... 140
Discussion: Experiment 2..... 140
General discussion..... 141
Limitations..... 147
Conclusion..... 148
References..... 149
Appendix..... 160

CHAPTER IV – IS INTERNAL SOURCE MEMORY RECOGNITION MODULATED BY EMOTIONAL ENCODING CONTEXTS?..... 171
Abstract..... 172
Experiment 1..... 178
Methods..... 179
Participants..... 179
Materials..... 180
Procedure..... 181
Data analysis..... 183
Results..... 184
Recognition accuracy..... 185
Item memory..... 185
Source memory..... 185
Response time..... 186
Study phase..... 186
Correct source responses..... 188
Manipulation check. 189
Discussion: Experiment 1..... 189
Experiment 2..... 191

Methods ... 191
Participants ... 191
Materials ... 192
Procedure ... 192
Data analysis ... 193
Results ... 193
Recognition accuracy ... 193
Item memory ... 193
Source memory ... 193
Response time ... 194
Study phase ... 194
Correct source responses ... 195
Manipulation check. ... 195
Discussion: Experiment 2 ... 196
General discussion ... 197
Limitations and future research ... 201
Conclusion ... 201
References ... 202

CHAPTER V – INTERACTIONS OF EMOTION AND SELF-REFERENCE IN SOURCE MEMORY: AN ERP STUDY ... 216
Abstract ... 217
Effects of emotion and self-reference during encoding: ERP evidence ... 218
Effects of emotion and self-reference on recognition memory: ERP evidence ... 220
The current study and hypotheses ... 223
Methods ... 224
Participants ... 224
Stimuli ... 225
Procedure ... 225
ERP data acquisition and analysis ... 228
Statistical analysis ... 229
Behavioral data ... 229

ERP data..... 230
Results..... 230
Behavioral results (recognition accuracy)..... 230
Item memory..... 230
Source memory..... 232
ERP results..... 233
Word encoding. 233
500 - 700 ms. 233
700 - 950 and 950 - 1200 ms. 233
Test phase..... 234
300 - 500 ms..... 234
500 - 800 ms..... 234
800 - 1200 ms..... 235
1200 - 1800 ms..... 236
Correlational analyses..... 236
Discussion..... 238
Behavioral findings..... 238
Emotion and self-reference effects during encoding..... 239
Emotion and self-reference effects during recognition..... 240
Limitations..... 243
Conclusion..... 244
References..... 244

CHAPTER VI – GENERAL DISCUSSION..... 259
Implications..... 269
Caveats and future directions..... 274
Concluding remarks..... 280
References..... 281

APPENDIX A – SUPPLEMENTARY MATERIALS - CHAPTER II..... 298
References..... 347

APPENDIX B – SUPPLEMENTARY MATERIALS - CHAPTER III 348
Method 357
Data analysis 357
Multinomial models 357
Goodman-Kruskal gamma correlation 358
Results 358
Experiment 1 359
Multinomial model results. 359
Incorrect source responses. 359
Do not know responses 359
Misses 360
Correct rejections and corrected false alarm rates 360
Response time 360
JOSs ratings 360
Source correct judgments 360
Correct rejections 360
Confidence ratings 361
Source correct judgments 361
Correct rejections 361
Experiment 2 361
Mean proportion of “yes” responses 361
Multinomial model results 361
Incorrect source responses 362
Do not know responses 362
Misses 363
Correct rejections and corrected false alarm rates 363
Response time 363
JOSs ratings 363
Source correct judgments 363
Correct rejections 364
Confidence ratings 364
Source correct judgments 364

Correct rejections..... 364
References..... 364

APPENDIX C – SUPPLEMENTARY MATERIALS - CHAPTER IV..... 366
Method..... 381
Materials: The case of a new pool of 645 European Portuguese adjectives..... 381
Data analysis..... 382
Multinomial models..... 382
Bayesian analysis..... 382
Results..... 383
Experiment 1..... 383
Multinomial model results..... 383
Bayesian results..... 384
Incorrect source responses..... 385
Do not know responses..... 386
Misses..... 386
Additional manipulation check..... 386
Experiment 2..... 387
Multinomial model results..... 387
Bayesian results..... 387
Incorrect source responses..... 388
Do not know responses..... 389
Misses..... 389
Correct rejections and false alarms..... 389
Additional manipulation check..... 389
Comparison across experiments..... 390
References..... 391

APPENDIX D – SUPPLEMENTARY MATERIALS - CHAPTER V..... 394
Method..... 404
Data analysis..... 404
Multinomial model..... 404

Goodman-Kruskal gamma correlation..... 404
Results..... 405
Multinomial model results..... 406
Response time during word encoding..... 407
Incorrect source memory responses..... 407
Do not know responses..... 408
Misses..... 408
Correct rejections and false alarms..... 408
Proportion of “yes” responses in the word encoding phase..... 409
Judgments of source (JOSs)..... 409
References..... 410

APPENDIX E – SUPPLEMENTARY MATERIALS - CHAPTER VI..... 413
References..... 418

APPENDIX F – APPROVAL BY THE SUBCOMMISSION OF LIFE AND HEALTH SCIENCES - UMINHO ETHICS COMMISSION..... 419

ABBREVIATIONS

ABC – Arousal-Biased Competition

AG – Angular Gyrus

ANEW – Affective Norms for English Words

ANOVA – Analysis of Variance

BDI-II – Beck Depression Inventory-II

BIC – Binding of Item and Context

BOLD – Blood Oxygen Level Dependent

CC – Cingulate Cortex

EEG - Electroencephalography

EEM – Emotion-Enhanced Memory

EOL – Ease-of-Learning

EP – European Portuguese

EPN – Early Posterior Negativity

ERP(s) – Event-Related Potential(s)

fMRI – Functional Magnetic Resonance Imaging

HEOG – Horizontal Electrooculogram

IADS – International Affective Digitized Sounds

IAPS – International Affective Picture System

ICA – Independent Component Analysis

JOLs - Judgments of Learning

JOSs – Judgments of Source

LPC – Late Positive Complex

LPN – Late Posterior Negativity

MEG – Magnetoencephalography

MCI – Mild Cognitive Impairment

mPFC – Medial Prefrontal Cortex

MTL – Medial Temporal Lobe

PANAS – Positive and Negative Affect Schedule

PET – Positron Emission Tomography

PFC – Prefrontal Cortex

P-PAL - Procura-PALavras

PPC – Posterior Parietal Cortex

PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses

RMANOVA – Repeated-Measures Analysis of Variance

ROI(s) – Region(s)-of-Interest

SAM – Self-Assessment Manikin

SPSS – Statistical Package for the Social Sciences

STAI-Y – State-Trait Anxiety Inventory - Form Y

SW(s) – Slow Wave(s)

TMS – Transcranial Magnetic Stimulation

VEOG – Vertical Electrooculogram

LIST OF FIGURES

CHAPTER I – INTRODUCTION AND THEORETICAL BACKGROUND

Figure 1. Schematic illustration of the old/new effects and summary of their latencies, topographies, and effects of emotion on their amplitude..... 25

CHAPTER II – HOW DOES EMOTION AFFECT SOURCE MEMORY? A META-ANALYSIS OF BEHAVIORAL EVIDENCE

Figure 1. Flow diagram of the search strategy according to the PRISMA guidelines (Moher et al., 2009)..... 79

Figure 2. Funnel plot graphs for each analysis (valence and arousal analyses on the left side, and arousal analyses on the right side)..... 86

Figure 3. Summary of the main findings obtained for each analysis. Of note, for the moderator and sub-group analyses, only the statistically significant results were presented..... 87

CHAPTER III – DIFFERENTIAL EFFECTS OF VALENCE AND ENCODING STRATEGY ON INTERNAL SOURCE MEMORY AND JUDGMENTS OF SOURCE: EXPLORING THE PRODUCTION AND THE SELF-REFERENCE EFFECT

Figure 1. Schematic illustration of the tasks used in Experiments 1 and 2..... 129

Figure 2. Behavioral results for JOSs, item memory and source memory in Experiments 1 and 2..... 136

CHAPTER IV – IS INTERNAL SOURCE MEMORY RECOGNITION MODULATED BY EMOTIONAL ENCODING CONTEXTS?

Figure 1. Schematic illustration of an experimental trial from encoding and retrieval phases..... 183

Figure 2. The recognition scores obtained for item (panel a – Experiment 1; panel b – Experiment 2) and source memory (panel c – Experiment 1; panel d – Experiment 2) are plotted on the y-axis as a function of source (self-reference/common) and valence of the background picture (negative/neutral/positive) presented during encoding..... 188

Figure 3. The response times obtained during the study phase (panel a – Experiment 1; panel b – Experiment 2) and during the recognition phase for source hits are plotted on the y-axis as a function of source (self-reference/common) and valence of the background picture (negative/neutral/positive) shown during encoding..... 189

CHAPTER V – INTERACTIONS OF EMOTION AND SELF-REFERENCE IN SOURCE MEMORY: AN ERP STUDY

Figure 1. Schematic illustration of the event-related potentials of interest in the current study, including the main findings reported in the literature concerning the effects of emotion and self-reference on the amplitude of each component..... 220

Figure 2. Schematic illustration of an experimental trial from the study and test phases. The number and type of stimuli regarding each encoding task (self *vs.* common) x valence (negative *vs.* neutral *vs.* positive) condition are also summarized..... 227

Figure 3. Behavioral results for item memory and source memory are shown in panel A and panel B, respectively..... 233

Figure 4. Grand average waveforms obtained during word encoding (panel A) together with their respective topographic maps (panel B) considering each encoding task (self *vs.* common) x valence (negative *vs.* neutral *vs.* positive) condition..... 234

Figure 5. Grand average waveforms and topographical maps obtained during the test phase considering the four-time windows analyzed..... 235

Figure 6. Correlation analyses between the mean amplitude obtained in the 500 – 800 ms (panel A) and in the 800 – 1200 ms (panel B) latency windows and the proportion of correct source responses for each encoding task (self *vs.* common) and valence (negative *vs.* neutral *vs.* positive) conditions..... 237

Figure 7. Schematic illustration of the event-related potentials of interest in the current study, including its main findings..... 237

CHAPTER VI – GENERAL DISCUSSION

Figure 1. Forest plots for the six comparisons conducted with the behavioral data collected in the context of this dissertation..... 263

Figure 2. Schematic illustration of the main findings of the studies included in this dissertation.. 269

APPENDIX A – SUPPLEMENTARY MATERIALS - CHAPTER II

Figure S1. Funnel plot graphs after implementing the trim-and-fill method..... 342

Figure S2. Forest plots for the six meta-analyses: negative *vs.* neutral (panel A); positive *vs.* neutral (panel B); negative *vs.* positive (panel C); high *vs.* low (panel D); medium *vs.* low (panel E); high *vs.* medium (panel F)..... 343

APPENDIX C – SUPPLEMENTARY MATERIALS - CHAPTER IV

Figure S1. Outputs of the model-based Bayesian analysis (BF_{10}) obtained with JASP for the comparison between experiments 1 and 2..... 379

APPENDIX D – SUPPLEMENTARY MATERIALS - CHAPTER V

Figure S1. In panel A, the mean response time (RT) obtained during word encoding is plotted on the y-axis as a function of encoding task (self-reference *vs.* common) and valence (negative *vs.* neutral *vs.* positive). In panel B, a plot of the mean judgments of source (JOSs) ratings obtained during the study phase is also presented..... 402

Figure S2. Illustration of grand average waveforms obtained during judgments of source (JOSs; panel A) together with the topographic maps (panels B) considering each encoding task (self *vs.* common) x valence (negative *vs.* neutral *vs.* positive) condition..... 403

APPENDIX E – SUPPLEMENTARY MATERIALS - CHAPTER VI.....

Figure S1. Schematic illustration of an experimental trial from the study and test phases of the fMRI study..... 415

Figure S2. Behavioral findings for item and source memory (panel A and B, respectively). The performance is plotted on the y-axis as a function of emotion (negative/neutral/positive) and source (self-reference/common)..... 417

LIST OF TABLES

CHAPTER I - INTRODUCTION AND THEORETICAL BACKGROUND

Table 1. *Examples of Features Commonly Manipulated in the Context of Source Memory Tasks....* 9

Table 2. *Summary of the Effects of Emotional Stimuli on Source Memory Accuracy Considering the Type of Source Memory Task and of Features.....* 13

CHAPTER II – HOW DOES EMOTION AFFECT SOURCE MEMORY? A META-ANALYSIS OF BEHAVIORAL EVIDENCE

Table 1. *Summary of the Moderator Variables Considered in this Study.....* 80

Table 2. *Summary of the Results of the Moderator Analyses for the Effects Based on Valence and Arousal.....* 92

Table 3. *Summary of the Results of the Moderator Analyses for the Effects Based on Arousal....* 97

CHAPTER III – DIFFERENTIAL EFFECTS OF VALENCE AND ENCODING STRATEGY ON INTERNAL SOURCE MEMORY AND JUDGMENTS OF SOURCE: EXPLORING THE PRODUCTION AND THE SELF-REFERENCE EFFECT

Table 1. *Behavioral Performance (Source x Valence) in Experiments 1 and 2.....* 135

CHAPTER IV – IS INTERNAL SOURCE MEMORY RECOGNITION MODULATED BY EMOTIONAL ENCODING CONTEXTS?

Table 1. *Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Source x Valence) Considering Both Experiments.....* 187

CHAPTER V – INTERACTIONS OF EMOTION AND SELF-REFERENCE IN SOURCE MEMORY: AN ERP STUDY

Table 1. *Descriptive Statistics of the Behavioral Performance as a Function of Encoding Task and Valence.....* 231

CHAPTER VI – GENERAL DISCUSSION

Table 1. *Main Results of the Meta-analysis for each Valence Comparison and Encoding Task.....* 262

APPENDIX A – SUPPLEMENTARY MATERIALS - CHAPTER II

Table S1. *Search Expressions Used in Each Database*..... 299

Table S2. *Formulas Used to Compute the Effect Sizes in the Current Meta-analysis*..... 300

Table S3. *Removed Experiments/Conditions after the Influential Outcomes Analysis*..... 301

Table S4. *Results of the Sensitivity Analysis for each Comparison*..... 302

Table S5. *Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based on Valence and Arousal (Organized by Alphabetic Order) – Part I*..... 303

Table S6. *Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based on Valence and Arousal (Organized by Alphabetic Order) – Part II*..... 316

Table S7. *Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based on Arousal (Organized by Alphabetic Order) – Part I*..... 333

Table S8. *Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based Arousal (Organized by Alphabetic Order) – Part II*..... 335

Table S9. *Methodological Quality Assessment of the Studies Included in the Meta-analysis Organized by Alphabetic Order*..... 338

APPENDIX B – SUPPLEMENTARY MATERIALS - CHAPTER III

Table S1. *Affective and Psycholinguistic Properties of the Stimuli Selected for Both Experiments*..... 349

Table S2. *Response Frequencies for Each Valence Condition in Experiment 1*..... 350

Table S3. *Response Frequencies for each Valence Condition in Experiment 2*..... 350

Table S4. *Multinomial Model Parameters for Item and Source Memory for each Experimental Condition (Source x Valence) in Experiments 1 and 2*..... 351

Table S5. *Descriptive Statistics of Confidence Ratings and Response Time (JOSs Ratings, Source Correct Judgments, Correct Rejections) in each Experimental Condition (Task x Emotion) of Experiments 1 and 2*..... 352

Table S6. *Results from the Pair-by-Pair Wilcoxon Tests Applied to the Proportion of Source Incorrect Responses, Do Not Know Responses, and Misses for Experiment 1 and Experiment 2*.. 353

Table S7. *Output of the Model-based Bayesian Analysis (BF_{10}) with JASP in Experiment 1*..... 355

Table S8. *Output of the Model-based Bayesian Analysis (BF_{10}) with JASP in Experiment 2*..... 356

APPENDIX C – SUPPLEMENTARY MATERIALS - CHAPTER IV

Table S1. *Main Characteristics of the Pool of Adjectives and Images Selected for Experiment 1 and Experiment 2*..... 367

Table S2. *Response Frequencies for Each Valence and Source Condition in Experiment 1*..... 368

Table S3. *Response Frequencies for Each Valence and Source Condition in Experiment 2*..... 368

Table S4. *Multinomial Model Parameters for Item and Source Memory for each Experimental Condition (Source x Valence) for Experiment 1 and Experiment 2*..... 369

Table S5. *Results from the Pair by Pair Wilcoxon Tests Applied to the Response Time of Source Correct Responses in Experiment 2*..... 370

Table S6. *Results from the Pair by Pair Wilcoxon Tests Applied to the Proportion of Source Incorrect Responses, Do Not Know Responses, and Misses for Experiment 1 and Experiment 2*..... 371

Table S7. *Means (and SDs) for the PANAS Scores Obtained in Experiment 2*..... 374

Table S8. *Output of the Model-based Bayesian Analysis (BF_{10}) Obtained with JASP for Experiment 1*..... 375

Table S9. *Output of the Model-based Bayesian Analysis (BF_{10}) Obtained with JASP for Experiment 2*..... 376

Table S10. *Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Congruency x Source x Valence) for Experiment 1*..... 377

Table S11. *Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Congruency x Source x Valence) for Experiment 2*..... 377

Table S12. *Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Source x Valence) for Experiment 1 and Experiment 2 Wherein Valence was Recoded According to the Participants' Ratings*..... 378

APPENDIX D – SUPPLEMENTARY MATERIALS - CHAPTER V

Table S1. *Main Characteristics of the Selected Lists of Adjectives*..... 395

Table S2. *Mean Number of Accepted Trials in the ERP Analysis During Study and Test Phases as a Function of Encoding Task and Valence Conditions*..... 396

Table S3. *Response Frequencies for Each Encoding Task and Valence Conditions in the Test Phase*..... 397

Table S4. *Multinomial Model Parameters for Item and Source Memory as a Function of Encoding Task and Valence Conditions*..... 398
Table S5. *Results from the Pair by Pair Wilcoxon Tests Applied to the Proportion of Incorrect Source Responses*..... 399
Table S6. *Mean Amplitude (μV) for the JOS Ratings in the Four Time Windows Analyzed for Middle Frontal (F1/Fz/F2) and Middle Parietal (P1/Pz/P2) ROIs*..... 400

APPENDIX E – SUPPLEMENTARY MATERIALS - CHAPTER VI

Table S1. *Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Source x Valence)*..... 416

CHAPTER I

INTRODUCTION AND THEORETICAL BACKGROUND

"(...) every thought we have, every word we speak, every action we engage in – indeed, our very sense of self and our sense of connectedness to others – we owe to our memory, to the ability of our brains to record and store our experiences. Memory is the glue that binds our mental life, the scaffolding that holds our personal history and that makes it possible to grow and change throughout life."

(Squire & Kandel, 2009)

"Thus, how emotion affects memory is more complicated than a simple statement to the effects that emotional events lead to impairment in memory, or conversely, that they consistently lead to enhancement of memory. (...) people remember different detail information from emotional events than they might remember from similar events that are neutral in character."

(Christianson & Loftus, 1991)

Memory has been puzzling and fascinating humanity across different eras and civilizations. Indeed, records from ancient civilizations offer several accounts of deities and mythological legends that exhorted the importance of memory. For instance, Mnemosyne, the goddess of memory and mother of muses, following the Greek ancient mythology, was evoked before reciting poems to grant full capacity to the reciter and assure no lapses. Mímir, known as "the remember" in the Norse mythology, was the omniscient guardian of a well whose waters contained much wisdom and knowledge. Odin, one of the most revered gods, even pledged his right eye to drink from the well aiming at a glimpse of such knowledge and wisdom. Although brief, these two examples convey the idea that keeping information is an asset, making possible to draw a parallel with our general view that memory is in some way a recorder and repository of information and experiences. In fact, a broader definition regards memory as the capacity to acquire, store, and retrieve information (Anderson, 1995; Baddeley, 1999; Tulving, 1987).

Even though the concept of memory has been addressed by distinct minds and thinkers over several centuries (see Burnham, 1888 for an overview of classic references), the more systematic study of memory only arrived in the early 1880s by the hand of a German philosopher: Hermann Ebbinghaus (1850 – 1909). He was the very first to probe memory using rigorous and innovative experimental methods, leaving an indelible mark in experimental psychology until this day (Anderson, 1995; Baddeley, Eysenck, & Anderson, 2015; Murdock, 1985; Shakow, 1930; Squire & Kandel, 2009). For instance, he devoted much of his work to serial learning, requiring himself to read aloud sequences of non-sense syllables with a consonant-vowel-consonant structure at steady, rapid rates (Baddeley, 1999; Baddeley et al., 2015; Kensinger & Schacter, 2008; Squire & Kandel, 2009). He repeated a sequence until he could pronounce it without errors or hesitation and measured the number of repetitions or time required to do this, and even to relearn the materials after distinct time intervals (Baddeley, 1999; Roediger, 1985). In his experimental endeavor, he tried to control for different factors, such as the influence of meaning and

previous knowledge as illustrated by the creation of pronounceable non-sense syllables, possible bias introduced by the experimenter, the rate of stimulus presentation, and the space and time of the day to keep constant the study conditions as much as possible (Anderson, 1995; Baddeley et al., 2015; Roediger, 1985). Remarkably, his empirical and methodological contributions were derived from extensive observations from a single subject: himself. In fact, as stated by Ramachandran and Blakeslee (1998), some of the great discoveries in the memory field, which paved the way for new experimental and theoretical grounds, were based on evidence from single cases.

In this context, another influential contribution arrived in the 1957 with a publication by the neurosurgeon William Beecher Scoville and the neuropsychologist Brenda Milner (Scoville & Milner, 1957). In this work, the authors reported a series of cases of medial temporal lobe (MTL) resections, including associated results from formal assessments of memory and other intellectual abilities. Among those cases was the young H. M. (1926 – 2008) that shaped the memory research enterprise by showing that much can be learned about memory and brain function in the context of amnesia (see Penfield & Milner, 1958; Rosenbaum et al., 2005; Schacter, 1996; Squire & Wixted, 2011; Tulving, 2002 for other case studies that were also fundamental in this regard). Specifically, H. M. was submitted to a bilateral resection of MTL structures at the age of 27 to alleviate his refractory epilepsy that debilitated his daily living. After the surgical intervention, the frequency and impact of the epileptic seizures was attenuated. The neuropsychological assessment showed the preservation of his general intellectual ability (there was even an increase in the final score), including his perceptual, reasoning, and language abilities (Eichenbaum, 2012; Milner, Squire, & Kandel, 1998; Schacter, 1996). Nonetheless, the surprising observation was that H. M. developed a profound anterograde amnesia accompanied by some degree of retrograde amnesia (he could recall remote memories from his younger years and even some events prior to the surgery; Milner et al., 1998; Scoville & Milner, 1957; Squire & Wixted, 2011). He could no longer form new memories concerning his personal daily experiences, even though he demonstrated attentional and immediate memory skills. H. M. could repeat a string of seven digits, to hold in mind a sequence of numbers by rehearsing them and applying elaborate memory strategies for 15 minutes, as well as to keep a fluid conversation, and this occurred as long as no distractions emerged (Annese et al., 2014; Diana & Wang, 2018; Eichenbaum, 2012; Milner et al., 1998; Squire, 2009). It was also discovered that, even though H. M. could no longer remember the circumstances surrounding specific learning episodes, his performance in specific tasks such as visuo-motor learning and priming tasks did not differ or was even better than controls (Rosenbaum et al., 2005; Schacter, 1996; Squire, 2009). Thus, there were aspects of his memory that were intact.

These observations hinted and even supported specific ideas regarding the organization of memory and its underlying brain mechanisms (Eichenbaum, 2012; Kensinger & Giovanello, 2006; Schacter, 1996; Squire, 2009; Squire & Kandel, 2009; Squire & Wixted, 2011): memory is a distinct cognitive ability that can be differentiated from other abilities such as perception, language, or reasoning; the H. M.'s amnesia impaired the transition of information from short-term memory into a more stable and durable memory trace (i.e., long-term memory), supporting the distinction between short-term and long-term memory (a distinction already implied in the writings of William James, 1890); not all memory-related activities were affected as H. M. was able to show learning and retention through his performance in certain tasks, indicating that memory could be composed of different systems; the MTL structures play an important role in memory functioning, even though these structures did not seem to be the last repository of information as H. M. did not lost all memories prior to the surgery, especially the remote ones.

Taken together, the H. M. case was fundamental to pursue the idea that memory is not a single, unitary, or monolithic ability. Even though not fully consensual (Baddeley, 1999; Tulving, 2002), the view that long-term memory is organized in multiple systems encompassing different brain mechanisms has been amassing favorable evidence for several decades, proving its heuristic value to the memory field (Baddeley et al., 2015; Squire, 1992, 2004). Following the classification purposed by Squire and colleagues (Squire, 1992, 2004; Squire & Knowlton, 1995; Squire, Knowlton, & Musen, 1993; Squire & Zola, 1996; but see Tulving, 1987 for other possible classifications), a major distinction lies between declarative (explicit) and nondeclarative (implicit) memory. The first refers to our ability to recollect and recognize facts and events in a conscious manner, that is, to intentionally bring to mind representations which seem to depend upon the good functioning of medial temporal (hippocampus, entorhinal, perirhinal, parahippocampal regions) and diencephalic structures (Baddeley et al., 2015; Milner et al., 1998; Squire, 2009; Squire & Knowlton, 1995; Squire & Zola-Morgan, 1988; Squire & Zola, 1996). In its turn, the second comprises all other expressions of memory that do not require a conscious reminiscence or recognition. The manifestation of learning and memory is rather expressed gradually through changes in performance, skill acquisition or response to stimuli, being even difficult to provide a conscious report of such changes (Baddeley et al., 2015; Squire, 1992, 2009; Squire & Wixted, 2011; Squire & Zola-Morgan, 1988; Tulving, 1987). Examples are the acquisition of perceptomotor skills, simple forms of classical conditioning, sensitization, habituation, and priming, whose operations seem to be supported by other brain regions, such as the neocortex, amygdala, cerebellum, and basal ganglia (Milner et al., 1998; Squire, 1992, 2009; Squire & Wixted, 2011; Squire & Zola, 1996).

This short overview of the multiple memory systems is important to situate, in terms of general memory theory, the research developed in the current dissertation. In this regard, within the declarative memory, the distinction between semantic and episodic memory is also relevant. Semantic memory is deemed as the memory for our general knowledge of the world that allows to evoke representations of objects, concrete and abstract concepts, facts, situations, events, social norms, among others, without the need to reminisce about details of the learning episodes from which the representations evolved (Baddeley et al., 2015;Tulving, 1987, 1993, 1995, 2001). Episodic memory, in its turn, is tied to personal learning episodes in the sense that allows to re-experience portions of our past experiences (Squire & Zola-Morgan, 1988; Tulving, 1985, 1987). To better illustrate this distinction, let us imagine that someone challenges us to remember all the information we can about "Pluto". We could say that we know that Pluto was one of the planets of our solar system, discovered in 1930, being the furthest away from the sun. Although we likely learnt this information as a child during our formal education, in later years (specifically in 2006), "Pluto" was no longer deemed as planet by the astronomical scientific community, being a matter of debate even nowadays. We could say that "Pluto" was an ancient Greek god which did not roam the skies but rather the underworld. "Pluto" could also allude to one peculiar dog, the Disney character. Even though we could recall different pieces of information associated with the name "Pluto", this does not mean that we are able to recall specific personal episodes related to the acquisition of such information. Thus, semantic memory was the main driver of information in this situation. In contrast, there are other pieces of information for which we can recollect associated personal experiences. Using a personal example, I remember one of my undergraduate classes on memory wherein the work of the British psychologist Frederic Bartlett (1886 – 1969) was mentioned. The professor asked who from the class had the opportunity to read "The War of the Ghosts" that he had provided as study material. I raised my hand timidly as I thought that I was not the only one. Apparently, I was the only raising the hand, so the professor used me as a participant to reminisce the story in front of the all class, as a simple demonstration of what Bartlett had done in his experiments. I remember that I was so nervous and feeling so unsure about what that I was recollecting. As expected, I even changed some of the elements of the story. Based on semantic memory, I know that Bartlett intended to study memory in a more naturalistic fashion using for example traditional tales, being one of the first arguing in favor of the reconstructive nature of memory (Baddeley, 1999; Baddeley et al., 2015; Johnson, 1997; Schacter & Addis, 2007; Squire & Kandel, 2009), but based on episodic memory I was able to re-experience and tell something from personal past also related to Bartlett. Thus, episodic memory implies the retrieval of contextual information (e.g., where; when; how; who) from the original episode (Diana & Wang, 2018; Tulving,

2001). As we will see in the following sections, this notion is fundamental to circumscribe the theoretical ground wherein the experimental work of this dissertation was developed.

Source memory: The nuts and bolts

Source memory has been conceptualized as the ability to retrieve, identify and make decisions about details/features/qualities that compose a given episode, thus allowing to specify its origins (Johnson, Hashtroudi, & Lindsay, 1993). Some authors argue that source memory contributes for the episodic character of declarative memory (Johnson, 1997; Johnson et al., 1993; Jurica & Shimamura, 1999; Squire & Kandel, 2009). For this reason, the study of source memory has been a fruitful research avenue to examine some of the processes that substantiate episodic recollection. More specifically, Marcia Johnson and colleagues (Johnson, 1997, 2006; Johnson, Bush, & Mitchell, 1998; Johnson et al., 1993; Johnson & Raye, 1981) proposed the source monitoring framework, which posits that when people are required to make judgments about the origins of an event, such as when they try to remember if a certain event occurred in reality or in a dream, they engage in heuristic and deliberative monitoring processes to reach a decision. These processes may include the evaluation of different qualities bound to the episode, such as perceptual (e.g., visual or auditory modality), contextual (e.g., when and where), affective (e.g., emotional reactions experienced during the event) or cognitive features (e.g., imagery, organization, elaboration strategies associated with the event), as well as an assessment of the information retrieved given previous knowledge (e.g., beliefs; schemas; stereotypes) and motivational goals (Johnson, 1997, 2006; Mather, Johnson, & De Leonardis, 1999; Raj & Bell, 2010). Other factors such as task demands and social context may also modulate source monitoring processes (e.g., in formal contexts, such as in professional ones, deliberative and stringent processes are more likely to be employed, whereas in informal contexts, such as a conversation with a friend, the engagement of heuristic and automatic processes is more likely; Johnson, 1997). The heuristic value of this framework has been immense. In fact, it has been applied to multiple research domains such as: eyewitness testimony (see Baddeley, 1999; Schacter, 1996 for overviews); inadvertent plagiarism also known as cryptomnesia (e.g., Gingerich & Sullivan, 2013; Macrae, Bodenhausen, & Calvini, 1999); auditory hallucinations in schizophrenia as they might be partially accounted for source memory biases, such as a tendency to attribute self-generated information to external sources (e.g., Brookwell, Bentall, & Varese, 2013; Fisher, McCoy, Poole, & Vinogradov, 2008; Keefe, Arnold, Bayen, McEvoy, & Wilson, 2002).

More specifically, this framework proposes distinct subsets of source monitoring processes that are engaged to discriminate between elements whose origin is outside (i.e., external to) the subject and

elements that are generated by (i.e., internal to) the subject (Johnson et al., 1993). Specifically, three subsets of source monitoring processes had been described: external source monitoring that allows to distinguish between features external to the subject (e.g., did I hear this news on the radio or on the television?); internal source monitoring that allows to discriminate between features that are generated by the subject (e.g., did I write this message or imagine that I did?); internal-external source monitoring, also known as reality monitoring, that makes possible to separate between externally- and internally-derived sources (e.g., did I say this or someone else did?). Other notions about source memory functioning were also drawn from this classification. Specifically, whereas external source memories are more likely to contain perceptual, spatiotemporal, semantic and affective details, internal source memories are more likely to contain details about cognitive operations that took place during the study episode (e.g., verbal thoughts; mental images; strategies to organize the information; associations; Ferguson, Hastroudi, & Johnson, 1992; Finke, Johnson, & Shyi, 1988; Johnson, 1988, 1997, 2006; Johnson & Raye, 1981; Raye & Johnson, 1980). Thus, when making source memory attributions if we recall more perceptual details about a given episode, we are more likely to deem the source as external to us; on the contrary, if the details are mainly mental activities, the source is likely to be perceived as internally-derived (Raye & Johnson, 1980). Furthermore, it has been suggested that both external and internal source memory decisions are more difficult than reality monitoring decisions as they might incorporate more overlapping and similar features, thus resulting in great likelihood of committing source misattributions (El Haj, Fasotti, & Allain, 2012; Johnson, 1997; Johnson & Foley, 1984; Raye & Johnson, 1980).

In terms of laboratory research, one common approach to study source memory is to manipulate two or more features during the encoding/study phase, followed by a retrieval/test phase in which the memory for such features is probed. Let us imagine the manipulation of two features corresponding to source A and source B, respectively (examples for external, internal, and reality monitoring are summarized on Table 1). This means that during encoding each stimulus (e.g., word; picture; face; sound) is related in some way with one of these two sources. During test, participants are then required to decide for each stimulus if it was presented in the context of source A or source B.

It is worth noting that the relationship between stimulus and source is not established in the same way as in typical associative memory paradigms, in which unique groups of stimuli are studied and tested. Two important distinctions between source memory and associative memory tasks should be considered: whereas in source memory tasks the same source is related to distinct stimuli (one-to-many mapping), in associative memory tasks each stimulus is particularly associated with other unique stimulus (one-to-one mapping; Hou, Grilli, & Glisky, 2019a, 2019b); due to the one-to-many mapping, the source variation

across learning trials is likely to be perceived as secondary, meaning that participants are more focused on the different stimuli rather than on source variations (Diana & Wang, 2018). This focus on the stimulus (*vs.* source), i.e., on the central (*vs.* peripheral) or on the content (*vs.* context), emphasizes other relevant distinction in this type of paradigms, which is between source and item memory. In this context, item memory refers to the ability to retrieve or recognize that a stimulus was previously encountered, whereas source memory requires the retrieval of features that accompanied this same stimulus during its acquisition (Raj & Bell, 2010). In more naturalistic settings, what is considered central/content in one moment can be peripheral/contextual in another, and this change is likely to be driven by the subjects' goals and agendas (Johnson, 1992). Although this is less problematic in laboratorial studies as the distinction between item and source is established experimentally, this poses some ecological concerns that should be acknowledged by experimental research. In a similar fashion, typical source memory tasks tend to manipulate and test specific stimulus features (see Table 1), meaning that although we might not be able to confidently decide between sources A or B for a certain stimulus, we might recall other details composing the stimulus study event. In other words, we might recover other source information that, at the time of test, do not inform the decision between source A or B, but that support the reconstruction of the original episode. Thus, experimental tasks might underestimate source memory performance. Even considering these limitations, experimental investigations on source memory have contributed not only to a better understanding of episodic recollection, but also to the identification of factors that modulate source memory functioning as we will exemplify in the following section.

When source memory is tinged with emotion: How emotion and self-referential encoding affect our memory for source features?

The interplay between emotion and memory has been a research topic gardening much interest and multiple ramifications over the years (e.g., Dolcos et al., 2017; Dolcos, Wang, & Mather, 2014; Levine & Pizarro, 2004; Tyng, Amin, Saad, & Malik, 2017; Uttl, Siegenthaler, & Ohta, 2006). A robust body of evidence had demonstrated that emotionally charged events can shape differently our memory experiences. Quintessential examples are flashbulb memories showing that an emotionally significant event can be recollected in a vivid and detailed manner (e.g., Brown & Kulik, 1977; Hirst & Phelps, 2016), as well as memory losses associated with traumatic events, demonstrating that emotionally intense events may also exert a deleterious effect on memory (e.g., Christianson & Loftus, 1987; Klein, Caspi, & Gil, 2003; Radulovic, Lee, & Ortony, 2018; Staniloiu & Markowitsch, 2014). Even though these represent

Table 1

Examples of Features Commonly Manipulated in the Context of Source Memory Tasks

Type of SM task (Johnson et al., 1993)	Examples
External SM	• Stimulus color (e.g., font color: yellow *vs*. blue) • Stimulus background (e.g., color of the border surrounding the stimuli: yellow *vs*. blue) • Spatial location (e.g., side of the screen in which the stimuli are presented: left *vs*. right) • Temporal order (e.g., list in which the stimuli are presented: list 1 *vs*. list 2) • Identity of the speaker (e.g., female *vs*. male)
Internal SM	• Encoding task (e.g., self-reference *vs*. common judgment task)
Reality monitoring	• Modality of presentation (e.g., seen *vs*. imagined) • Identity of who generated the information (e.g., self- *vs*. other-generated)

Note. SM = Source Memory.

more extreme instances of the effects of emotion on memory, we deal with a diverse spectrum of emotional information on a daily basis, implying that emotion is a prominent player in understanding memory functioning. Furthermore, the study of memory in these naturalistic situations, although of incommensurable value, may also have some disadvantages as it is difficult to measure the accuracy of self-reported and personal memories (Reisberg & Heuer, 2004). Thus, together with the study of memory in real world situations, it is also paramount to pursue emotion-related effects in a more rigorous and systematic manner in laboratorial settings, even acknowledging that more intense emotions are unlikely to be induced in such contexts (Schacter, 1996).

The concept of emotion *per se* has been intensely debated by the scientific community. Nonetheless, the answer to the question "What is in an emotion?" posed by William James (1884) more than one hundred years ago is far from being settled (Frijda, 2008; Russell, 2003; Sander, 2013). Generally, emotion has been used as an umbrella term being related to other phenomena such as affect, mood, feeling, sentiment, motivation, among others (Sander, 2013). Nevertheless, there seems to be some agreement that emotion *per se* is a multicomponent phenomenon that entails temporary and fast alterations in the neurophysiology, cognition, and behavior of an organism in response to internal or external events (Dolcos et al., 2020; Gardhouse & Anderson, 2013; Moors, 2009; Okon-Singer, Hendler, Pessoa, & Shackman, 2015; Sander, 2013). The challenges of studying emotion are also reflected in the variety of theoretical frameworks that have been proposed (see Moors, 2009; Sander, 2013 for

overviews). Major examples are: the categorical models (e.g., emotion can be separated in basic, universal, discrete emotional states such as joy, surprise, fear, sadness, anger, disgust, each one with unique behavioral and physiological manifestations, shaped by evolutionary forces, and supported by distinct brain mechanisms; Ekman, 1992a, 1992b); the appraisal models (e.g., emotions emerge and are modulated by the relevance that one attributes to certain environmental events, which might help or hinder the achievement of active goals and needs; Levine & Pizarro, 2006; Moors, Ellsworth, Scherer, & Frijda, 2013); the dimensional models (e.g., from a bidimensional perspective, arousal and valence are deemed as the building blocks of our affective experiences; Lang, Greenwald, Bradley, & Hamm, 1993; Posner, Russell, & Peterson, 2005; Russell, 1980, 2003; Russell & Pratt, 1980).

The dimensional models have been the most used to probe the emotion-memory dynamic relationship in laboratorial settings. More specifically, different emotions are represented in a circular dimensional space, resulting from the combination of varying degrees of valence and arousal (Posner et al., 2005; Russell & Pratt, 1980). Whereas the valence dimension refers to the pleasantness, ranging from negative/unpleasant to positive/pleasant, the arousal dimension represents the activity level, ranging from low/calm to high/excited. The questions of how these dimensions relate to one another and whether negative/positive valence or low/high arousal might be placed within the same continuum are still under debate (Kuppens, Tuerlinckx, Russell, & Barrett, 2013; Kuppens et al., 2017; Niedenthal, 2008; Petrolini & Viola, 2020). For instance, whereas some authors advance with the orthogonality and bipolarity of the valence and arousal dimensions (e.g., Russell, 1980; Russell & Pratt, 1980), others admit some degree of association between the two dimensions (e.g., Lang & Bradley, 2013; Lang et al., 1993). Such association is supported by evidence showing a small yet consistent "U", "V", or "boomerang" shape relationship between arousal and valence, that is, the more positive or negative a certain stimulus is perceived, the more arousing it is (Kuppens et al., 2013, 2017). Furthermore, the dimensional approach have been anchored to the development and the widespread use of diverse stimulus databases, such as the Affective Norms for English Words (ANEW; Bradley & Lang, 1999; European Portuguese – EP version: Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), the International Affective Digitized Sounds (IADS-2; Bradley & Lang, 2007; EP version: Soares et al., 2013), and the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008; EP version: Soares et al., 2015). Specifically, one of the most fruitful approaches to probe the effects of emotion on memory has relied on the contrast between emotional stimuli with distinct affective properties, such as negative and positive high arousing stimuli *vs.* neutral low arousal stimuli. Thus, in the current dissertation, whenever the expression

"emotional stimuli" or similar is used, it is assumed that, according to the dimensional approach, the stimuli induce in the participants emotions with certain hedonic and arousal characteristics.

In this context, when testing item memory, it has been demonstrated that emotional stimuli appear to be in advantage in relation to neutral ones considering different pictorial (e.g., D'Argembeau & Van der Linden, 2005; Kensinger, Addis, & Atapattu, 2011; Nashiro & Mather, 2011; Rimmele, Davachi, Petrov, Dougal, & Phelps, 2011; Schmidt, Patnaik, & Kensinger, 2011; Yick, Buratto, & Schaefer, 2015) and verbal materials (single words: e.g., Doerksen & Shimamura, 2001; Dougal, Phelps, & Davachi, 2007; Le Bigot, Knutsen, & Gil, 2018; Maddock & Frein, 2009; Robin, Cébron, Letellier, & Nizard, 2018; word pairs: e.g., Maddox, Naveh-Benjamin, Old, & Kilb, 2012; sentences: e.g., Davidson, McFarland, & Glisky, 2006, Experiment 3). This advantage, also designated as emotion-enhanced memory (EEM) effect, is observed in terms of accuracy, as well as in terms of subjective vividness of recollection of details from the study episode (e.g., Kensinger, 2007; Kensinger & Corkin, 2003; Ritchey, Wang, Yonelinas, & Ranganath, 2019). Several non-exclusive encoding and post-encoding mechanisms have been advanced to explain the EEM effect (Gutchess & Kensinger, 2018; Hamann, 2001). Specifically, during encoding emotional stimuli appear to engage greater attentional, elaborative, and rehearsal resources (Kensinger & Corkin, 2003; Reisberg, 2006; Reisberg & Heuer, 2004). Also, they appear as more distinct and salient when mixed with non-emotional stimuli, and are more prone to be organized into specific semantic categories (Hunt & McDaniel, 1993; Schmidt & Saari, 2007; Talmi, 2013; Talmi, Luk, McGarry, & Moscovitch, 2007; Talmi & McGarry, 2012; Watts, Buratto, Brotherhood, Barnacle, & Schaefer, 2014), as supported by the observation that emotional stimuli are more likely to be falsely recalled (e.g., Barnacle, Tsivilis, Schaefer, & Talmi, 2018; Davidson et al., 2006; Johansson, Mecklinger, & Treese, 2004; Langeslag & Van Strien, 2008; Maratos, Allan, & Rugg, 2000). Emotional stimuli are also benefited in terms of consolidation, reflected in an enhanced EEM effect at long *vs.* short retention intervals (Quevedo et al., 2003; Schümann, Bayer, Talmi, & Sommer, 2018; Sharot & Yonelinas, 2008; Waring & Kensinger, 2009).

Even though the extant evidence suggests a mnemonic advantage of emotional (*vs.* neutral) information in the case of item memory, the effect of emotion on memory might be regarded as a "double-edge sword" (Dolcos et al., 2014). It seems that memory enhancing effects are more likely to occur for content or central aspects, whereas the memory for peripheral or contextual elements tends to be hindered by the presence of emotion, despite discrepant evidence in this regard (see Reisberg, 2006; Reisberg & Heuer, 2004 for overviews). This bring us back to the importance of studying source memory in emotional conditions as it might contribute to clarify how emotional information shapes the objective

memory performance for features that make part of the study episode. Nonetheless, as we will present in different instances of the current dissertation (see Chapter II and Appendix included in Chapter III), studies focusing on the effects of emotional stimuli on source memory performance have shown mixed findings, including enhancements, impairments, and lack of effects (see also Table 2 for a summary). In this context, one important step is to raise and test hypotheses that might explain the conflicting results considering not only theoretical predictions, but also the identification of factors that may influence the emotion-source memory relationship. Different theoretical proposals have been adopted in this regard, such as the Yerkes-Dodson law (Yerkes & Dodson, 1908), the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959), the priority-binding mechanism (Hadley & MacKay, 2006; MacKay et al., 2004), the object-based binding theory (Mather, 2007), and the arousal-biased competition theory (Mather & Sutherland, 2011). These proposals will be addressed in more detail in Chapter II. Moreover, different factors associated with study materials and study-test conditions (e.g., type and modality of the stimuli; encoding intentionality and strategies; study-test delay interval; type of memory test), as well as individual characteristics (e.g., age; personality) might modulate source memory performance (e.g., Cansino et al., 2019; Murphy & Isaacowitz, 2008; Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995; see also Chapter II). Nonetheless, the systematic study of these factors in the domain of source memory is still scarce, especially in emotional conditions. It is in this context that the current dissertation intends to bring some contributions, particularly considering how distinct encoding tasks might influence the effects of emotional information on objective source memory accuracy, as well as on subjective metamemory judgments.

In this fashion, the encoding task appears to be a relevant candidate to consider (see Leynes & Crawford, 2018; Leynes & Mok, 2017 for a similar idea) as illustrated by the classical Levels of Processing framework, which claims that information encoded based on meaning leads to better memory performance than information encoded based on perceptual features (Craik & Lockhart, 1972; Craik & Tulving, 1975; see Yonelinas, 2002). Indeed, the brain processes mobilized during memory formation (Padovani, Koenig, Brandeis, & Perrig, 2011; Park, Uncapher, & Rugg, 2008) and memory recognition (Peters & Daum, 2009; Senkfor, Van Petten, & Kutas, 2008) vary with the requirements of the encoding task. In the case of source memory, prior evidence also indicated that a better source memory performance was found when participants were asked to perform affect *vs.* size ratings (Mathews, Ridgeway, & Holmes, 2013, Experiment 3).

Table 2

Summary of the Effects of Emotional Stimuli on Source Memory Accuracy Considering the Type of Source Memory Task and of Features

Type of SM task (Johnson et al., 1993)/Type of SM feaure	Emo > Neu	Emo = Neu	Emo < Neu	Supporting studies
External SM				
Stimulus color	✓	✓	✓	D'Argembeau & Van der Linden, 2004; Davidson et al., 2006; Doerksen & Shimamura, 2001; Dougal et al., 2007; Kensinger & Corkin, 2003; Rimmele, Davachi, & Phelps, 2012; Wang, 2012; Wang & Fu, 2011
Stimulus background	✓	✗	✓	Doerksen & Shimamura, 2001; Jia, Gao, Cui, & Guo, 2018; Kaida, Niki, & Born, 2015; MacKenzie, Powell, & Donaldson, 2015; W. B. Mao, An, & Yang, 2017; Rimmele et al., 2011
Spatial location	✓	✓	✗	Borg, Leroy, Favre, Laurent, & Thomas-Antérion, 2011; A. D'Argembeau & Van der Linden, 2004; Koenig & Mecklinger, 2008; Novak & Mather, 2009; Rimmele et al., 2012; Schmidt et al., 2011
Temporal order	✓	✓	✓	Aupée, 2007; Koenig & Mecklinger, 2008; Minor & Herzmann, 2019; Rimmele et al., 2012; Schmidt et al., 2011; Yick et al., 2015
Identity of the speaker	✓	✗	✗	Davidson et al., 2006
Internal SM				
Encoding task	✗	✓	✓	Blix & Brennen, 2011; Ferré, Comesaña, & Guasch, 2019; Gallant, Spaniol, & Yang, 2019; Gallant & Yang, 2014; Kensinger & Schacter, 2006a; X. Mao, You, Li, & Guo, 2015; Otani, Jaffa, Libkuman, Goernert, & Kato, 2012; Otani, Libkuman, et al., 2012; Pereira, Sampaio, & Pinheiro, 2019; Ritchey et al., 2019; Sharot & Yonelinas, 2008
Reality monitoring				
Modality of presentation	✓	✓	✓	Cook, Hicks, & Marsh, 2007; Dutton & Carroll, 2001; Kensinger, O'Brien, Swanberg, Garoff-Eaton, & Schacter, 2007; Kensinger & Schacter, 2005a, 2006b; Robin et al., 2018
Identity of who generated the information	✓	✓	✓	Diaz-Asper, Malley, Genderson, Apud, & Elvevåg, 2008; Larøi, Linden, & Marczewski, 2004; Le Bigot et al., 2018; McKague, McAnally, Skovron, Bendall, & Jackson, 2012; Morrison & Haddock, 1997

Note. Emo = Emotional; Neu = Neutral; SM = Source Memory; ✓ = Evidence in favor; ✗ = No evidence found.

Accordingly, some authors have argued that the encoding strategy influence the degree of integration between item and contextual details and may, therefore, contribute for the enhancement, impairment, or

null findings documented in source memory studies using emotional information (Chiu, Dolcos, Gonsalves, & Cohen, 2013; Riegel, Wierzba, Grabowska, Jednoróg, & Marchewka, 2016).

More specifically, in the current dissertation, we intended to probe encoding tasks typically known to boost memory performance, but that were less studied in the context of source memory. Indeed, the implementation of specific encoding strategies such as interactive imagery, sentence generation (e.g., Kuhlmann & Touron, 2012, 2017), semantic clustering (e.g., Wegesin, Jacobs, Zubin, Ventura, & Stern, 2000), and unitization (i.e., relate in a meaningful way, in a single unit item and contextual details; e.g., El Haj, Kessels, & Allain, 2016; Tu & Diana, 2016; Tu, Alty, & Diana, 2017) had shown beneficial effects on source memory. Here, we focused particularly on self-referential encoding, even though the production effect (i.e., the memory advantage observed for information read aloud *vs.* read silently; Bodner & Taikh, 2012; Dodson & Schacter, 2001; C. M. MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Ozubko & MacLeod, 2010) was also examined in one of the experiments (see Chapter III, Experiment 1). The self-reference effect has been described as a memory advantage for information that is associated with the self during encoding is comparison with other encoding tasks (e.g., other-self-referent; semantic; perceptual; phonemic; Conway, Dewhurst, Pearson, & Sapute, 2001; D'Argembeau, Comblain, & Van der Linden, 2005; Derry & Kuiper, 1981; Dewhurst & Conway, 1995; Dewhurst, Anderson, Grace, & Boland, 2017; Fossati et al., 2004; Kelley et al., 2002; Yaoi, Osaka, & Osaka, 2015; Yin, Ma, Xu, & Yang, 2019). This advantage has been supported by extensive evidence coming from laboratorial, clinical and daily settings, being corroborated across the life-span (Derry & Kuiper, 1981; Glisky & Marquine, 2009; Grilli & McFarland, 2011; Kuiper & Rogers, 1979; Symons & Johnson, 1997).

The study of self-reference in emotional conditions may be of particular interest as some authors argue that the memory benefits associated with this encoding strategy may be optimal when the information is emotional (Miall, 1986). In fact, more recently, Gutchess and Kensinger (2018) proposed that self-referential and emotional information rely on similar cognitive mechanisms during encoding that result in a greater likelihood of remembering the information. These mechanisms include attentional prioritization, elaboration, and rehearsal, which are more likely to be mobilized for both self-referential and emotional information *vs.* non-self-referential and non-emotional. The authors also argue in favor of converging neural mechanisms, namely greater activation of the medial prefrontal cortex (mPFC) and greater amplitude of the late positive complex (LPC) of the event-related potential response derived from the electroencephalogram. Also, prior evidence had suggested that the way information is encoded, such as when using a self-reference task, plays a role in how emotional stimuli are processed in the sense that the degree of emotionality effects may be heightened or attenuated as a function of the encoding task

(e.g., Delaney-Busch, Wilkie, & Kuperberg, 2016; Fischler & Bradley, 2006; Ritchey, Bessette-Symons, Hayes, & Cabeza, 2012; Zhou et al., 2017). In a similar fashion, emotional information interferes with the degree of self-reference as in typical conditions one is more prone to accept positive information as being self-related when compared to negative information (e.g., Derry & Kuiper, 1981; Lewis, Critchley, Rotshtein, & Dolan, 2007; Moran, Macrae, Heatherton, Wyland, & Kelley, 2006; Pauly, Finkelmeyer, Schneider, & Habel, 2013).

Thus, it is plausible to assume that the effect of emotional stimuli on source memory might also differ as a function of the encoding task and that the interaction between self-reference and emotion might shape source memory performance and underlying brain responses in specific ways. Whereas prior evidence had demonstrated the benefits of self-referential encoding in the case of source memory (Dulas, Newsome, & Duarte, 2011; Durbin, Mitchell, & Johnson, 2017; Hou et al., 2019b; Leshikar, Dulas, & Duarte, 2015; X. Mao, Wang, Wu, & Guo, 2017; Pereira et al., 2019; Yin et al., 2019; Zhang, Pan, Li, & Guo, 2018), the use of emotional stimuli together with self-referential processing has received less attention, not only in terms of source memory performance, but also with respect to brain mechanisms and subjective metamemory judgments (as we intend to convey in the following sections).

Delving into prospective metamemory judgments of source: How does one perceive the effects of emotion on future source memory performance?

Measuring source memory performance (e.g., accuracy) in the context of experimental tasks with embedded emotionally charged stimuli can be deemed as an objective measure of source memory. Additionally, other way to examine how emotion and encoding tasks may influence source memory is to use subjective measures such as asking the participants to assess the memorability of the information while they are studying it. This subjective approach, somewhat based on the expectations, knowledge, and predictions of the individuals, has been used in the domain of metamemory research. Specifically, metamemory refers to beliefs and conceptions each one of us has concerning how memory works (Flavell, 1979). Metamemory is essential to guide our daily behavior as it helps us to monitor and implement strategies that maximize information encoding and mitigate memory failures. For example, if we usually forget appointments and we are aware of the impact of these lapses, we are more likely to engage in external or internal compensatory strategies such as written reminders or mnemonics, respectively, to assure we do not miss appointments. Similar examples come from patients with neurological disorders such as amnesia (e.g., Pannu & Kaszniak, 2005; Schacter, 1996), as well as from educational settings (e.g., Dunlosky & Rawson, 2012; Rhodes, 2016; Soderstrom, Yue, & Bjork, 2016). In the first case, as

amnesic patients display different degrees of awareness regarding their memory deficits, the study and development of interventions targeting metamemory are critical to help these patients to circumvent their daily memory difficulties (Schacter, 1996). In the second case, the better the students monitor and predict the successful acquisition of the study materials, the more likely they are to optimize their study plan and assure a good academic performance (Soderstrom et al., 2016).

In metamemory research, one common approach is to ask participants to perform prospective or retrospective metamemory judgments (Dunlosky, Mueller, & Thiede, 2016; Nelson & Narens, 1990; Rhodes, 2016). Whereas in the first case participants are required to evaluate how likely they are to remember/forget a given information in a later memory test, in the second case participants assess the memorability of information during its retrieval. In the current dissertation, we focused on prospective metamemory judgments as measured by judgments of source (JOSs), even though there are other types of prospective judgments, such as ease-of-learning (EOL) judgments (i.e., individuals are invited to predict if a given item will be easy or difficult to acquire) and judgments of learning (JOLs; i.e., individuals are required to evaluate in a rating or numeric scale the memorability of specific items; Dunlosky et al., 2016; Fairfield, Mammarella, Palumbo, & Di Domenico, 2015; Pannu & Kaszniak, 2005). Indeed, JOSs are similar to JOLs in the sense they provide a forecast concerning future memory performance during encoding, but instead of focusing on the memorability of items in general, they entail the memorability and discriminability for specific source features. Also, JOSs have received less attention when compared to JOLs. In this regard, extensive research had demonstrated that JOLs are sensitive to different characteristics of the study materials, including associative relatedness and font-size (e.g., Luna, Martín-Luengo, & Albuquerque, 2018; Undorf & Erdfelder, 2015; see Dunlosky, Mueller, & Tauber, 2015 and Rhodes, 2016 for overviews). Thus, the emotional properties of the stimuli are no exception, even though this topic has amassed limited evidence (see Fairfield et al., 2015 for a brief review). Prior studies with JOLs had already shown that emotional stimuli tend to be perceived as more memorable than neutral ones (words: Gallant et al., 2019, Experiment 1; Hourihan, Fraundorf, & Benjamin, 2017; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010; pictures: Hourihan & Bursey, 2017; Tauber, Dunlosky, Urry, & Opitz, 2017; faces: Nomi, Rhodes, & Cleary, 2013; Witherby & Tauber, 2018).

Moreover, following the cue-utilization hypothesis (Koriat, 1997), which proposed a classification for cues supporting subjective metamemory decisions, the emotional properties of the stimuli can be deemed as intrinsic cues. In other words, intrinsic cues refer to inherent stimulus characteristics that inform and contribute to JOLs. This perspective also proposes other types of cues, namely extrinsic (i.e., task-related conditions during encoding or retrieval) and mnemonic cues (i.e., individual experiences with

a specific item). Accordingly, whereas the emotional quality of the stimuli may function as an intrinsic cue, the encoding task may be considered as an extrinsic cue. Even though there is some degree of convergence among cues and conditions that modulate both memory and metamemory, this is not always the case (see Rhodes, 2016). As such, the effects of multiple cues such as emotion and encoding task might differ between subjective and objective memory measures. Likewise, it is not clear whether the same cues influence JOLs and JOSs in similar ways, so the effects of emotional information on JOLs described here might differ in the case of JOSs. Thus, more research is needed to unravel how multiple cues such as emotion and encoding task affect JOSs.

In fact, when considering the effects of emotional information and encoding strategies on JOSs, few studies were conducted so far and two of them are included in this dissertation (see Chapter III and Appendix D). Carroll and colleagues (Carroll, Mazzoni, Andrews, & Pocock, 1999) devised one of the first studies focusing on the effects of different encoding strategies (seen *vs.* imagined; self- *vs.* other-generated actions) on JOSs. Two years later, they also examined the effects of three different arousing conditions, showing that highly arousing conditions were accompanied by an agreement between JOLs and JOSs predictions and objective memory performance (Dutton & Carroll, 2001). Thus, even though source memory is the focus of this dissertation, one of its the novel contributions certainly pertains to JOSs.

Neural correlates of source memory

The history of memory research is a vivid example of how the study of brain-behavior synergies has shaped an entire field, as conveyed early in this introduction. With the advent of different brain imaging techniques, such as the electroencephalography (EEG), the magnetoencephalography (MEG), the positron emission tomography (PET) the functional magnetic resonance imaging (fMRI), the transcranial magnetic stimulation (TMS), among others, new waves of research emerged offering innovative ways to examine and advance our knowledge on brain, behavior, and cognition. Even though these diverse techniques have unique specificities and characteristics (see Baddeley et al., 2015, p. 37, for a summary), they are complementary in many ways and have allowed researchers to probe brain responses while people engage in tasks tapping, for example, perception, attention, motor skills, language, executive functions, and of course memory, among many other abilities (Bunge & Kahn, 2009; Schacter, 1996). Particularly, in this dissertation, we focus on two of the most used non-invasive techniques: the EEG/Event-related potential (ERP) technique and the fMRI. Nonetheless, before delving into more details concerning the neural correlates of source memory, it is important to present how these brain signals are obtained.

The EEG is a method of registering brain-derived activity by means of electrodes placed on the head's surface. It was applied for the first time in a series of experiments conducted by Hans Berger (1873 – 1941) between 1910 and 1929 (Cohen, 2014; Keil, 2013; Niedermeyer, 2005). At this time, researchers pioneered the investigation of spontaneous oscillatory activity emerging from the recorded signal (i.e., the electroencephalogram), but soon the possibility of exploring this signal in response to specific events was accomplished and progressed over several decades (Fabiani, Gratton, & Federmeier, 2007). Particularly, the ERPs correspond to positive and negative voltage deflections that result from segmenting and averaging portions of the electroencephalogram time-locked to specific external (e.g., stimulus presentation) or internal events (e.g., preparation of a behavioral response; Johnson, 1997; Kappenman & Luck, 2012, 2016; Keil, 2013). For each event, the waveform is composed not only of brain activity related to the event processing, but also from unrelated activity including spontaneous/background EEG activity not time-locked to the event, extracranial electrical potentials (e.g., muscles; skin; heart), and even electrical activity from surrounding devices which is capture by the EEG equipment (Keil, 2013; Luck, 2014). In this context, averaging across several repetitions of the same event type is a fundamental step to increase the activity of interest while minimizing other types of activity occurring concomitantly yet randomly with the events (Fabiani et al., 2007). In the context of spontaneous electroencephalogram activity, ERPs display very small amplitudes, usually between 2 and 20 microvolts, so averaging allows the event-related activity to become noticeable relative to the background activity (Altenmüller, Münte, & Gerloff, 2005; Baddeley et al., 2015; Kappenman & Luck, 2012; see Figure 1 for an example of averaged ERP waveforms and Chapter V for more details on data collection and data analysis procedures).

The ERPs are considered a direct, real-time measure of brain activity in the sense that they are spontaneous manifestations of the electrochemical processes that characterize the communication between neurons, being detected at the surface of the scalp almost instantaneously, with a millisecond temporal precision (Bunge & Kahn, 2009; Cohen, 2014; Kappenman & Luck, 2012; Luck, 2014). Nonetheless, the neural origins of the ERPs show that only a portion of brain activity is captured by the EEG/ERP technique. It is agreed that the majority of voltage fluctuations recorded with the electroencephalogram results from excitatory/inhibitory postsynaptic potentials, i.e., alterations in the flux of ions across the cell membrane prompted by the connection of neurotransmitters to postsynaptic receptors (Kappenman & Luck, 2016; Wilding & Ranganath, 2012). More specifically, it results from the synchronous activity of large populations of neurons similarly aligned in terms of spatial direction (i.e., perpendicular to the cortical sheath), which enables the summation rather than the cancelation of the

electric fields generated by neurons (Cohen, 2017; Fabiani et al., 2007; Kappenman & Luck, 2012, 2016; Luck, 2014). Only specific neural populations comply with these conditions, the pyramidal cells, which represent a significant percentage (70 - 90 %) of the cortical neurons (Bunge & Kahn, 2009; Keil, 2013). This implies that ERPs are unlikely to result from action potentials, or to have their origins in cortical interneurons or in close-field structures, such as midbrain nuclei, whose electric activity tends to cancel out (Fabiani et al., 2007; Luck, 2014).

As we can see, one of the main advantages of the EEG/ERP technique is its excellent temporal resolution and, for this reason, it is considered as the "goal standard" amongst the non-invasive brain imaging techniques in this regard (Dolcos, LaBar, & Cabeza, 2006; Fabiani et al., 2007; Johnson, 1997). Thus, this technique is particularly suited to evaluate and compare the time-course of covert processing between different experimental conditions, capturing activity without or even before behavioral responses are made (Dolcos et al., 2020; Friedman & Johnson, 2000; Rugg, Allan, & Birch, 2000). Accordingly, it has been particularly useful to collect experimental or clinical data from participants for whom is difficult or even impossible to obtain a behavioral response (Luck, 2014), such as infants (e.g., Cruz, Crego, Ribeiro, Gonçalves, & Sampaio, 2015; Riva et al., 2018; Shany & Berger, 2011), children (e.g., Taylor & Baldeweg, 2002), or patients with disorders of consciousness (e.g. Luauté, Morlet, & Mattout, 2015; Risetti et al., 2013). Also, when compared to other techniques, it is less expensive and it is associated with a lower level of discomfort to the participants as it is not necessary to use radioactive traces as it is in the case of PET, nor to be enclosed in a restricted and noisy environment as it is in the case of fMRI (Kappenman & Luck, 2016). However, EEG data collection can be somehow burdensome to the participants. Sessions are usually long due to the time needed to do the proper montage of the equipment and to complete the experimental task, which often requires several repetitions of the same stimulus category to increase the signal-to-noise ratio (Luck, 2014). Participants are also asked to stay still as much as possible to reduce the contamination of the signal by artifacts, which may be a source of discomfort. Two other relevant drawbacks of this technique are the superposition problem (i.e., the waveforms are composed of different overlapping components, being difficult to disentangle and isolate each underlying component) and its weak spatial resolution (i.e., there are many anatomical source candidates that could account for the patterns observed at the surface of the scalp, being a conundrum to adjudicate between possible combinations of neural generators; Altenmüller et al., 2005; Bunge & Kahn, 2009; Fabiani et al., 2007; Luck, 2014).

In its turn, the fMRI technique offers excellent spatial resolution and allows to explore neural networks recruited when one is engaged in a specific task (Baddeley et al., 2015; Dolcos et al., 2006,

2020; Johnson, 1997). Specifically, fMRI captures alterations in hemodynamic responses relying on the assumption that when brain areas are required to support a given operation, there is an increase in the need and consumption of oxygen, glucose, and other nutrients (Baddeley et al., 2015; Wager, Hernandez, Jonides, & Lindquist, 2007). This implies an increase in the local blood flow or volume and, particularly, fMRI makes use of the Blood Oxygen Level Dependent (BOLD) method to gauge the ratio of oxygenated *vs.* deoxygenated hemoglobin in the blood (Bunge & Kahn, 2009; Dolcos et al., 2006; Wilding & Ranganath, 2012). Oxyhemoglobin and deoxyhemoglobin have different magnetic properties: the first is mildly repelled by magnets whereas the second is attracted by them, hence changes in the oxy/deoxyhemoglobin ratio are associated with alterations in the local magnetic properties (Armony & Han, 2013). Thus, when contrasting different experimental conditions, such as an "active" *vs.* "rest" condition, differences between the engaged neural activity are expected and, consequently, differences in the intensity of the BOLD signal can also be detected (Armony & Han, 2013; Baddeley et al., 2015; Wager et al., 2007). Due to these characteristics, the fMRI technique is deemed as an indirect measure of brain activity (Bunge & Kahn, 2009; Wager et al., 2007).

Even though this technique has great spatial resolution with full-brain coverage, allowing to capture activity of deep subcortical structures, which is not the case when using the EEG/ERP technique, its temporal resolution is low, in the order of seconds (Armony & Han, 2013; Dolcos et al., 2006, 2020). This happens because hemodynamic responses are slow, taking some seconds to reach their peak; for example, the hemodynamic response in the visual cortex in response to a brief target takes around 4 – 6 seconds to peak (Bunge & Kahn, 2009; Wager et al., 2007). Moreover, in comparison with the EEG equipment, the use of an MRI scanner implies greater costs in terms of machinery, supplies, and technical personnel (Luck, 2014). In fact, the MRI scanner is a sophisticated piece of engineering and its magnetic field can be 1×10^4 to 1×10^5 times more powerful than the earth's field (Tsai, Grant, Mortele, Kung, & Smith, 2015; Wager et al., 2007). It has unique safety hazards, such interactions with implants and devices, burns, acoustic injuries, among others; hence, safety measures are of uttermost importance (Hartwig et al., 2009; Tsai et al., 2015). From the information presented so far, EEG and fMRI are complementary techniques due to their temporal and spatial resolution, respectively (Citron, 2012; Dolcos et al., 2006; Fabiani et al., 2007). Their combination represents quite a challenge and it is under development even nowadays (Luck, 2014). Nonetheless, it is already possible to use fMRI to aid with the source location of EEG/ERP signals or to acquire concomitantly fMRI and EEG data (e.g., Bunge & Kahn, 2009; Lei, Wu, & Valdes-Sosa, 2015; Moore, Shafer, Bakhtiari, Dolcos, & Singhal, 2019).

When using ERP and fMRI techniques to probe memory processes, event-related designs are commonly adopted as they allow to group and compare the neural activity recorded trial-by-trial according to specific experimental conditions (Armony & Han, 2013). The main aim is to isolate activity of interest, that is, memory-related activity during encoding and retrieval (Dolcos et al., 2006; Wager et al., 2007). More specifically, in a typical memory recognition task, participants are usually presented with several stimuli, which are later mixed with unstudied/new stimuli. At this stage, participants are required to decide for each stimulus if it was seen previously (old item) or not (new item), i.e., to make old/new judgments. Thus, when studying encoding processes, one common approach is to divide and compare the encoding activity as a function of the latter recognition status, i.e., accurately recognized *vs.* forgotten (Altenmüller et al., 2005; Dolcos et al., 2006, 2020; Kensinger & Schacter, 2008; Luck, 2014). For example, the difference between remembered *vs.* forgotten encoding neural activity is called difference due to/in subsequent memory or differential neural activity based on memory (Dm effect or subsequent memory effect; e.g., Dolcos & Cabeza, 2002; Guo, Duan, Li, & Paller, 2006; Leynes & Crawford, 2018; Yick et al., 2015). In a similar fashion, during retrieval, it is possible to compare the neural activity underlying successful and unsuccessful memory attempts, as well as old and new probes (Altenmüller et al., 2005; Baddeley et al., 2015; Kensinger & Schacter, 2008; Luck, 2014). With these possible comparisons in mind, the two subsequent sections provide a brief overview of ERP and fMRI findings in the domain of source memory retrieval.

Temporal dynamics of source memory: Event-related potential signatures

By comparing the neural activity of old and new probes recorded during memory recognition, different old/new effects have been documented with distinct temporal courses, polarities, spatial topographies, and putative functional roles (see Altenmüller et al., 2005; Wilding & Ranganath, 2012; Figure 1 for a summary). The two most studied old/new effects are the early frontal old/new effect (or frontal N400-like component: FN400) and the parietal old/new effect (Curran, 2000; Leynes & Phillips, 2008). The first is a wave deflection that is more negative for new in comparison with old items, emerging around 300 - 500 ms post-stimulus, and reaching its maximum response over frontal scalp areas. The second is a positive deflection appearing between 500 - 800 ms, whose positivity is larger for old *vs.* new items, with maximum response over parietal electrodes. These effects have been particularly studied in the context of dual-process memory frameworks that posit familiarity and recollection as two distinct processes underpinning memory recognition (Jacoby, 1991; Rugg & Curran, 2007; Yonelinas, 2002). Whereas familiarity has been conceptualized as a faster and more automatic process that signals that an

item was previously encountered, but without the retrieval of details from the encoding episode, recollection is deemed as slower and more intentional, effortful process related with the retrieval of specific details (Friedman & Johnson, 2000; Yonelinas, 2002; Yonelinas & Ritchey, 2015).

In this context, the early frontal old/new effect has been advanced as an index of familiarity as the repetition of stimuli *per se* engenders the effect (Czernochowski, Mecklinger, Johansson, & Brinkmann, 2005). Also, the effect appears for items that share perceptual or conceptual similarities with the studied targets (Curran, 2000; Curran & Dien, 2003). In this regard, the amplitude of the early frontal old/new effect is modulated by the presence of conceptually related *vs*. conceptually unrelated primers, suggesting that conceptual priming plays a role during recognition tests, and that the functional specificity of this effect as a correlate of familiarity remains under debate (Voss & Federmeier, 2011). The parietal old/new effect, in its turn, has been consistently associated with recollection as its amplitude is larger for responses associated with conscious and accurate retrieval of details of the encoding episode when compared to incorrect source responses or even to responses associated with familiarity judgments (Addante, Ranganath, & Yonelinas, 2012; Allan & Rugg, 1998; Cansino & Trejo-Morales, 2008; Leynes, Askin, & Landau, 2017; C. A. MacLeod & Donaldson, 2017; Peters & Daum, 2009; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999; Wilding & Rugg, 1997; Woroch & Gonsalves, 2010). Furthermore, its amplitude is modulated by the amount of information retrieved from the study episode, being larger when more details are retrieved (e.g., Leynes & Crawford, 2018; Leynes, Crawford, & Bink, 2005; Wilding, 2000; Woroch & Gonsalves, 2010) and even attenuated when participants are instructed to actively suppress memories from prior learning episodes, such as in the case of Think/No-Think tasks (e.g., López-Caneda, Crego, Campos, González-Villar, & Sampaio, 2019).

Besides these old/new effects, other two effects have been reported specially in the context of source memory tasks: the late right frontal old/new effect and the late posterior negativity (LPN; e.g., Cycowicz & Friedman, 2003; Donaldson & Rugg, 1998; Johansson & Mecklinger, 2003; Mecklinger, Rosburg, & Johansson, 2016). The first is a positive-going deflection that commences around 800 - 1000 ms and that lasts until the end of the epoch under analysis; its amplitude is larger for studied items in contrast to unstudied ones, with maximum response over right frontal sites (Friedman & Johnson, 2000). The functional role of this effect stills under debate, even so several authors have suggested that it is a correlate of evaluative and monitoring processes acting upon the information retrieved from memory to support memory decisions (Donaldson & Rugg, 1998; Leynes, 2002; Leynes, Grey, & Crawford, 2006; Trott et al., 1999; Van Petten, Senkfor, & Newberg, 2000; Wilding & Rugg, 1996). Such postretrieval monitoring role has been attributed to both successful (Wilding & Rugg, 1997) and unsuccessful memory

decisions, including incorrect source judgments and false alarms (Trott et al., 1999; Vallesi & Shallice, 2006). Also, this effect is modulated by distinct retrieval demands as the amplitude is larger for items that are difficult to recognize, such as items that were subjected to shallow encoding conditions (e.g., judged if the first and last letters of a certain word are in alphabetic order; Rugg et al., 2000).

The LPN is a late onset negativity that emerges around 1000 ms with a sustained time course, with maximum response over mid-posterior sites (Pz), with a more negative amplitude for old in comparison with new items, and that is observed regardless of the accuracy of the source memory response (Friedman, Cycowicz, & Bersick, 2005; Mecklinger, Johansson, Parra, & Hanslmayr, 2007; Mecklinger et al., 2016). Concerning its functional role, the LPN has been associated with the reconstruction, maintenance, and evaluation of item and contextual features, which may be mobilized when facing difficult source memory decisions (i.e., decisions that involve a more fine-grained examination of the information to discriminate between sources with overlapping features) or even when task-relevant details are not readily available (Johansson & Mecklinger, 2003; Leynes et al., 2006; Mecklinger et al., 2016; Wilding & Rugg, 1997). Furthermore, the number of contextual attributes activated in service of the reconstructions of the study episode may be reflected on the LPN amplitude, with larger amplitudes indicating more available information (Johansson & Mecklinger, 2003; Mecklinger et al., 2007).

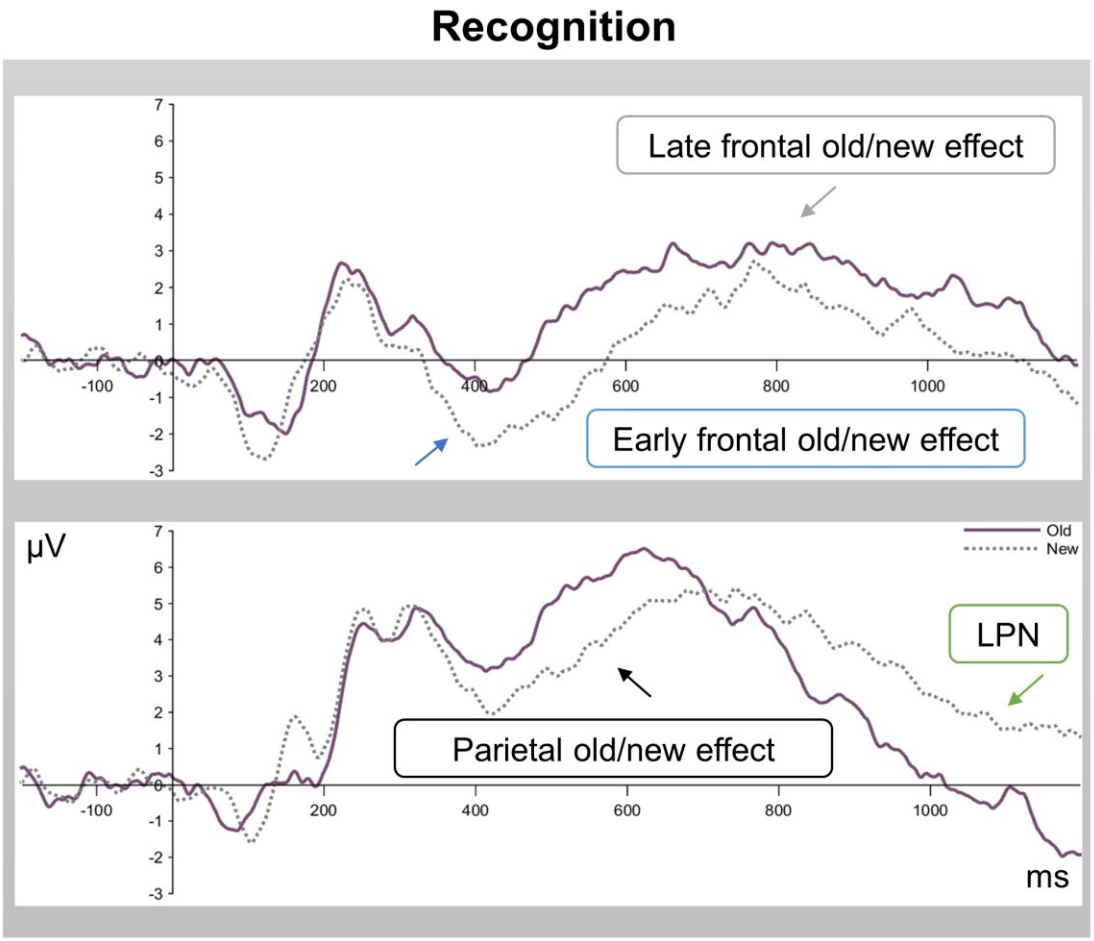
Regarding the influence of emotional information on these old/new effects and to avoid repetitions, this topic is addressed in more detail in Chapter V of this dissertation. Even so, the main effects of emotion on the amplitude of the former electrophysiological responses are by now outlined in Figure 1.

Neural basis of source memory: fMRI evidence

During the first pages of this dissertation two fundamental ideas were presented: the case H. M. was crucial to launch and pursue the notion that MTL areas play a pivotal role in memory functioning (Eichenbaum, 2012; Kensinger & Giovanello, 2006; Schacter, 1996; Squire, 2009; Squire & Kandel, 2009; Squire & Wixted, 2011); declarative/explicit memory relies on the functioning of medial temporal and diencephalic brain regions (Baddeley et al., 2015; Milner et al., 1998; Squire, 2009; Squire & Knowlton, 1995; Squire & Zola-Morgan, 1988; Squire & Zola, 1996). Thus, the involvement of neuroanatomical structures such as the hippocampus, the entorhinal, perirhinal, parahippocampal cortices, and the amygdala in source memory is also examined and acknowledged (Diana & Wang, 2018; Eichenbaum, 2012; Gold et al., 2006; Johnson, 2006; Mitchell & Johnson, 2009). Other brain region that is often implicated in source memory is the prefrontal cortex (PFC), whose relevance had been hinted by both neuropsychological studies and age-related research. Regarding the first, it has been documented

that patients with frontal lobe lesions are likely to display source amnesia, that is, they are able to learn specific information without recalling details from the original learning episode (Schacter, 1996). Additionally, some authors have argued that executive functions (e.g., planning; inhibition; mental flexibility), which have been intimately related to the frontal lobes (e.g., Alvarez & Emory, 2006; Nowrangi, Lyketsos, Rao, & Munro, 2014), play a role on source memory, even though such relationship remains unclear (El Haj & Allain, 2012; Glisky, Rubin, & Davidson, 2001). In favor of this argument, neurological disorders that imply alterations in terms of frontal lobe and executive functions, such as Parkinson's disease, are often associated with poorer source memory performance (e.g., Drag, Bieliauskas, Kaszniak, Bohnen, & Glisky, 2009). Regarding age-related contributions, the developmental trajectory of source memory follows an inverted "U" shape function similar to what has been documented for different cognitive abilities (e.g., Deary et al., 2009; Salthouse, 2004, 2009). This means that source memory tends to improve from childhood to adulthood (e.g., Cycowicz, Friedman, Snodgrass, & Duff, 2001; Raj & Bell, 2010; Sprondel, Kipp, & Mecklinger, 2011) and to decline from adulthood to late adulthood (e.g., Chua, Chen, & Park, 2006; Dennis et al., 2008; Dulas & Duarte, 2014; Dulas et al., 2011; Glisky et al., 2001). Some evidence even implies that different source memory tasks also pose distinct challenges across development; for example, the discrimination between internally-derived sources appears to be particularly difficult to children (Foley & Johnson, 1985; Johnson & Foley, 1984), as well as to older adults (Degl'Innocenti & Bäckman, 1996). In this context, one of the factors that is hypothesized to contribute to these developmental fluctuations is precisely the developmental trajectory of the PFC (also following an inverted "U" shape function), whose maturation occurs for several years extending into adulthood (e.g., Tanaka, Matsui, Uematsu, Noguchi, & Miyawaki, 2012) and whose signs of anatomical and functional decline are detected throughout aging (e.g., Cabeza & Dennis, 2013; Zanto & Gazzaley, 2019).

With respect to fMRI evidence, it supports not only the involvement of MTL and PFC on source memory (see Mitchell, 2016; Mitchell & Johnson, 2009 for reviews), but also the engagement of parietal regions during both encoding and retrieval (Spaniol et al., 2009). In this context, fMRI have been fostering a more fine-grained analysis of the neural networks underlying encoding and retrieval mechanisms. For example, during retrieval, it has been proposed that specific MTL regions support item and source memory (Ross & Slotnick, 2008; Weis et al., 2004). Whereas some regions are engaged in both item and source memory tasks, including the hippocampus and perirhinal cortex (Cansino et al., 2015; Diana, Yonelinas, & Ranganath, 2007; MacKenzie et al., 2015; Mugikura et al., 2010; Ross & Slotnick, 2008),



Old/new effects	Latency (ms)	Topography	Effects of emotion on amplitude
Early frontal	300 - 500	Frontal	X
Parietal	500 - 800	Parietal	↑ Emo
Late frontal	800 - 1000	Frontal	X or ↑ Emo
LPN	> 1000	Parietal	X or ↑ Emo

Figure 1. Schematic illustration of the old/new effects and summary of their latencies, topographies, and effects of emotion on their amplitude. Note: Emo = Emotional; LPN = Late Posterior Negativity; X = No effect on amplitude; ↑ = Increased amplitude.

there are other regions such as the parahippocampal cortex, whose activity seems to be particularly engaged during the retrieval of source information (Diana et al., 2007; Kahn, Davachi, & Wagner, 2004; Ross & Slotnick, 2008; Rugg et al., 2012). Furthermore, hippocampal activity is not only associated with the successful retrieval of source information, but also appears to co-vary with the amount of contextual

information retrieved from the study episode (Rugg et al., 2012; Suzuki, 2012). PFC areas are also expected to support the control of source monitoring processes (Mitchell, 2016). In this regard, the ventrolateral PFC has been hypothesized to be involved in various processes, such as search and evaluation of source features (Lundstrom, Ingvar, & Petersson, 2005; Mitchell et al., 2008), retrieval attempts associated with semantic details of the study items (Dobbins, Foley, Schacter, & Wagner, 2002; Raposo, Han, & Dobbins, 2009), and post-retrieval monitoring processes (Donaldson, Wheeler, & Petersen, 2010; Hayes, Buchler, Stokes, Kragel, & Cabeza, 2011; Takahashi, Ohki, & Miyashita, 2002). Also, the dorsolateral PFC has been implicated in the detection and monitoring of specific source features (Henson, Shallice, & Dolan, 1999; Mitchell, Raye, Johnson, & Greene, 2006; Mitchell et al., 2008; Rugg, Henson, & Robb, 2003). With respect to the parietal cortex, even though its functional role remains under debate, some of the major roles postulated for the posterior parietal cortex (PCC) entail (Berryhill, 2012; Mitchell & MacPherson, 2017; Simons, Peers, Mazuz, Berryhill, & Olson, 2010; see Mitchell, 2016 for review): subjective experiences that may occur during retrieval (e.g., confidence and vividness judgments); processing of multiple features that constitute a given event. It is noteworthy that the involvement of different MTL, PFC, parietal, and other regions also depends upon the type of information and specific task demands imposed during encoding and retrieval. Differences in the activity of these regions had been registered as a function of the modality of presentation (e.g., visual *vs.* auditory; Peters, Koch, Schwarz, & Daum, 2007; Peters, Suchan, Köster, & Daum, 2007), as well as of the type of study items (e.g., words; objects; pictures; Dulas & Duarte, 2011; Park, Abellanoza, Schaeffer, & Gandy, 2014). Moreover, differences had been recorded between types of source memory tasks and between specific source features (e.g., Brandt, Bergström, Buda, Henson, & Simons, 2014; Diana et al., 2007; Duarte, Henson, Knight, Emery, & Kim, 2010; Liang & Preston, 2017; Simons, Henson, Gilbert, & Fletcher, 2008; Suzuki et al., 2002; Turner, Simons, Gilbert, Frith, & Burgess, 2008).

When analyzing subsequent memory effects assessed during the encoding phase of source memory tasks, the neural networks recruited overlap to a great extent with the ones already described in the context of retrieval. Specifically, in contrast to incorrect source decisions, accurate source decisions were related with heightened activation in the hippocampus and parahippocampal cortex (Davachi, Mitchell, & Wagner, 2003; Diana, 2017; Gold et al., 2006; Peters, Suchan, et al., 2007). Even though hippocampal and perirhinal cortex activity has been documented in both item and source memory encoding (Dougal et al., 2007; Gold et al., 2006; Song, Jeneson, & Squire, 2011), it seems that hippocampal together with parahippocampal activity are particularly related to subsequent source recollection, which further suggests that specific MTL areas work in supplementary ways to support the

encoding of different aspects of a learning episode (Davachi et al., 2003; Ritchey et al., 2019). The recruitment of subsets of PFC regions seem to vary with the type of encoding tasks and related cognitive operations, as well as with the type of materials used during encoding (e.g., Dulas & Duarte, 2014; Sugimori, Mitchell, Raye, Greene, & Johnson, 2014). Furthermore, parietal and other posterior regions appear particularly engaged as a function of the online processing required by the perceptual features of the stimulus (e.g., Cansino, Maquet, Dolan, & Rugg, 2002; Kensinger & Schacter, 2006b; Sugimori et al., 2014; Uncapher, Otten, & Rugg, 2006). For instance, the intra-parietal sulcus is proposed to play a role on perceptual binding (i.e., integration of distinct perceptual features; Uncapher et al., 2006) and the lateral occipital cortex, an area especially responsive to objects, display greater activity during the study of objects and their locations that are later successfully retrieved (Cansino et al., 2002).

The participation of amygdala on emotion-related processes is well-established in the literature (see LeDoux & Phelps, 2008 for overview). Accordingly, the integration of the amygdala in the brain systems underpinning memory processes for emotion-laden information is algo recognized (see Buchanan & Adolphs, 2004; Dolcos et al., 2006; Holland & Kensinger, 2013; LaBar & Cabeza, 2006 for examples of comprehensive reviews). It has been proposed that amygdala assumes somehow a mediator role of the brain systems supporting memory, exerting influence on processes such as perception, attention, and semantic elaboration (Murty, Ritchey, Adcock, & LaBar, 2010). Specifically, greater amygdala-hippocampus connectivity had been observed during encoding and retrieval of emotional when compared to neutral items (Fastenrath et al., 2014; Smith, Stephan, Rugg, & Dolan, 2006). Also, greater activity in the amygdala, hippocampus, and entorhinal cortex was noticed during the retrieval of emotional *vs.* neutral items, and the emotion effect was enhanced in amygdalar-hippocampal areas for responses associated with the recollection of details from the study episode (Dolcos, Labar, & Cabeza, 2005; Ritchey et al., 2019). The activity of amygdala and other areas such as the orbitofrontal cortex were shown to support better source memory performance specifically for emotional items, while other areas such as lateral inferior PFC and parahippocampal cortex were reported for neutral items (Kensinger & Schacter, 2005b).

Furthermore, distinct effects of valence and arousal on memory-related brain networks had been postulated (Kensinger, 2004; Kensinger & Corkin, 2004): stimulus arousal is hypothesized to contribute to the EEM effect by mobilizing amygdalar-hippocampal areas and by increasing the likelihood of devoting attentional resources during encoding and of transforming the information into a more stable representation over time during consolidation; stimulus valence, in its turn, is hypothesized to engage more PFC-hippocampal areas, possibly associated with controlled encoding and elaborative processes.

Differences in the involvement of amygdalar-hippocampal and PFC-hippocampal networks was also documented as a function of valence and level of processing with the amygdala contributing especially to the memory of negative information in shallow encoding conditions and the PFC contributing to positive information in deep encoding conditions (Ritchey, LaBar, & Cabeza, 2011). Moreover, the encoding of negative information appears to recruit more occipito-temporal regions related to sensory processes, while positive information recruits more frontal regions related to elaborative processes (Steinmetz & Kensinger, 2009). It is suggested, once more, that memory-related neural correlates vary with emotion and encoding task. Of relevance for the current dissertation are the effects of stimulus valence and encoding tasks, especially when the stimuli are encoded with reference to the self, on the neural correlates of source memory.

As mentioned earlier, mPFC has been argued to be implicated in the processing of both emotional and self-referenced information (Gutchess & Kensinger, 2018). Considering the evidence reviewed so far regarding emotion effects together with the bulk of evidence implicating this area in self-referential tasks (e.g., D'Argembeau et al., 2007; Fossati et al., 2003; Kelley et al., 2002; Macrae, Moran, Heatherton, Banfield, & Kelley, 2004; Philippi, Duff, Denburg, Tranel, & Rudrauf, 2012), the involvement of the mPFC on source memory for self-related information is also highlighted (Leshikar & Duarte, 2012, 2014; see Mitchell, 2016). This is illustrated, for example, by studies indicating that the mPFC is particularly engaged when participants are required to discriminate between self-generated and external features such as perceived items or items generated by an external agent (e.g., Kensinger & Schacter, 2006b; Simons et al., 2008; Turner et al., 2008). Even though the interactive effects of emotion and self-referential processing on source memory have been scarcely addressed in the literature, some insights may be derived from studies probing these two factors in contexts other than source memory. Specifically, when contrasting the activity of emotional self-related stimuli *vs.* other encoding tasks, the mPFC and the cingulate cortex (CC) emerged as areas of interest (Herbert, Pauli, & Herbert, 2011). Additionally, other areas such as angular gyrus (AG), anterior temporal cortex, and amygdala have been implicated, especially in the case of negative self-referent processing (e.g., De Pisapia, Barchiesi, Jovicich, & Cattaneo, 2019; Yoshimura et al., 2009). Thus, areas such as the mPFC, CC, AG, and amygdala may support memory-related processes regarding emotional self-referential information (e.g., Yaoi et al., 2015), yet more studies are warranted to clarify this issue in the case of source memory.

A brief note on methodological choices, main objectives, and outline of this dissertation

As source memory is the key concept of the current dissertation, the way source memory features are manipulated during encoding and tested during retrieval might influence the behavioral performance, as well as related neural correlates. In typical source memory paradigms, the manipulation of source features during encoding is similar across studies in the sense that there is a variation of two or more features (e.g., source A *vs.* source B), even though these features might vary in nature (see Table 1). On the contrary, the test of source memory is quite variable from study to study as the test might be "direct" or "indirect" (see Chapter II for more details). Direct when the test focus primarily on source memory and indirect when the test first probes item memory and then source memory. Examples of indirect approaches include asking about source memory after old/new recognition tasks, remember/know judgments, or confidence ratings. Considering that the way memory is tested might influence the outcomes (e.g., Gao et al., 2016; Raposo, Frade, & Alves, 2016) and that the key concept is source memory, we used a direct approach given the main goal of the research project encompassing this dissertation: to probe the effects of emotional valence and encoding tasks in internal source memory attributions, using not only behavioral outcomes, but also brain-derived responses, namely ERPs and fMRI. For instance, if we were to use a standard old/new recognition, we knew that successful item recognition could occur without the retrieval of details from study episode, just based on familiarity processes (Kensinger & Giovanello, 2006; Mandler, 1980; Squire & Wixted, 2011; Tulving, 2001; Yonelinas, 2002). These old/new decisions could also incorporate the retrieval of details that do not inform about the features specifically manipulated during encoding. Thus, if we were to examine neural responses during old/new judgments, we would have responses with contextual details informative or not of the features manipulated during encoding, as well as responses based on familiarity. In other words, we would be increasing the possibility of mixing responses of interest with other possible responses. Alternatively, if we were to examine the neural responses not in the old/new recognition window, but in a subsequent source judgment window, we could not really ensure we would capture activity of interest regarding source monitoring processes. As participants get to know the gist of the memory test (i.e., first old/new, and then source if old), it is likely that source monitoring processes would take place even during old/new judgments. So, the activity of interest could also be missed. For similar reasons, we did not use a dissociation procedure (Jacoby, 1991). Specifically, this procedure allows to separate familiarity and recollection processes (i.e., when details of the study episode are retrieved) by requiring the participants to make inclusion/exclusion decisions. Participants are required to respond "yes" only to items that during encoding were presented with a given feature (inclusion), while responding "no" to items presented

with other features or to new items (exclusion). Although this is a different kind of source memory test, if we were to use it, besides decreasing the number of accurate source memory responses as just one source would be targeted, the exclusion of the other source could be based on the retrieval of informative details, uninformative details, or guess responses. Once again, the possibility of mixing activity of interest with other types of activity would be increased.

Considering that the source features manipulated during encoding in the current dissertation were related to encoding tasks and that most of the experiments employed two specific tasks, a self-referential and common judgment task (i.e., if the item represents something commonly used by people in everyday life), it is also pertinent to justify why the self-referential task was compared to a semantic task and not to other tasks. As mentioned earlier in this introduction, the advantage of self-referential processing on memory has been reported relative to other tasks, such as other-referent, semantic, and perceptual tasks (e.g., Fossati et al., 2003, 2004; Miall, 1986; Yoshimura et al., 2009; Zhou et al., 2017). By selecting a semantic task and not a perceptual one as a comparison, we followed prior literature (e.g., Durbin et al., 2017; Leshikar et al., 2015) and assured that participants would access the meaning of the items and elaborate on them as it happens in the case of self-reference tasks. At the same time, by not using an other-referent task, we anticipated some degree of differentiation between the cognitive operations engaged, since it had been suggested that both other- and self-referencing may rely on similar processes (e.g., Denny, Kober, Wager, & Ochsner, 2012; Yoshimura et al., 2009). As such, these qualitative differences between tasks may be relevant when studying internal source memory given it mainly relies on discriminating between details pertaining to cognitive operations (Ferguson et al., 1992; Finke et al., 1988; Johnson, 1988, 1997, 2006; Johnson & Raye, 1981; Raye & Johnson, 1980). The more the cognitive operations differ, the more the participants are likely to show a reasonable source memory performance and avoid floor effects, which is an important element to consider in internal source memory tasks as they might be difficult due to the similarities and overlap amongst features (Johnson, 1997; Johnson & Foley, 1984; Raye & Johnson, 1980).

Moreover, in the experiments included in this dissertation, there was an effort to focus on stimulus valence (negative *vs.* neutral *vs.* positive) while trying to control for arousal to better isolate valence-related effects. Indeed, most of the studies examining the effects of emotional information on source memory tend to use negative and positive stimuli that are more arousing than neutral ones (see Chapter II), thus placing a particular emphasis on arousal whilst valence remains less explored in the literature (Holland & Kensinger, 2013). Also, the effects of emotion on source memory have been mainly achieved through the manipulation of the intrinsic affective properties of the stimuli. Nonetheless, other approach that has

been used in item memory research relies on the manipulation of emotional cues that are presented together with neutral study materials, i.e., a manipulation of the encoding context (see Chapter IV for more details). Few studies have implemented such manipulation in the case of source memory (see Macri, Pavard, & Versace, 2018; Ventura-Bort et al., 2017; Xie & Zhang, 2017 for examples), so the current dissertation also aimed to give some contribute in this regard. The inclusion of these two approaches allowed to select one of them to be further examined with non-invasive brain imaging techniques.

Finally, as aforementioned, the series of studies included in this dissertation integrated an extended project whose main goal was to probe the effects of emotion and encoding task on internal source memory including both behavioral and neuroimaging data. Thus, this goal is transversal to all studies included here, even though specific objectives are presented for each study. Also, as a secondary goal, some of the studies included prospective metamemory judgments of source, i.e., JOSs, to examine how the former factors may affect the participants' beliefs regarding source memorability (see Chapters II and V). More specifically, besides the current introductory chapter (Chapter I), this dissertation includes five more chapters. The Chapter II is a complement to this general introduction as it presents a meta-analytic review of the effects of stimulus emotion on source memory performance. Being a systematic review, it includes further details regarding research data and main theoretical proposals developed in this domain. The following two chapters, Chapter III and Chapter IV, correspond to behavioral studies examining the effects of stimulus emotion and encoding strategy on source memory. In the first, emotion was manipulated based on the intrinsic affective properties of the stimuli, whereas in the second emotion was manipulated as an affective cue extrinsic and presented together with the stimuli. Chapter V addresses electrophysiological correlates, considering both encoding and retrieval phases. In the last chapter, Chapter VI, a general discussion of the main findings is provided.

References

Addante, R. J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. *NeuroImage*, *62*(1), 439–450. https://doi.org/10.1016/j.neuroimage.2012.04.031

Allan, K., & Rugg, M. D. (1998). Neural correlates of cued recall with and without retrieval of source memory. *NeuroReport*, *9*(15), 3463–3466. https://doi.org/10.1097/00001756-199810260-00023

Altenmüller, E. O., Münte, T. F., & Gerloff, C. (2005). Neurocognitive function and the EEG. In E.

Niedermeyer & da Silva, F. L. (Eds.). *Electroencephalography: Basic principles, clinical applications, and related fields* (5th ed., 661-682). Philadelphia, USA: Lippincott Williams & Wilkins.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, *16*(1), 17–42. https://doi.org/10.1007/s11065-006-9002-x

Anderson, J. R. (1995). *Learning and memory: An integrated approach*. New Jersey, USA: John Wiley & Sons.

Annese, J., Schenker-Ahmed, N. M., Bartsh, H., Maechler, P., Sheh, C., Thomas, N., ... Corkin, S. (2014). Postmortem examination of patient H. M.'s brain based on histological sectioning and digital 3D reconstruction. *Nature Communications*, *5*, 3122. https://doi.org/10.1038/ncomms4122

Armony, J., & Han, J. E. (2013). PET and fMRI: Basic principles and applications in affective neuroscience research. In J. Armony & P. Vuilleumier (Eds), *The Cambridge handbook of human affective neuroscience* (pp. 133-153). New York, USA: Cambridge University Press.

Aupée, A. M. (2007). A detrimental effect of emotion on picture recollection. *Scandinavian Journal of Psychology*, *48*(1), 7–11. https://doi.org/10.1111/j.1467-9450.2006.00533.x

Baddeley, A. D. (1999). *Cognitive psychology: A modular course. Essentials of human memory.* East Sussex, UK: Psychology Press/Taylor & Francis.

Baddeley, A., Eysenck, M. W. & Anderson, M. C. (2015). *Memory* (2nd ed.). New York, USA: Psychology Press.

Barnacle, G. E., Tsivilis, D., Schaefer, A., & Talmi, D. (2018). Local context influences memory for emotional stimuli but not electrophysiological markers of emotion-dependent attention. *Psychophysiology*, *55*(4), 1–14. https://doi.org/10.1111/psyp.13014

Berryhill, M. E. (2012). Insights from neuropsychology: Pinpointing the role of the posterior parietal cortex in episodic and working memory. *Frontiers in Integrative Neuroscience*, *6*, 31. https://doi.org/10.3389/fnint.2012.00031

Blix, I., & Brennen, T. (2011). Intentional forgetting of emotional words after trauma: A study with victims of sexual assault. *Frontiers in Psychology*, *2*, 235. https://doi.org/10.3389/fpsyg.2011.00235

Bodner, G. E., & Taikh, A. (2012). Reassessing the basis of the production effect in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1711–1719. https://doi.org/10.1037/a0028466

Borg, C., Leroy, N., Favre, E., Laurent, B., & Thomas-Antérion, C. (2011). How emotional pictures influence visuospatial binding in short-term memory in ageing and Alzheimer's disease? *Brain*

and Cognition, *76*(1), 20–25. https://doi.org/10.1016/j.bandc.2011.03.008

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Bradley, M. M., & Lang, P. J. (2007). *International Affective Digitized Sounds* (*2nd Edition*; *IADS-2*): *Affective ratings of sounds and instruction manual* (Technical Report No. B-3). Gainesville, FL: University of Florida, NIMH Center for the Study of Emotion and Attention.

Brandt, V. C., Bergström, Z. M., Buda, M., Henson, R. N. A., & Simons, J. S. (2014). Did i turn off the gas? Reality monitoring of everyday actions. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 209–219. https://doi.org/10.3758/s13415-013-0189-z

Brookwell, M. L., Bentall, R. P., & Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: A meta-analytic review. *Psychological Medicine*, *43*(12), 2465–2475. https://doi.org/10.1017/S0033291712002760

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*(1), 73–99. https://doi.org/10.1016/0010-0277(77)90018-X

Buchanan, T. W., & Adolphs, R. (2004). Memory for emotional events. In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 42-75). New York, USA: Oxford University Press.

Bunge, S. A., & Kahn, I. (2009). Cognition: An overview of neuroimaging technique. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (pp. 1063-1067). Oxford, USA: Academic Press.

Burnham, W. H. (1888). Memory , historically and experimentally considered . I . An historical sketch of the older conceptions of memory. *The American Journal of Psychology*, *2*(1), 39–90. https://doi.org/10.2307/1411406

Cabeza, R., & Dennis, N. A. (2013). Frontal lobes and aging: Deterioration and compensation. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (2nd ed., pp. 628–652). New York, USA: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195134971.001.0001

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral Cortex*, *12*(10), 1048–1056. https://doi.org/10.1093/cercor/12.10.1048

Cansino, S., Torres-Trejo, F., Estrada-Manilla, C., Ramírez-Barajas, L., Pérez-Loyda, M., Nava-Chaparro, A., … Ruiz-Velasco, S. (2019). Predictors of source memory success and failure in older adults. *Frontiers in Aging Neuroscience*, *10*, 17. https://doi.org/10.3389/fnagi.2019.00017

Cansino, S., & Trejo-Morales, P. (2008). Neurophysiology of successful encoding and retrieval of source memory. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 85–98. https://doi.org/10.3758/CABN.8.1.85

Cansino, S., Trejo-Morales, P., Estrada-Manilla, C., Pasaye-Alcaraz, E. H., Aguilar-Castañeda, E., Salgado-Lujambio, P., & Sosa-Ortiz, A. L. (2015). Brain activity during source memory retrieval in young, middle-aged and old adults. *Brain Research*, *1618*, 168–180. https://doi.org/10.1016/j.brainres.2015.05.032

Carroll, M., Mazzoni, G., Andrews, S., & Pocock, P. (1999). Monitoring the future: Object and source memory for real and imagined events. *Applied Cognitive Psychology*, *13*(4), 373–390. https://doi.org/10.1002/(SICI)1099-0720(199908)13:4<373::AID-ACP605>3.0.CO;2-F

Chiu, Y. C., Dolcos, F., Gonsalves, B. D., & Cohen, N. J. (2013). On opposing effects of emotion on contextual or relational memory. *Frontiers in Psychology*, *4*, 103. https://doi.org/10.3389/fpsyg.2013.00103

Christianson, S.-Å., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, *1*(4), 225–239. https://doi.org/10.1002/acp.2350010402

Christianson, S.-Å., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition and Emotion*, *5*(2), 81–108. https://doi.org/10.1080/02699939108411027

Chua, H. F., Chen, W., & Park, D. C. (2006). Source memory, aging and culture. *Gerontology*, *52*(5), 306–313. https://doi.org/10.1159/000094612

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*(3), 211–226. https://doi.org/10.1016/j.bandl.2011.12.007

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. Massachusetts, EUA: MIT Press.

Cohen, M. X. (2017). Where does EEG come from and what does it mean? *Trends in Neurosciences*, *40*(4), 208–218. https://doi.org/10.1016/j.tins.2017.02.004

Conway, M. A., Dewhurst, S. A., Pearson, N., & Sapute, A. (2001). The self and recollection reconsidered: How a "failure to replicate" failed and why trace strength accounts of recollection are untenable. *Applied Cognitive Psychology*, *15*(6), 673–686. https://doi.org/10.1002/acp.740

Cook, G. I., Hicks, J. L., & Marsh, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition*, *35*(2), 222–230. https://doi.org/10.3758/BF03193443

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294. https://doi.org/10.1037/0096-3445.104.3.268

Cruz, S., Crego, A., Ribeiro, E., Gonçalves, Ó., & Sampaio, A. (2015). A VEP study in sleeping and awake one-month-old infants and its relation with social behavior. *International Journal of Developmental Neuroscience*, *41*, 37–43. https://doi.org/10.1016/j.ijdevneu.2014.12.006

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923–938. https://doi.org/10.3758/BF03209340

Curran, T., & Dien, J. (2003). Differentiating amodal familiarity from modality-specific memory processes: An ERP study. *Psychophysiology, 40*(6), 979–988. https://doi.org/10.1111/1469-8986.00116

Cycowicz, Y. M., & Friedman, D. (2003). Source memory for the color of pictures: Event-related brain potentials (ERPs) reveal sensory-specific retrieval-related activity. *Psychophysiology*, *40*(3), 455–464. https://doi.org/10.1111/1469-8986.00047

Cycowicz, Y. M., Friedman, D., Snodgrass, J. G., & Duff, M. (2001). Recognition and source memory for pictures in children and adults. *Neuropsychologia*, *39*(3), 255–267. https://doi.org/10.1016/S0028-3932(00)00108-1

Czernochowski, D., Mecklinger, A., Johansson, M., & Brinkmann, M. (2005). Age-related differences in familiarity and recollection: ERP evidence from a recognition memory study in children and young adults. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(4), 417–433. https://doi.org/10.3758/CABN.5.4.417

D'Argembeau, A., Comblain, C., & Van der Linden, M. (2005). Affective valence and the self-reference effect: Influence of retrieval conditions. *British Journal of Psychology*, *96*(4), 457–466. https://doi.org/10.1348/000712605X53218

D'Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Balteau, E., Luxen, A., ... Salmon, E. (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *Journal of Cognitive Neuroscience*, *19*(6), 935–944. https://doi.org/10.1162/jocn.2007.19.6.935

D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4*(2), 173–188. https://doi.org/10.1037/1093-4510.

D'Argembeau, A., & Van der Linden, M. (2005). Influence of emotion on memory for temporal information. *Emotion*, *5*(4), 503–507. https://doi.org/10.1037/1528-3542.5.4.503

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 2157–2162. https://doi.org/10.1073/pnas.0337195100

Davidson, P. S. R., McFarland, C. P., & Glisky, E. L. (2006). Effects of emotion on item and source memory in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 306–322. https://doi.org/10.3758/CABN.6.4.306

De Pisapia, N., Barchiesi, G., Jovicich, J., & Cattaneo, L. (2019). The role of medial prefrontal cortex in processing emotional self-referential information: a combined TMS/fMRI study. *Brain Imaging and Behavior*, *13*(3), 603–614. https://doi.org/10.1007/s11682-018-9867-3

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., ... Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135–152. https://doi.org/10.1093/bmb/ldp033

Degl'Innocenti, A., & Bäckman, L. (1996). Aging and source memory: Influences of intention to remember and associations with frontal lobe tests. *Aging, Neuropsychology, and Cognition*, *3*(4), 307–319. https://doi.org/10.1080/13825589608256632

Delaney-Busch, N., Wilkie, G., & Kuperberg, G. (2016). Vivid: How valence and arousal influence word processing under different task demands. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(3), 415–432. https://doi.org/10.3758/s13415-016-0402-y

Dennis, N. A., Hayes, S. M., Prince, S. E., Madden, D. J., Huettel, S. A., & Cabeza, R. (2008). Effects of aging on the neural correlates of successful item and source memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 791–808. https://doi.org/10.1037/0278-7393.34.4.791

Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *24*(8), 1742–1752. https://doi.org/10.1162/jocn_a_00233

Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, *90*(4), 286–297. https://doi.org/10.1037/0021-843X.90.4.286

Dewhurst, S A, & Conway, M. A. (1995). The self and recollective experience. *Applied Cognitive Psychology*, *9*(1), 1–19. https://doi.org/10.1002/acp.2350090102

Dewhurst, S. A., Anderson, R. J., Grace, L., & Boland, J. (2017). Survival processing versus self-reference: A memory advantage following descriptive self-referential encoding. *Journal of Memory and Language*, *94*, 291–304. https://doi.org/10.1016/j.jml.2017.01.003

Diana, R. A. (2017). Parahippocampal cortex processes the nonspatial context of an event. *Cerebral Cortex*, *27*(3), 1808–1816. https://doi.org/10.1093/cercor/bhw014

Diana, R. A., & Wang, F. (2018). Episodic memory. In J. T. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed., pp. 1–33). New Jersey, USA: John Wiley & Sons. https://doi.org/10.1002/9781119170174.epcn103

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Diaz-Asper, C., Malley, J., Genderson, M., Apud, J., & Elvevåg, B. (2008). Context binding in schizophrenia: Effects of clinical symptomatology and item content. *Psychiatry Research*, *159*(3), 259–270. https://doi.org/10.1016/j.psychres.2007.02.018

Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: Multiple prefrontal processes subserve source memory. *Neuron*, *35*(5), 989–996. https://doi.org/10.1016/S0896-6273(02)00858-9

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, *8*(1), 155–161. https://doi.org/10.3758/BF03196152

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1*(1), 5–11. https://doi.org/10.1037/1528-3542.1.1.5

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: Encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(3), 252–263. https://doi.org/10.3758/CABN.2.3.252

Dolcos, F., Katsumi, Y., Moore, M., Berggren, N., de Gelder, B., Derakshan, N., ... Dolcos, S. (2020). Neural correlates of emotion-attention interactions: From perception, learning, and memory to social cognition, individual differences, and training interventions. *Neuroscience & Biobehavioral Reviews*, *108*, 559–601. https://doi.org/10.1016/j.neubiorev.2019.08.017

Dolcos, F., Katsumi, Y., Weymar, M., Moore, M., Tsukiura, T., & Dolcos, S. (2017). Emerging directions

in emotional episodic memory. *Frontiers in Psychology*, *8*, 1867. https://doi.org/10.3389/fpsyg.2017.01867

Dolcos, F., LaBar, K. S., & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), 2626–2631. https://doi.org/10.1073/pnas.0409848102

Dolcos, F., LaBar, K. S.., & Cabeza, R. (2006). Memory for emotional episodes: The strengths and limits of arousal-based accounts. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 107-133). Oxford, United Kingdom: Blackwell Publishing.

Dolcos, F., Wang, L., & Mather, M. (2014). Current research and emerging directions in emotion-cognition interactions. *Frontiers in Integrative Neuroscience*, *8*, 83. https://doi.org/10.3389/fnint.2014.00083

Donaldson, D. I., & Rugg, M. D. (1998). Recognition memory for new associations: Electrophysiological evidence for the role of recollection. *Neuropsychologia*, *36*(5), 377–395. https://doi.org/10.1016/S0028-3932(97)00143-7

Donaldson, D. I., Wheeler, M. E., & Petersen, S. E. (2010). Remember the source: Dissociating frontal and parietal contributions to episodic memory. *Journal of Cognitive Neuroscience*, *22*(2), 377–391. https://doi.org/10.1162/jocn.2009.21242

Dougal, S., Phelps, E. A., & Davachi, L. (2007). The role of medial temporal lobe in item recognition and source recollection of emotional stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(3), 233–242. https://doi.org/10.3758/CABN.7.3.233

Drag, L. L., Bieliauskas, L. A., Kaszniak, A. W., Bohnen, N. I., & Glisky, E. L. (2009). Source memory and frontal functioning in Parkinson's disease. *Journal of the International Neuropsychological Society, 15*(3), 399-406. https://doi.org/10.1017/S1355617709090572

Duarte, A., Henson, R. N., Knight, R. T., Emery, T., & Kim, S. (2010). The orbitofrontal cortex is necessary for temporal context memory. *Journal of Cognitive Neuroscience*, *22*(8), 1819–1831. https://doi.org/10.1162/jocn.2009.21316

Dulas, M. R., & Duarte, A. (2011). The effects of aging on material-independent and material-dependent neural correlates of contextual binding. *NeuroImage*, *57*(3), 1192–1204. https://doi.org/10.1016/j.neuroimage.2011.05.036

Dulas, M. R., & Duarte, A. (2014). Aging affects the interactio between attentional control and source memory: An fMRI study. *Journal of Cognitive Neuroscience*, *26*(12), 2653–2669.

https://doi.org/10.1162/jocn_a_00663

Dulas, M. R., Newsome, R. N., & Duarte, A. (2011). The effects of aging on ERP correlates of source memory retrieval for self-referential information. *Brain Research*, *1377*, 84–100. https://doi.org/10.1016/j.brainres.2010.12.087

Dunlosky, J., Mueller, M. L., & Tauber, S. K. (2015). The contribution of processing fluency (and beliefs) to people's judgments of learning. In D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger, III (Eds.), *Remembering: Attributions, processes, and control in human memory* (pp. 46-64). New York, USA: Psychology Press.

Dunlosky, J., Mueller, M. L., & Thiede, K. W. (2016). Methodology for investigating human metamemory: Problems and pitfalls. In J. Dunlosky & U. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 23-37). New York, USA: Oxford University Press.

Dunlosky, J, & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271–280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory*, *25*(9), 1191–1200. https://doi.org/10.1080/09658211.2017.1282517

Dutton, A., & Carroll, M. (2001). Eyewitness testimony: Effects of source. *Australian Journal of Psychology*, *53*(2), 83–91. https://doi.org/10.1080/00049530108255128

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183–201. https://doi.org/10.1037/h0047707

Eichenbaum, H. (2012). What H. M. taught us. *Journal of Cognitive Neuroscience*, *25*(1), 14–21. https://doi.org/10.1162/jocn_a_00285

Ekman, P. (1992a). An argument for basic emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P. (1992b). Are there basic emotions? *Psychological Review*, *99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

El Haj, M., & Allain, P. (2012). What do we know about the relationship between source monitoring deficits and executive dysfunction? *Neuropsychological Rehabilitation*, *22*(3), 449–472. https://doi.org/10.1080/09602011.2012.658267

El Haj, M., Fasotti, L., & Allain, P. (2012). Source monitoring in Alzheimer's Disease. *Brain and Cognition*, *80*(2), 185–191. https://doi.org/10.1016/j.bandc.2012.06.004

El Haj, M., Kessels, R. P. C., & Allain, P. (2016). Source memory rehabilitation: A Review toward recommendations for setting up a strategy training aimed at the "what, where, and when" of episodic retrieval. *Applied Neuropsychology: Adult*, *23*(1), 53–60. https://doi.org/10.1080/23279095.2014.992071

Fabiani, M., Gratton, G., & Federmeier, K. D. (2007). Event-related brain potentials: Methods, theory, and applications. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 85–119). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511546396.002

Fairfield, B., Mammarella, N., Palumbo, R., & Di Domenico, A. (2015). Emotional meta-memories: A review. *Brain Sciences*, *5*(4), 509–520. https://doi.org/10.3390/brainsci5040509

Fastenrath, M., Coynel, D., Spalek, K., Milnik, A., Gschwind, L., Roozendaal, B., ... de Quervain, D. J. F. (2014). Dynamic modulation of amygdala-hippocampal connectivity by emotional arousal. *Journal of Neuroscience*, *34*(42), 13935–13947. https://doi.org/10.1523/JNEUROSCI.0786-14.2014

Ferguson, S. A., Hastroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging*, *7*(3), 443–452. https://doi.org/10.1037/0882-7974.7.3.443

Ferré, P., Comesaña, M., & Guasch, M. (2019). Emotional content and source memory for language: Impairment in an incidental encoding task. *Frontiers in Psychology*, *10*, 65. https://doi.org/10.3389/fpsyg.2019.00065

Finke, R. A., Johnson, M. K., & Shyi, G. C. W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory & Cognition*, *16*(2), 133–137. https://doi.org/10.3758/BF03213481

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: Words, phrases, and task effects. *Progress in Brain Research*, *156*, 185–203. https://doi.org/10.1016/S0079-6123(06)56009-1

Fisher, M., McCoy, K., Poole, J. H., & Vinogradov, S. (2008). Self and other in schizophrenia: A cognitive neuroscience perspective. *American Journal of Psychiatry*, *165*(11), 1465–1472. https://doi.org/10.1176/appi.ajp.2008.07111806

Flavell, J. H. (1979). metacognition and cognitive monitoring: a new area of cognitive-development inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined

actions: A developmental comparison. *Child Development*, *56*(5), 1145–1155. https://doi.org/10.2307/1130229

Fossati, P., Hevenor, S. J., Graham, S. J., Grady, C., Keightley, M. L., Craik, F., & Mayberg, H. (2003). In search of the emotional self: An fMRI study using positive and negative emotional words. *American Journal of Psychiatry*, *160*(11), 1938–1945. https://doi.org/10.1176/appi.ajp.160.11.1938

Fossati, P., Hevenor, S. J., Lepage, M., Graham, S. J., Grady, C., Keightley, M. L., ... Mayberg, H. (2004). Distributed self in episodic memory: Neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage*, *22*(4), 1596–1604. https://doi.org/10.1016/j.neuroimage.2004.03.034

Friedman, D., Cycowicz, Y. M., & Bersick, M. (2005). The late negative episodic memory effect: The effect of recapitulating study details at test. *Cognitive Brain Research*, *23*(2–3), 185–198. https://doi.org/10.1016/j.cogbrainres.2004.10.005

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique*, *51*(1), 6–28. https://doi.org/https://doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R

Frijda, N. H. (2008). The psychologists' point of view. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 68-87). New York, USA: The Guilford Press.

Gallant, S. N., Spaniol, J., & Yang, L. (2019). Age differences in cue utilization during prospective and retrospective memory monitoring. *Psychology and Aging*, *34*(4), 545–557. https://doi.org/10.1037/pag0000352

Gallant, S. N., & Yang, L. (2014). Positivity effect in source attributions of arousal-matched emotional and non-emotional words during item-based directed forgetting. *Frontiers in Psychology*, *5*, 1334. https://doi.org/10.3389/fpsyg.2014.01334

Gao, C., Rosburg, T., Hou, M., Li, B., Xiao, X., & Guo, C. (2016). The role of retrieval mode and retrieval orientation in retrieval practice: insights from comparing recognition memory testing formats and restudying. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(6), 977–990. https://doi.org/10.3758/s13415-016-0446-z

Gardhouse, K., & Anderson, A. K. (2013). Objective and subjective measurements in affective science. In J. Armony & P. Vuilleumier (Eds), *The Cambridge handbook of human affective neuroscience* (pp. 57-81). New York, USA: Cambridge University Press.

Gingerich, A. C., & Sullivan, M. C. (2013). Claiming hidden memories as one's own: A review of inadvertent plagiarism. *European Journal of Cognitive Psychology*, *25*(8), 903–916. https://doi.org/10.1080/20445911.2013.841674

Glisky, E. L., & Marquine, M. J. (2009). Semantic and self-referential processing of positive and negative trait adjectives in older adults. *Memory*, *17*(2), 144–157. https://doi.org/10.1080/09658210802077405

Glisky, E. L., Rubin, S. R., & Davidson, P. S. R. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1131–1146. https://doi.org/10.1037//0278-7393.27.5.1131

Gold, J. J., Smith, C. N., Bayley, P. J., Shrager, Y., Brewer, J. B., Stark, C. E. L., ... Squire, L. R. (2006). Item memory, source memory, and the medial temporal lobe: Concordant findings from fMRI and memory-impaired patients. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(24), 9351–9356. https://doi.org/10.1073/pnas.0602716103

Grilli, M. D., & McFarland, C. P. (2011). Imagine that: Self-imagination improves prospective memory in memory-impaired individuals with neurological damage. *Neuropsychological Rehabilitation*, *21*(6), 847–859. https://doi.org/10.1080/09602011.2011.627263.

Guo, C., Duan, L., Li, W., & Paller, K. A. (2006). Distinguishing source memory and item memory: Brain potentials at encoding and retrieval. *Brain Research*, *1118*(1), 142–154. https://doi.org/10.1016/j.brainres.2006.08.034

Gutchess, A., & Kensinger, E. A. (2018). Shared mechanisms may support mnemonic benefits from self-referencing and emotion. *Trends in Cognitive Sciences*, *22*(8), 712–724. https://doi.org/10.1016/j.tics.2018.05.001

Hadley, C. B., & MacKay, D. G. (2006). Does emotion help or hinder immediate memory? Arousal versus priority-binding mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 79–88. https://doi.org/10.1037/0278-7393.32.1.79

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394–400. https://doi.org/10.1016/S1364-6613(00)01707-1

Hartwig, V., Giovannetti, G., Vanello, N., Lombardi, M., Landini, L., & Simi, S. (2009). Biological effects and safety in magnetic resonance imaging: A review. *International Journal of Environmental Research and Public Health*, *6*(6), 1778–1798. https://doi.org/10.3390/ijerph6061778

Hayes, S. M., Buchler, N., Stokes, J., Kragel, J., & Cabeza, R. (2011). Neural correlates of confidence during item recognition and source memory retrieval: Evidence for both dual-process and strenght

memory theories. *Journal of Cognitive Neuroscience*, *23*(12), 3959–3971. https://doi.org/10.1162/jocn_a_00086

Henson, R. N. A., Shallice, T., & Dolan, R. J. (1999). Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. *Brain*, *122*(7), 1367–1381. https://doi.org/10.1093/brain/122.7.1367

Herbert, C., Pauli, P., & Herbert, B. M. (2011). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Social Cognitive and Affective Neuroscience*, *6*(5), 653–661. https://doi.org/10.1093/scan/nsq082

Hirst, W., & Phelps, E. A. (2016). Flashbulb memories. *Current Directions in Psychological Science*, *25*(1), 36–41. https://doi.org/10.1177/0963721415622487

Holland, A. C., & Kensinger, E. A. (2013). Emotion in episodic memory: The effects of emotional content, emotional state, and motivational goals. In J. Armony & P. Vuilleumier (Eds), *The Cambridge handbook of human affective neuroscience* (pp. 465-488). New York, USA: Cambridge University Press.

Hou, M., Grilli, M. D., & Glisky, E. L. (2019a). Self-Reference enhances memory for multi-element events judged likely to happen in young and older adults. *Memory*, *27*(10), 1451–1461. https://doi.org/10.1080/09658211.2019.1674336

Hou, M., Grilli, M. D., & Glisky, E. L. (2019b). Self-reference enhances relational memory in young and older adults. *Aging, Neuropsychology, and Cognition*, *26*(1), 105–120. https://doi.org/10.1080/13825585.2017.1409333

Hourihan, K. L., & Bursey, E. (2017). A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory*, *25*(1), 35–43. https://doi.org/10.1080/09658211.2015.1122809

Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory & Cognition*, *45*(1), 121–136. https://doi.org/10.3758/s13421-016-0646-3

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*(4), 421–445. https://doi.org/10.1006/jmla.1993.1023

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

James, W. (1884). II. - What is an emotion? *Mind*, *9*(34), 188–205. https://doi.org/10.1093/mind/os-

IX.34.188

James, W. (1890). *The principles of psychology*. New York, USA: Henry Holt and Co.

Jia, X., Gao, C., Cui, L., & Guo, C. (2018). Does emotion arousal influence the benefit received from testing: Insights from neural correlates of retrieval mode effect. *NeuroReport*, *29*(17), 1449–1455. https://doi.org/10.1097/WNR.0000000000001130

Johansson, M., & Mecklinger, A. (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. *Biological Psychology*, *64*(1–2), 91–117. https://doi.org/10.1016/S0301-0511(03)00104-2

Johansson, M., Mecklinger, A., & Treese, A.-C. (2004). Recognition memory for emotional and neutral faces : An event-related potential study. *Journal of Cognitive Neuroscience*, *16*(10), 1840–1853. https://doi.org/10.1162/0898929042947883

Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General*, *117*(4), 390–394. https://doi.org/10.1037/0096-3445.117.4.390

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.1992.4.3.268

Johnson, M. K. (1997). Source monitoring and memory distortion. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *352*(1362), 1733–1745. https://doi.org/10.1098/rstb.1997.0156

Johnson, M. K. (2006). Memory and reality. *American Psychologist*, *61*(8), 760–771. https://doi.org/10.1037/0003-066X.61.8.760

Johnson, M. K., Bush, J. G., & Mitchell, K. J. (1998). Interpersonal reality monitoring: Judging the sources of other people's memories. *Social Cognition, 16*(2), 199–224. https://doi.org/10.1521/soco.1998.16.2.199

Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues*, *40*(2), 33–50. https://doi.org/10.1111/j.1540-4560.1984.tb01092.x

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85. https://doi.org/10.1037/0033-295X.88.1.67

Jurica, P. J., & Shimamura, A. P. (1999). Monitoring item and source information: Evidence for a negative

generation effect in source memory. *Memory & Cognition*, *27*(4), 648–656. https://doi.org/10.3758/BF03211558

Kahn, I., Davachi, L., & Wagner, A. D. (2004). Functional-neuroanatomic correlates of recollection: Implications for models of recognition memory. *Journal of Neuroscience*, *24*(17), 4172–4180. https://doi.org/10.1523/JNEUROSCI.0624-04.2004

Kaida, K., Niki, K., & Born, J. (2015). Role of sleep for encoding of emotional memory. *Neurobiology of Learning and Memory*, *121*, 72–79. https://doi.org/10.1016/j.nlm.2015.04.002

Kappenman, E. S., & Luck, S. J. (2012). ERP components: The ups and downs of brainwave recordings. In Luck, S. J. & Kappenman, E. S. (Eds.), *Oxford library of psychology. The Oxford handbook of event-related potential components* (pp. 407-434). New York, USA: Oxford University Press.

Kappenman, E. S., & Luck, S. J. (2016). Best practices for event-related potential research in clinical populations. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *1*(2), 110–115. https://doi.org/10.1016/j.bpsc.2015.11.007

Keefe, R. S. E., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophrenia Research*, *57*(1), 51–67. https://doi.org/10.1016/S0920-9964(01)00306-1

Keil, A. (2013). Electro- and magneto-electroencephalography in the study of emotion. In J. Armony & P. Vuilleumier (Eds), *The Cambridge handbook of human affective neuroscience* (pp. 107-132). New York, USA: Cambridge University Press.

Kelley, A. W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794. https://doi.org/10.1162/08989290260138672

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*(4), 241–251. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science*, *16*(4), 213–218. https://doi.org/10.1111/j.1467-8721.2007.00506.x

Kensinger, E. A., Addis, D. R., & Atapattu, R. K. (2011). Amygdala activity at encoding corresponds with memory vividness and with memory for select episodic details. *Neuropsychologia*, *49*(4), 663–673. https://doi.org/10.1016/j.neuropsychologia.2011.01.017

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A, & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, *101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., & Giovanello, K. S. (2006). The status of semantic and episodic memory in amnesia. In F. J. Chen (Ed.), *Brain mapping and language*. New York, USA: Nova Science Publishers.

Kensinger, E. A., O'Brien, J. L., Swanberg, K., Garoff-Eaton, R. J., & Schacter, D. L. (2007). The effects of emotional content on reality-monitoring performance in young and older adults. *Psychology and Aging*, *22*(4), 752–764. https://doi.org/10.1037/0882-7974.22.4.752

Kensinger, E. A., & Schacter, D. L. (2005a). Emotional content and reality-monitoring ability: fMRI evidence for the influences of encoding processes. *Neuropsychologia*, *43*(10), 1429–1443. https://doi.org/10.1016/j.neuropsychologia.2005.01.004

Kensinger, E. A., & Schacter, D. L. (2005b). Retrieving accurate and distorted memories: Neuroimaging evidence for effects of emotion. *NeuroImage*, *27*(1), 167–177. https://doi.org/10.1016/j.neuroimage.2005.03.038

Kensinger, E. A., & Schacter, D. L. (2006a). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *The Journal of Neuroscience*, *26*(9), 2564–2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

Kensinger, E. A., & Schacter, D. L. (2006b). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition*, *34*(2), 251–260. https://doi.org/10.3758/BF03193403

Kensinger, E. A., & Schacter, D. L. (2008). Emotional networks in the brain. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 601-617). New York, USA: The Guilford Press.

Klein, E., Caspi, Y., & Gil, S. (2003). The relation between memory of the traumatic event and PTSD: Evidence from studies of traumatic brain injury. *The Canadian Journal of Psychiatry*, *48*(1), 28–33. https://doi.org/10.1177/070674370304800106

Koenig, S., & Mecklinger, A. (2008). Electrophysiological correlates of encoding and retrieving emotional events. *Emotion*, *8*(2), 162–173. https://doi.org/10.1037/1528-3542.8.2.162

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments

of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1352–1364. https://doi.org/10.1037/a0027863

Kuhlmann, B. G., & Touron, D. R. (2017). Relate it! Objective and subjective evaluation of mediator-based strategies for improving source memory in younger and older adults. *Cortex*, *91*, 25–39. https://doi.org/10.1016/j.cortex.2016.11.015

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self-other differences. *Journal of Personality and Social Psychology*, *37*(4), 499–514. https://doi.org/10.1037/0022-3514.37.4.499

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940. https://doi.org/10.1037/a0030811

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The relation between valence and arousal in subjective experience varies with personality and culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

LaBar, K. S., Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*(1), 54-64. https://doi.org/10.1038/nrn1825

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual* (Technical Report No. A-8). Gainesville, FL: University of Florida, NIMH Center for the Study of Emotion and Attention.

Lang, P. J., & Bradley, M. M. (2013). Appetitive and defensive motivation: Goal-directed or goal-determined? *Emotion Review*, *5*(3), 230–234. https://doi.org/10.1177/1754073913477511

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Langeslag, S. J. E., & Van Strien, J. W. (2008). Age differences in the emotional modulation of ERP old/new effects. *International Journal of Psychophysiology*, *70*(2), 105–114. https://doi.org/10.1016/j.ijpsycho.2008.07.022

Larøi, F., Linden, M., & Marczewski, P. (2004). The effects of emotional salience, cognitive effort and meta-cognitive beliefs on a reality monitoring task in hallucination-prone subjects. *British Journal*

of Clinical Psychology, *43*(3), 221–233. https://doi.org/10.1348/0144665031752970

LeDoux, J. E., & Phelps, E. A. (2008). Emotional networks in the brain. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 159-179). New York, USA: The Guilford Press.

Le Bigot, L., Knutsen, D., & Gil, S. (2018). I remember emotional content better, but I'm struggling to remember who said it! *Cognition*, *180*, 52–58. https://doi.org/10.1016/j.cognition.2018.07.001

Lei, X., Wu, T., & Valdes-Sosa, P. A. (2015). Incorporating priors for EEG source imaging and connectivity analysis. *Frontiers in Neuroscience*, *9*, 284. https://doi.org/10.3389/fnins.2015.00284

Leshikar, E. D., & Duarte, A. (2012). Medial prefrontal cortex supports memory accuracy for self-referenced items. *Social Neuroscience*, *7*(2), 126–145. https://doi.org/10.1080/17470919.2011.585242

Leshikar, E. D., & Duarte, A. (2014). Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 236–252. https://doi.org/10.3758/s13415-013-0198-y

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Aging Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition*, *22*(5), 530–554. https://doi.org/10.1521/soco.22.5.530.50767

Levine, L. J., & Pizarro, D. A. (2006). Memory for emotional episodes: The strengths and limits of arousal-based accounts. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 37-58). Oxford, United Kingdom: Blackwell Publishing.

Lewis, P. A., Critchley, H. D., Rotshtein, P., & Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral Cortex*, *17*(3), 742–748. https://doi.org/10.1093/cercor/bhk024

Leynes, P. A. (2002). The effect of specific test queries on source-monitoring event-related potentials. *Brain and Cognition*, *50*(2), 218–233. https://doi.org/10.1016/S0278-2626(02)00505-5

Leynes, P. A., Askin, B., & Landau, J. D. (2017). Visual perspective during remembering: ERP evidence of familiarity-based source monitoring. *Cortex*, *91*, 157–168. https://doi.org/10.1016/j.cortex.2016.12.003

Leynes, P. A., & Crawford, C. J. (2018). Event-related potential (ERP) evidence that encoding focus alters

recollected features. *Brain and Cognition*, *127*, 42–50. https://doi.org/10.1016/j.bandc.2018.09.005

Leynes, P. A., Crawford, J. T., & Bink, M. L. (2005). Interrupted actions affect output monitoring and event-related potentials (ERPs). *Memory*, *13*(7), 759–772. https://doi.org/10.1080/09658210444000377

Leynes, P. A., Grey, J. A., & Crawford, J. T. (2006). Event-related potential (ERP) evidence for sensory-based action memories. *International Journal of Psychophysiology*, *62*(1), 193–202. https://doi.org/10.1016/j.ijpsycho.2006.04.003

Leynes, P. A., & Mok, B. A. (2017). Encoding focus alters diagnostic recollection and event-related potentials (ERPs). *Brain and Cognition*, *117*, 1–11. https://doi.org/10.1016/j.bandc.2017.06.011

Leynes, P. A., & Phillips, M. C. (2008). Event-Related Potential (ERP) Evidence for Varied Recollection During Source Monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 741–751. https://doi.org/10.1037/0278-7393.34.4.741

Liang, J. C., & Preston, A. R. (2017). Medial temporal lobe reinstatement of content-specific details predicts source memory. *Cortex*, *91*, 67–78. https://doi.org/10.1016/j.cortex.2016.09.011

López-Caneda, E., Crego, A., Campos, A. D., González-Villar, A., & Sampaio, A. (2019). The think/no-think alcohol task: A new paradigm for assessing memory suppression in alcohol-related contexts. *Alcoholism: Clinical and Experimental Research*, *43*(1), 36–47. https://doi.org/10.1111/acer.13916

Luauté, J., Morlet, D., & Mattout, J. (2015). BCI in patients with disorders of consciousness: Clinical perspectives. *Annals of Physical and Rehabilitation Medicine*, *58*(1), 29–34. https://doi.org/10.1016/j.rehab.2014.09.015

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). Massachusetts, USA: MIT Press.

Luna, K., Martín-Luengo, B., & Albuquerque, P. B. (2018). Do delayed judgements of learning reduce metamemory illusions? A meta-analysis. *Quarterly Journal of Experimental Psychology*, *71*(7), 1626–1636. https://doi.org/10.1080/17470218.2017.1343362

Lundstrom, B. N., Ingvar, M., & Petersson, K. M. (2005). The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. *NeuroImage*, *27*(4), 824–834. https://doi.org/10.1016/j.neuroimage.2005.05.008

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between

emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

MacKenzie, G., Powell, T. F., & Donaldson, D. I. (2015). Positive emotion can protect against source memory impairment. *Cognition and Emotion*, *29*(2), 236–250. https://doi.org/10.1080/02699931.2014.911145

MacLeod, C. A., & Donaldson, D. I. (2017). Investigating the functional utility of the left parietal ERP old/new effect: Brain activity predicts within but not between participant variance in episodic recollection. *Frontiers in Human Neuroscience*, *11*, 580. https://doi.org/10.3389/fnhum.2017.00580

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 671–685. https://doi.org/10.1037/a0018785

Macrae, C. N., Bodenhausen, G. V., & Calvini, G. (1999). Contexts of cryptomnesia: May the source be with you. *Social Cognition*, *17*(3), 273–297. https://doi.org/10.1521/soco.1999.17.3.273

Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., & Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex*, *14*(6), 647–654. https://doi.org/10.1093/cercor/bhh025

Macri, A., Pavard, A., & Versace, R. (2018). The beneficial effect of contextual emotion on memory: The role of integration. *Cognition and Emotion*, *32*(6), 1355–1361. https://doi.org/10.1080/02699931.2017.1387101

Maddock, R., & Frein, S. (2009). Reduced memory for the spatial and temporal context of unpleasant words. *Cognition and Emotion*, *23*(1), 96–117. https://doi.org/10.1080/02699930801948977

Maddox, G. B., Naveh-Benjamin, M., Old, S., & Kilb, A. (2012). The role of attention in the associative binding of emotionally arousing words. *Psychonomic Bulletin & Review*, *19*(6), 1128–1134. https://doi.org/10.3758/s13423-012-0315-x

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

Mao, W. B., An, S., & Yang, X. F. (2017). The effects of goal relevance and perceptual features on emotional items and associative memory. *Frontiers in Psychology*, *8*, 1223. https://doi.org/10.3389/fpsyg.2017.01223

Mao, X., Wang, Y., Wu, Y., & Guo, C. (2017). Self-referential information alleviates retrieval inhibition of

directed forgetting effects—an ERP evidence of source memory. *Frontiers in Behavioral Neuroscience*, *11*, 187. https://doi.org/10.3389/fnbeh.2017.00187

Mao, X., You, Y., Li, W., & Guo, C. (2015). Emotion impairs extrinsic source memory-An ERP study. *Biological Psychology*, *110*, 182–189. https://doi.org/10.1016/j.biopsycho.2015.07.005

Maratos, E. J., Allan, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, *38*(11), 1452–1465. https://doi.org/10.1016/S0028-3932(00)00061-0

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, *16*(3–5), 437–458. https://doi.org/10.1080/026432999380870

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competitition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

Mathews, A., Ridgeway, V., & Holmes, E. A. (2013). Feels like the real thing: Imagery is both more realistic and emotional than verbal thought. *Cognition and Emotion*, *27*(2), 217–229. https://doi.org/10.1080/02699931.2012.698252

McKague, M., McAnally, K. I., Skovron, M., Bendall, S., & Jackson, H. J. (2012). Source monitoring and proneness to auditory-verbal hallucinations: A signal detection analysis. *Cognitive Neuropsychiatry*, *17*(6), 544–562. https://doi.org/10.1080/13546805.2012.676311

Mecklinger, A., Johansson, M., Parra, M., & Hanslmayr, S. (2007). Source-retrieval requirements influence late ERP and EEG memory effects. *Brain Research*, *1172*(1), 110–123. https://doi.org/10.1016/j.brainres.2007.07.070

Mecklinger, A., Rosburg, T., & Johansson, M. (2016). Reconstructing the past: The late posterior negativity (LPN) in episodic memory studies. *Neuroscience & Biobehavioral Reviews*, *68*, 621–638. https://doi.org/10.1016/j.neubiorev.2016.06.024

Miall, D. S. (1986). Emotion and the self: The context of remembering. *British Journal of Psychology*, *77*(3), 389–397. https://doi.org/10.1111/j.2044-8295.1986.tb02205.x

Milner, B., Squire, L. R., & Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron*, *20*(3), 445–468. https://doi.org/10.1016/S0896-6273(00)80987-3

Minor, G., & Herzmann, G. (2019). Effects of negative emotion on neural correlates of item and source

memory during encoding and retrieval. *Brain Research*, *1718*, 32–45. https://doi.org/10.1016/j.brainres.2019.05.001

Mitchell, K. J. (2016). The cognitive neuroscience of source monitoring. In J. Dunlosky & U. Tauber (Eds.), *The Oxford Handbook of Metamemory* (pp. 425-449). New York, USA: Oxford University Press.

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*(4), 638–677. https://doi.org/10.1037/a0015849

Mitchell, K. J., & MacPherson, S. E. (2017). The cognitive neuroscience of source memory: Moving the ball forward. *Cortex*, *91*, 1–8. https://doi.org/10.1016/j.cortex.2017.04.010

Mitchell, K. J., Raye, C. L., Johnson, M. K., & Greene, E. J. (2006). An fMRI investigation of short-term source memory in young and older adults. *NeuroImage*, *30*(2), 627–633. https://doi.org/10.1016/j.neuroimage.2005.09.039

Mitchell, K. J., Raye, C. L., McGuire, J. T., Frankel, H., Greene, E. J., & Johnson, M. K. (2008). Neuroimaging evidence for agenda-dependent monitoring of different features during short-term source memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 780–790. https://doi.org/10.1037/0278-7393.34.4.780

Moore, M., Shafer, A. T., Bakhtiari, R., Dolcos, F., & Singhal, A. (2019). Integration of spatio-temporal dynamics in emotion-cognition interactions: A simultaneous fMRI-ERP investigation using the emotional oddball task. *NeuroImage*, *202*, 116078. https://doi.org/10.1016/j.neuroimage.2019.116078

Moors, A. (2009). Theories of emotion causation: A review. *Cognition and Emotion, 23*(4), 625-662. https://doi.org/10.1080/02699930802645739

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, *5*(2), 119–124. https://doi.org/10.1177/1754073912468165

Moran, J. M., Macrae, C. N., Heatherton, T. F., Wyland, C. L., & Kelley, W. M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, *18*(9), 1586–1594. https://doi.org/10.1162/jocn.2006.18.9.1586

Morrison, A. P., & Haddock, G. (1997). Cognitive factors in source monitoring and auditory hallucinations. *Psychological Medicine*, *27*(3), 669–679. https://doi.org/10.1017/S003329179700487X

Mugikura, S., Abe, N., Suzuki, M., Ueno, A., Higano, S., Takahashi, S., & Fujii, T. (2010). Hippocampal activation associated with successful external source monitoring. *Neuropsychologia*, *48*(6),

1543–1550. https://doi.org/10.1016/j.neuropsychologia.2010.01.021

Murdock, B. B. (1985). The contributions of Hermann Ebbinghaus. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(3), 469–471. https://doi.org/10.1037/0278-7393.11.3.469

Murphy, N. A., & Isaacowitz, D. M. (2008). Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging*, *23*(2), 263–286. https://doi.org/10.1037/0882-7974.23.2.263

Murty, V. P., Ritchey, M., Adcock, R. A., & LaBar, K. S. (2010). FMRI studies of successful emotional memory encoding: A quantitative meta-analysis. *Neuropsychologia*, *48*(12), 3459–3469. https://doi.org/10.1016/j.neuropsychologia.2010.07.030

Nashiro, K., & Mather, M. (2011). The effect of emotional arousal on memory binding in normal aging and Alzheimer's disease. *American Journal of Psychology*, *124*(3), 301–312.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, *26*, 125–173. https://doi.org/10.1016/S0079-7421(08)60053-5

Niedenthal, P. M. (2008). Emotion concepts. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 587-600). New York, USA: The Guilford Press.

Niedermeyer, E. (2005). Historical aspects. In E. Niedermeyer & da Silva, F. L. (Eds.). *Electroencephalography: Basic principles, clinical applications, and related fields* (5th ed., 1-15). Philadelphia, USA: Lippincott Williams & Wilkins.

Nomi, J. S., Rhodes, M. G., & Cleary, A. M. (2013). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition and Emotion*, *27*(1), 141–149. https://doi.org/10.1080/02699931.2012.679917

Novak, D. L., & Mather, M. (2009). The tenacious nature of memory binding for arousing negative items. *Memory & Cognition*, *37*(7), 945–952. https://doi.org/10.3758/MC.37.7.945

Nowrangi, M. A., Lyketsos, C., Rao, V., & Munro, C. A. (2014). Systematic review of neuroimaging correlates of executive functioning: Converging evidence from different clinical populations. *Journal of Neuropsychiatry and Clinical Neurosciences*, *26*(2), 114–125. https://doi.org/10.1176/appi.neuropsych.12070176

Okon-Singer, H., Hendler, T., Pessoa, L., & Shackman, A. J. (2015). The neurobiology of emotion-cognition interactions: Fundamental questions and strategies for future research. *Frontiers in Human Neuroscience*, *9*, 58. https://doi.org/10.3389/fnhum.2015.00058

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104

Otani, H., Jaffa, M., Libkuman, T. M., Goernert, P. N., & Kato, K. (2012). Does source memory impairment associated with emotionally arousing stimuli occur at encoding or retrieval? *Journal of Cognitive Psychology*, *24*(7), 796–801. https://doi.org/10.1080/20445911.2012.693468

Otani, H., Libkuman, T. M., Goernert, P. N., Kato, K., Migita, M., Freehafer, S. E., & Landow, M. P. (2012). Emotion, directed forgetting, and source memory. *British Journal of Psychology*, *103*(3), 343–358. https://doi.org/10.1111/j.2044-8295.2011.02078.x

Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1543–1547. https://doi.org/10.1037/a0020604

Padovani, T., Koenig, T., Brandeis, D., & Perrig, W. J. (2011). Different brain activities predict retrieval success during emotional and semantic encoding. *Journal of Cognitive Neuroscience*, *23*(12), 4008–4021. https://doi.org/10.1162/jocn_a_00096

Pannu, J. K., & Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychology Review*, *15*(3), 105–130. https://doi.org/10.1007/s11065-005-7091-6

Park, H., Abellanoza, C., Schaeffer, J., & Gandy, K. (2014). Source recognition by stimulus content in the MTL. *Brain Research*, *1553*, 59–68. https://doi.org/10.1016/j.brainres.2014.01.029

Park, H., Uncapher, M. R., & Rugg, M. D. (2008). Effects of study task on the neural correlates of source encoding. *Learning and Memory*, *15*(6), 417–425. https://doi.org/10.1101/lm.878908

Pauly, K., Finkelmeyer, A., Schneider, F., & Habel, U. (2013). The neural correlates of positive self-evaluation and self-related memory. *Social Cognitive and Affective Neuroscience*, *8*(8), 878–886. https://doi.org/10.1093/scan/nss086

Penfield, W., & Milner, B. (1958). Memory defict produced by bilateral lesions in the hippocampal zone. *A.M.A. Archives of Neurology and Psychiatry*, *79*, 475–497. https://doi.org/10.1001/archneurpsyc.1958.02340050003001

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy on internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology*, *10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

Peters, J., & Daum, I. (2009). Frontal but not parietal positivity during source recollection is sensitive to

episodic content. *Neuroscience Letters*, *454*(3), 182–186. https://doi.org/10.1016/j.neulet.2009.03.019

Peters, J., Koch, B., Schwarz, M., & Daum, I. (2007). Domain-specific impairment of source memory following a right posterior medial temporal lobe lesion. *Hippocampus*, *17*(9), 505–509. https://doi.org/10.1002/hipo.20297

Peters, J., Suchan, B., Köster, O., & Daum, I. (2007). Domain-specific retrieval of source information in the medial temporal lobe. *European Journal of Neuroscience*, *26*(5), 1333–1343. https://doi.org/10.1111/j.1460-9568.2007.05752.x

Petrolini, V., & Viola, M. (2020). Core affect dynamics: Arousal as a modulator of valence. *Review of Philosophy and Psychology*, Advance online publication. https://doi.org/10.1007/s13164-020-00474-w

Philippi, C. L., Duff, M. C., Denburg, N. L., Tranel, D., & Rudrauf, D. (2012). Medial PFC damage abolished the self-reference effect. *Journal of Cognitive Neuroscience*, *24*(2), 475–481. https://doi.org/10.1162/jocn_a_00138

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734. https://doi.org/10.1017/S0954579405050340

Quevedo, J., Sant' Anna, M. K., Madruga, M., Lovato, I., De-Paris, F., Kapczinski, F., … Cahill, L. (2003). Differential effects of emotional arousal in short- and long-term memory in healthy adults. *Neurobiology of Learning and Memory*, *79*(2), 132–135. https://doi.org/10.1016/S1074-7427(02)00034-5

Radulovic, J., Lee, R., & Ortony, A. (2018). State-dependent memory: Neurobiological advances and prospects for translation to dissociative amnesia. *Frontiers in Behavioral Neuroscience*, *12*, 259. https://doi.org/10.3389/fnbeh.2018.00259

Raj, V., & Bell, M. A. (2010). Cognitive processes supporting episodic memory formation in childhood: The role of source memory, binding, and executive functioning. *Developmental Review*, *30*(4), 384–402. https://doi.org/10.1016/j.dr.2011.02.001

Ramachandran, V. S., & Blakeslee, V. S. (1998). *Phantoms in the brain: Probing the mysteries of the human mind*. New York, USA: William Morrow.

Raposo, A., Frade, S., & Alves, M. (2016). Framing memories: How the retrieval query format shapes the neural bases of remembering. *Neuropsychologia*, *89*, 309–319. https://doi.org/10.1016/j.neuropsychologia.2016.06.036

Raposo, A., Han, S., & Dobbins, I. G. (2009). Ventrolateral prefrontal cortex and self-initiated semantic elaboration during memory retrieval. *Neuropsychologia*, *47*(11), 2261–2271. https://doi.org/10.1016/j.neuropsychologia.2008.10.024

Raye, C. L., & Johnson, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society*, *15*(6), 405–408. https://doi.org/10.3758/BF03334572

Reisberg, D. (2006). Memory for emotional episodes: The strengths and limits of arousal-based accounts. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 15-36). Oxford, United Kingdom: Blackwell Publishing.

Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 3-41). New York, USA: Oxford University Press.

Rhodes, M. G. (2016). Judgments of learning: Methods, data, and theory. In J Dunlosky & S. K. Tauber (Eds.), *Oxford library of psychology. The Oxford handbook of metamemory* (1st ed., pp. 65–80). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199336746.013.4

Riegel, M., Wierzba, M., Grabowska, A., Jednoróg, K., & Marchewka, A. (2016). Effect of emotion on memory for words and their context. *Journal of Comparative Neurology*, *524*(8), 1636–1645. https://doi.org/10.1002/cne.23928

Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion*, *11*(3), 553-562. https://doi.org/10.1037/a0024246

Rimmele, U., Davachi, L., & Phelps, E. A. (2012). Memory for time and place contributes to enhanced confidence in memories for emotional events. *Emotion*, *12*(4), 834–846. https://doi.org/10.1037/a0028003

Risetti, M., Formisano, R., Toppi, J., Quitadamo, L. R., Bianchi, L., Astolfi, L., ... Mattia, D. (2013). On ERPs detection in disorders of consciousness rehabilitation. *Frontiers in Human Neuroscience*, *7*, 775. https://doi.org/10.3389/fnhum.2013.00775

Ritchey, M., Bessette-Symons, B., Hayes, S. M., & Cabeza, R. (2012). Emotion processing in the aging brain is modulated by semantic elaboration. *Neuropsychologia*, *49*(4), 640–650. https://doi.org/10.1016/j.neuropsychologia.2010.09.009

Ritchey, M., LaBar, K. S., & Cabeza, R. (2011). Emotional memory formation. *Journal of Cognitive Neuroscience*, *23*(4), 757–771. https://doi.org/10.1162/jocn.2010.21487

Ritchey, M., Wang, S. F., Yonelinas, A. P., & Ranganath, C. (2019). Dissociable medial temporal pathways for encoding emotional item and context information. *Neuropsychologia*, *124*, 66–78. https://doi.org/10.1016/j.neuropsychologia.2018.12.015

Riva, V., Cantiani, C., Mornati, G., Gallo, M., Villa, L., Mani, E., ... Molteni, M. (2018). Distinct ERP profiles for auditory processing in infants at-risk for autism and language impairment. *Scientific Reports*, *8*(1), 715. https://doi.org/10.1038/s41598-017-19009-y

Robin, F., Cébron, T., Letellier, M., & Nizard, J. (2018). The effects of emotional content on source monitoring in fibromyalgia patients. *Journal of General Psychology*, *145*(4), 392–414. https://doi.org/10.1080/00221309.2018.1494536

Roediger, H. L. (1985). Remembering Ebbinghaus. *Contemporary Psychology: A Journal of Reviews*, *30*(7), 519–523. https://doi.org/10.1037/023895

Rosenbaum, R. S., Köhler, S., Schacter, D. L., Moscovitch, M., Westmacott, R., Black, S. E., ... Tulving, E. (2005). The case of K.C.: Contributions of a memory-impaired person to memory theory. *Neuropsychologia*, *43*(7), 989–1021. https://doi.org/10.1016/j.neuropsychologia.2004.10.007

Ross, R. S., & Slotnick, S. D. (2008). The hippocampus is preferentially associated with memory for spatial context. *Journal of Cognitive Neuroscience*, *20*(3), 432–446. https://doi.org/10.1162/jocn.2008.20035

Rugg, M. D., Allan, K., & Birch, C. S. (2000). Electrophysiological evidence for the modulation of retrieval orientation. *Journal of Cognitive Neuroscience*, *12*(4), 664–678. https://doi.org/10.1162/089892900562291

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. https://doi.org/10.1016/j.tics.2007.04.004

Rugg, M. D., Henson, R. N. A., & Robb, W. G. K. (2003). Neural correlates of retrieval processing in the prefrontal cortex during recognition and exclusion tasks. *Neuropsychologia*, *41*(1), 40–52. https://doi.org/10.1016/S0028-3932(02)00129-X

Rugg, M. D., Vilberg, K. L., Mattson, J. T., Yu, S. S., Johnson, J. D., & Suzuki, M. (2012). Item memory, context memory and the hippocampus: fMRI evidence. *Neuropsychologia*, *50*(13), 3070–3079. https://doi.org/10.1016/j.neuropsychologia.2012.06.004

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*,

110(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Russell, J. A., & Pratt, G. (1980). A description of the affective quality attributed to environments. *Journal of Personality and Social Psychology*, *38*(2), 311–322. https://doi.org/10.1037//0022-3514.38.2.311

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140–144. https://doi.org/10.1111/j.0963-7214.2004.00293.x

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*(4), 507–514. https://doi.org/10.1016/j.neurobiolaging.2008.09.023

Sander, D. (2013). Models of emotion: The affective neuroscience approach. In J. Armony & P. Vuilleumier (Eds), *The Cambridge handbook of human affective neuroscience* (pp. 5-53). New York, USA: Cambridge University Press.

Schacter, D. L. (1996). *Searching for memory: The brain, the mind, and the past*. New York, USA: Basic Books.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 773–786. https://doi.org/10.1098/rstb.2007.2087

Schmidt, K., Patnaik, P., & Kensinger, E. A. (2011). Emotion's influence on memory for spatial and temporal context. *Cognition and Emotion*, *25*(2), 229–243. https://doi.org/10.1080/02699931.2010.483123

Schmidt, S. R., & Saari, B. (2007). The emotional memory effect: Differential processing or item distinctiveness? *Memory & Cognition*, *35*(8), 1905–1916. https://doi.org/10.3758/BF03192924

Schümann, D., Bayer, J., Talmi, D., & Sommer, T. (2018). Dissociation of immediate and delayed effects of emotional arousal on episodic memory. *Neurobiology of Learning and Memory*, *148*, 11–19. https://doi.org/10.1016/j.nlm.2017.12.007

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, *20*(1), 11–21. https://doi.org/10.1136/jnnp.20.1.11

Senkfor, A. J., Van Petten, C., & Kutas, M. (2008). Enactment versus conceptual coding: Equivalent item memory but different source memory. *Cortex*, *44*(6), 649–664. https://doi.org/10.1016/j.cortex.2007.12.004

Shakow, D. (1930). Hermann Ebbinghaus. *The American Journal of Psychology*, *42*(4), 505–518. https://doi.org/10.2307/1414874

Shany, E., & Berger, I. (2011). Neonatal electroencephalography: Review of a practical approach. *Journal of Child Neurology*, *26*(3), 341–355. https://doi.org/10.1177/0883073810384866

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Simons, J. S., Henson, R. N. A., Gilbert, S. J., & Fletcher, P. C. (2008). Separable forms of reality monitoring supported by anterior prefrontal cortex. *Journal of Cognitive Neuroscience*, *20*(3), 447–457. https://doi.org/10.1162/jocn.2008.20036

Simons, J. S., Peers, P. V., Mazuz, Y. S., Berryhill, M. E., & Olson, I. R. (2010). Dissociation between memory accuracy and memory confidence following bilateral parietal lesions. *Cerebral Cortex*, *20*(2), 479–485. https://doi.org/10.1093/cercor/bhp116

Smith, A. P. R., Stephan, K. E., Rugg, M. D., & Dolan, R. J. (2006). Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron*, *49*(4), 631–638. https://doi.org/10.1016/j.neuron.2005.12.025

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Soares, A. P., Pinheiro, A. P., Costa, A., Frade, C. S., Comesaña, M., & Pureza, R. (2013). Affective auditory stimuli: Adaptation of the International Affective Digitized Sounds (IADS-2) for European Portuguese. *Behavior Research Methods, 45*(4), 1168-1181. https://doi.org/10.3758/s13428-012-0310-1

Soares, A. P., Pinheiro, A. P., Costa, A., Frade, C. S., Comesaña, M., & Pureza, R. (2015). Adaptation of the International Affective Picture System (IAPS) for European Portuguese. *Behavior Research Methods, 47*(4), 1159-1177. https://doi.org/10.3758/s13428-014-0535-2

Soderstrom, N. C., Yue, C. L., & Bjork, E. L. (2016). Metamemory and education. In J Dunlosky & S. K. Tauber (Eds.), *Oxford library of psychology. The Oxford handbook of metamemory* (1st ed., pp. 197–215). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199336746.013.6

Song, Z., Jeneson, A., & Squire, L. R. (2011). Medial temporal lobe function and recognition memory: A novel approach to separating the contribution of recollection and familiarity. *Journal of Neuroscience*, *31*(44), 16026–16032. https://doi.org/10.1523/JNEUROSCI.3012-11.2011

Spaniol, J., Davidson, P. S. R., Kim, A. S. N., Han, H., Moscovitch, M., & Grady, C. L. (2009). Event-

related fMRI studies of episodic encoding and retrieval: Meta-analyses using activation likelihood estimation. *Neuropsychologia*, *47*(8–9), 1765–1779. https://doi.org/10.1016/j.neuropsychologia.2009.02.028

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging*, *10*(4), 527–539. https://doi.org/10.1037/0882-7974.10.4.527

Sprondel, V., Kipp, K. H., & Mecklinger, A. (2011). Developmental changes in item and source memory: evidence from an ERP recognition memory study with children, adolescents, and adults. *Child Development*, *82*(6), 1938–1953. https://doi.org/10.1111/j.1467-8624.2011.01642.x

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, *4*(3), 232–243. https://doi.org/10.1162/jocn.1992.4.3.232

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

Squire, L. R. (2009). The legacy of patient H. M. for neuroscience. *Neuron*, *61*(1), 6–9. https://doi.org/10.1016/j.neuron.2008.12.023.

Squire, L. R.., & Kandel, E. C. (2009). *Memory: From mind to molecules* (2nd ed.). Colorado, USA: Roberts and Company Publishers.

Squire, L. R., & Knowlton, B. J. (1995). Memory, hippocampus, and brain systems. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 825–837). Massachusetts, USA: The MIT Press.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology*, *44*, 453–495. https://doi.org/10.1146/annurev.ps.44.020193.002321

Squire, L. R., & Wixted, J. T. (2011). The cognitive neuroscience of human memory since H. M. *Annual Review of Neuroscience*, *34*(1), 259–288. https://doi.org/10.1146/annurev-neuro-061010-113720

Squire, L. R., & Zola-Morgan, S. (1988). Memory: Brain systems and behavior. *Trends in Neurosciences*, *11*(4), 170–175. https://doi.org/10.1016/0166-2236(88)90144-0

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(24), 13515–13522. https://doi.org/10.1073/pnas.93.24.13515

Staniloiu, A., & Markowitsch, H. J. (2014). Dissociative amnesia. *The Lancet Psychiatry*, *1*(3), 226–241. https://doi.org/10.1016/S2215-0366(14)70279-2

Steinmetz, K. R. M., & Kensinger, E. A. (2009). The effects of valence and arousal on the neural activity leading to subsequent memory. *Psychophysiology*, *46*(6), 1190–1199. https://doi.org/10.1111/j.1469-8986.2009.00868.x.

Sugimori, E., Mitchell, K. J., Raye, C. L., Greene, E. J., & Johnson, M. K. (2014). Brain mechanisms underlying reality monitoring for heard and imagined words. *Psychological Science*, *25*(2), 403–413. https://doi.org/10.1177/0956797613505776

Suzuki, M. (2012). Recollection of episodic memory and the hippocampus: Evidence from fMRI. *Psychologia*, *55*(2), 63–79. https://doi.org/10.2117/psysoc.2012.63

Suzuki, M., Fujii, T., Tsukiura, T., Okuda, J., Umetsu, A., Nagasaka, T., ... Yamadori, A. (2002). Neural basis of temporal context memory: A functional MRI study. *NeuroImage*, *17*(4), 1790–1796. https://doi.org/10.1006/nimg.2002.1303

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta- analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Takahashi, E., Ohki, K., & Miyashita, Y. (2002). The role of the parahippocampal gyrus in source memory for external and internal events. *NeuroReport*, *13*(15), 1951–1956. https://doi.org/10.1097/00001756-200210280-00024

Talmi, D. (2013). Enhanced emotional memory. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Talmi, D., Luk, B. T. C., McGarry, L. M., & Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *Journal of Memory and Language*, *56*(4), 555–574. https://doi.org/10.1016/j.jml.2007.01.002

Talmi, D., & McGarry, L. M. (2012). Accounting for immediate emotional memory enhancement. *Journal of Memory and Language*, *66*(1), 93–108. https://doi.org/10.1016/j.jml.2011.07.009

Tanaka, C., Matsui, M., Uematsu, A., Noguchi, K., & Miyawaki, T. (2012). Developmental trajectories of the fronto-temporal lobes from infancy to early adulthood in healthy individuals. *Developmental Neuroscience*, *34*(6), 477–487. https://doi.org/10.1159/000345152

Tauber, S. K., & Dunlosky, J. (2012). Can older adults accurately judge their learning of emotional information? *Psychology and Aging*, *27*(4), 924–933. https://doi.org/10.1037/a0028447

Tauber, S. K., Dunlosky, J., Urry, H. L., & Opitz, P. C. (2017). The effects of emotion on younger and older adults' monitoring of learning. *Aging, Neuropsychology, and Cognition*, *24*(5), 555–574. https://doi.org/10.1080/13825585.2016.1227423

Taylor, M. J., & Baldeweg, T. (2002). Application of EEG, ERP and intracranial recordings to the

investigation of cognitive functions in children. *Developmental Science*, *5*(3), 318–334. https://doi.org/10.1111/1467-7687.00372

Trott, C. T., Friedman, D., Ritter, W., Fabiani, M., & Snodgrass, J. G. (1999). Episodic priming and memory for temporal source: Event-related potentials reveal age-related differences in prefrontal functioning. *Psychology and Aging*, *14*(3), 390–413. https://doi.org/10.1037/0882-7974.14.3.390

Tsai, L. L., Grant, A. K., Mortele, K. J., Kung, J. W., & Smith, M. P. (2015). A practical guide to MR imaging safety: What radiologists need to know. *Radiographics*, *35*(6), 1722–1737. https://doi.org/10.1148/rg.2015150108

Tu, H.-W., Alty, E. E., & Diana, R. A. (2017). Event-related potentials during encoding: Comparing unitization to relational processing. *Brain Research*, *1667*, 46–54. https://doi.org/10.1016/j.brainres.2017.05.003

Tu, H.-W., & Diana, R. A. (2016). Two are not better than one: Combining unitization and relational encoding strategies. *Journal of Experimental Psychology: Learning Memory & Cognition*, *42*(1), 114–126. https://doi.org/10.1037/xlm0000170

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*(4), 385–398. https://doi.org/10.1037/0003-066x.40.4.385

Tulving, E. (1987). Multiple memory systems and consciousness. *Human Neurobiology*, *6*(2), 67–80.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, *2*(3), 67–70. https://doi.org/10.1111/1467-8721.ep10770899

Tulving, E. (1995). Organization of memory: Quo Vadis? In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 839–847). Massachusetts, USA: The MIT Press.

Tulving, E. (2001). Episodic memory and common sense: How far apart? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*(1413), 1505–1515. https://doi.org/10.1098/rstb.2001.0937

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Turner, M. S., Simons, J. S., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2008). Distinct roles for lateral and medial rostral prefrontal cortex in source monitoring of perceived and imagined events. *Neuropsychologia*, *46*(5), 1442–1453. https://doi.org/10.1016/j.neuropsychologia.2007.12.029

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning

and memory. *Frontiers in Psychology*, *8*, 1454. https://doi.org/10.3389/fpsyg.2017.01454

Uncapher, M. R., Otten, L. J., & Rugg, M. D. (2006). Episodic encoding is more than the sum of its parts: An fMRI investigation of multifeatural contextual encoding. *Neuron*, *52*(3), 547–556. https://doi.org/10.1016/j.neuron.2006.08.011

Undorf, M., & Erdfelder, E. (2015). The relatedness effect on judgments of learning: A closer look at the contribution of processing fluency. *Memory & Cognition*, *43*(4), 647–658. https://doi.org/10.3758/s13421-014-0479-x

Uttl, B. U., Siegenthaler, A. L., & Ohta, N. (2006). Memory and emotion from interdisciplinary perspectives. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 1-12). Oxford, United Kingdom: Blackwell Publishing.

Vallesi, A., & Shallice, T. (2006). Prefrontal involvement in source memory: An electrophysiological investigation of accounts concerning confidence and accuracy. *Brain Research*, *1124*(1), 111–125. https://doi.org/10.1016/j.brainres.2006.09.076

Van Petten, C., Senkfor, A. J., & Newberg, W. M. (2000). Memory for drawings in locations: Spatial source memory and event-related potentials. *Psychophysiology*, *37*(4), 551–564. https://doi.org/10.1017/S0048577200990541

Ventura-Bort, C., Dolcos, F., Wendt, J., Wirkner, J., Hamm, A. O., & Weymar, M. (2017). Item and source memory for emotional associates is mediated by different retrieval processes. *Neuropsychologia*, Advance online publication. https://doi.org/10.1016/j.neuropsychologia.2017.12.015

Voss, J. L., & Federmeier, K. D. (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology*, *48*(4), 532–546. https://doi.org/10.1111/j.1469-8986.2010.01085.x

Wager, T. D., Hernandez, L., Jonides, J., & Lindquist, M. (2007). Elements of functional neuroimaging. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 19–55). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511546396.002

Wang, B. (2012). Females' superiority in item memory, but not source memory for neutral and emotional Chinese words. *Personality and Individual Differences*, *52*(8), 925–929. https://doi.org/10.1016/j.paid.2012.02.005

Wang, B., & Fu, X. (2011). Time course of effects of emotion on item memory and source memory for Chinese words. *Neurobiology of Learning and Memory*, *95*(4), 415–424. https://doi.org/10.1016/j.nlm.2011.02.001

Waring, J. D., & Kensinger, E. A. (2009). Effects of emotional valence and arousal upon memory trade-offs with aging. *Psychology and Aging*, *24*(2), 412–422. https://doi.org/10.1037/a0015526

Watts, S., Buratto, L. G., Brotherhood, E. V., Barnacle, G. E., & Schaefer, A. (2014). The neural fate of neutral information in emotion-enhanced memory. *Psychophysiology*, *51*(7), 673–684. https://doi.org/10.1111/psyp.12211

Wegesin, D. J., Jacobs, D. M., Zubin, N. R., Ventura, P. R., & Stern, Y. (2000). Source memory and encoding strategy in normal aging. *Journal of Clinical and Experimental Neuropsychology*, *22*(4), 455–464. https://doi.org/10.1076/1380-3395(200008)22:4;1-0;FT455

Weis, S., Specht, K., Klaver, P., Tendolkar, I., Willmes, K., Ruhlmann, J., ... Fernandez, G. (2004). Process dissociation between contextual retrieval and item recognition. *NeuroReport*, *15*(18), 2729–2733. https://doi.org/10.3410/f.1022996.265645

Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology*, *35*(1), 81–87. https://doi.org/10.1016/S0167-8760(99)00095-1

Wilding, E. L., & Ranganath, C. (2012). ERP components: The ups and downs of brainwave recordings. In Luck, S. J. & Kappenman, E. S. (Eds.), *Oxford library of psychology. The Oxford handbook of event-related potential components* (pp. 407-434). New York, USA: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0187

Wilding, E. L., Rugg, M. D. (1996). An event-related potential study of recognition with and without retrieval of source. *Brain*, *33*(6), 889–905. https://doi.org/10.1093/brain/119.3.889

Wilding, E. L., & Rugg, M. D. (1997). An event-related potential study of memory for spoken and heard words. *Neuropsychologia*, *35*(9), 1185–1195.

Witherby, A. E., & Tauber, S. K. (2018). Monitoring of learning for emotional faces: how do fine-grained categories of emotion influence participants' judgments of learning and beliefs about memory? *Cognition and Emotion*, *32*(4), 860–866. https://doi.org/10.1080/02699931.2017.1360252

Woroch, B., & Gonsalves, B. D. (2010). Event-related potential correlates of item and source memory strength. *Brain Research*, *1317*, 180–191. https://doi.org/10.1016/j.brainres.2009.12.074

Xie, W., & Zhang, W. (2017). Negative emotion enhances mnemonic precision and subjective feelings of remembering in visual long-term memory. *Cognition*, *166*, 73–83. https://doi.org/10.1016/j.cognition.2017.05.025

Yaoi, K., Osaka, M., & Osaka, N. (2015). Neural correlates of the self-reference effect: Evidence from evaluation and recognition processes. *Frontiers in Human Neuroscience*, *9*, 383.

https://doi.org/10.3389/fnhum.2015.00383

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Yick, Y. Y., Buratto, L. G., & Schaefer, A. (2015). The effects of negative emotion on encoding-related neural activity predicting item and source recognition *Neuropsychologia*, *73*, 48–59. https://doi.org/10.1016/j.neuropsychologia.2015.04.030

Yin, X., Ma, Y., Xu, X., & Yang, H. (2019). The effect of self-referencing on memory for different kinds of source information. *Memory*, *27*(4), 519–527. https://doi.org/10.1080/09658211.2018.1532009

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: An emotional binding account. *Trends in Cognitive Sciences*, *19*(5), 259–267. https://doi.org/10.1016/j.tics.2015.02.009

Yoshimura, S., Ueda, K., Suzuki, S., Onoda, K., Okamoto, Y., & Yamawaki, S. (2009). Self-referential processing of negative stimuli within the ventral anterior cingulate gyrus and right amygdala. *Brain and Cognition*, *69*(1), 218–225. https://doi.org/10.1016/j.bandc.2008.07.010

Zanto, T. P., & Gazzaley, A. (2019). Aging of the frontal lobe. In M. D'Esposito & J. H. Grafman (Eds.), *Handbook of Clinical Neurology* (Vol. 163, pp. 369–389). Amsterdarm, Netherlands: Elsevier B.V. https://doi.org/10.1016/B978-0-12-804281-6.00020-3

Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-Serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology*, *65*(4), 236–244. https://doi.org/10.1027/1618-3169/a000409

Zhou, H., Guo, J., Ma, X., Zhang, M., Liu, L., Feng, L., … Zhong, N. (2017). Self-reference emerges earlier than emotion during an implicit self-referential emotion processing task: Event-related potential evidence. *Frontiers in Human Neuroscience*, *11*, 451. https://doi.org/10.3389/fnhum.2017.00451

Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*, *62*(3), 240–253. https://doi.org/10.1016/j.jml.2009.11.004

CHAPTER II

HOW DOES EMOTION AFFECT SOURCE MEMORY?

A META-ANALYSIS OF BEHAVIORAL EVIDENCE

How does emotion affect source memory? A meta-analysis of behavioral evidence[1,2]

Abstract

The current meta-analysis examined the effects of valence and arousal on source memory accuracy, including the identification of variables that moderate the magnitude and direction of those effects. Forty-eight studies comprising 80 individual experiments (N = 2809 participants) were selected. Three separate analyses focusing on valence and arousal effects (valence-based: negative-neutral; positive-neutral; negative-positive) and other three focusing exclusively on arousal (arousal-based: high-low; medium-low; high-medium) were considered. Whereas no reliable effects of emotion on source memory were observed in the case of valence-based analyses (d_{unb} between - 0.11 and - 0.01), in the case of arousal-based analyses source memory was improved for high and medium arousal (*vs.* low arousal; d_{unb} = 0.27; d_{unb} = 0.49, respectively), with no difference between high and medium arousal stimuli (d_{unb} = - 0.12). Importantly, emotion effects on source memory were moderated by factors such as type of features and of source memory tasks. These factors may account for the discrepant findings typically found in source memory research and should be addressed in future studies.

Keywords: source memory; emotion; arousal; valence; meta-analysis; systematic review.

A robust body of evidence demonstrates that emotions affect cognitive functions, including how we encode, store, and retrieve information in our daily life (e.g., Dolcos et al., 2017; Levine & Pizarro, 2004; Tyng, Amin, Saad, & Malik, 2017). Specifically, memory for emotional stimuli appears to be improved when compared to neutral ones, an effect that has been reported for both immediate and delayed retrieval conditions, spanning from a few minutes to years (D'Argembeau & Van der Linden, 2004; Davidson, Mcfarland, & Glisky, 2006; Kensinger & Schacter, 2006b; Steinmetz & Kensinger, 2009; Waring & Kensinger, 2009; Weymar, Löw, & Hamm, 2011; Wirkner, Ventura-Bort, Schulz, Hamm, & Weymar, 2018; Yick, Buratto, & Schaefer, 2015). This phenomenon, also known as the emotion-enhanced memory (EEM) effect, is not only characterized by enhanced overall accuracy, but also by an enhanced subjective experience of

[1] Pereira, D. R., Teixeira-Santos, A. C., Sampaio, A., & Pinheiro, A. P. (submitted). How does emotion affect source memory? A meta-analysis of behavioral evidence.

[2] The Supplementary Materials associated with this chapter are presented in Appendix A.

remembering details from the study episode (Ford & Kensinger, 2019; Ford, Morris, & Kensinger, 2014; Kensinger & Corkin, 2003; Ochsner, 2000; Ritchey, Wang, Yonelinas, & Ranganath, 2019; Weymar, Löw, Schwabe, & Hamm, 2010). Nonetheless, it remains to be clarified whether this sense of subjective vividness for elements of the study episode is also reflected in objective memory performance.

The study of emotion effects on source memory, that is the memory for the origins of an event (Johnson, Hashtroudi, & Lindsay, 1993), may grant valuable contributions. Of note, emotion may operate differently on source memory compared to the memory for the stimulus itself (i.e., item memory; Johnson et al., 1993; Jurica & Shimamura, 1999). The source monitoring framework, the main theoretical proposal in the field, asserts that the recollection of source-specifying information relies on decision-making processes that may include previous knowledge, stereotypes and beliefs (e.g., "I think it was Mary who told me this, she always has strong convictions about this matter"), as well as the evaluation of different qualitative features of the event (Johnson et al., 1993; Mather, Johnson, & De Leonardis, 1999). Examples of these qualitative features are perceptual (e.g., color; identity of the speaker), spatial, and temporal characteristics, as well as thoughts, mental images, strategies, and feelings that occurred during encoding. The discrimination between these distinct qualitative features has been the focus of research concerning the influence of emotional information on source memory, which yielded mixed findings (see Pereira, Sampaio, & Pinheiro, 2019 for a selective overview). Even though previous meta-analyses have addressed source memory, namely the impact of age-related processes (Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995), the impact of emotion remains to be specified. Accordingly, the proposal of the current study is to fill this gap by presenting an up-to-date systematic review of the effects of emotional stimuli on source memory accuracy.

Emotion effects on source memory

In a typical source memory paradigm, participants are usually invited to learn a list of items while qualitative features of the stimulus are manipulated, such as stimulus color, color or shape surrounding the stimulus, location on the screen, list or block of appearance, as well as modality of presentation (e.g., visual *vs.* auditory), who generated the stimulus (e.g., male *vs.* female; participant-generated *vs.* experimenter-generated), or encoding task (see Appendix of Pereira et al., 2019 for more examples). After a learning period and a delay interval, memory for the qualitative features manipulated during encoding is tested. For example, imagine that during the study phase some items are presented in color A and others in color B (e.g., Dougal, Phelps, & Davachi, 2007; Kensinger & Corkin, 2003; Wang, 2012; Wang & Fu, 2011). During the test phase,

the experimenter can ask the participants if a given item was previously presented in color A or B, or if it is a novel item. Source memory can be tested directly as in the former example or after testing the memory for the item itself, such as in the case of free recall, cued-recall, old/new recognition, or remember/know judgments. With respect to the manipulation of the emotional quality of a stimulus, the dimensional conceptualization of emotion has been widely adopted. It postulates that dimensions such as arousal (i.e., the degree of activation prompted by the stimulus, ranging from low to high) and valence (i.e., the degree of pleasantness of the stimulus, ranging unpleasant to pleasant) characterize our affective experience (Bradley & Lang, 1994; Lang, Greenwald, Bradley, & Hamm, 1993; Russel, 1980). Based on valence and arousal properties, a list of negative, positive and neutral stimuli with distinct levels of arousal might be selected to examine their influence on source memory performance.

So far, studies that examined the effects of emotion on source memory have focused on a variety of source features and provided mixed effects, showing that emotion may enhance, exert no effect or even hinder source memory performance (see Appendix of Pereira et al., 2019). Specifically, considering memory for the perceptual qualities of a stimulus such as font color, some studies revealed that participants are more likely to remember color when the stimulus is emotional (D'Argembeau & Van der Linden, 2004, Experiment 3; Doerksen & Shimamura, 2001, Experiment 1; Kensinger & Corkin, 2003), whereas others failed to identify emotion effects on source memory (D'Argembeau & Van der Linden, 2004, Experiments 1 and 2; Davidson et al., 2006, Experiment 2; Wang & Fu, 2011; but see also MacKenzie, Powell, & Donaldson, 2015, Experiment 2 for an impairment effect for negative stimuli). In contrast, when color is manipulated as the frame or background surronding the stimulus, deleterious effects of emotion on source memory are more consistently observed (Boywitt, 2015, Experiment 1; MacKenzie et al., 2015, Experiment 1; Rimmele, Davachi, Petrov, Dougal, & Phelps, 2011, Experiment 1).

Regarding memory for spatio-temporal features, such as stimulus location on the screen and the list/block of appearance during encoding, most studies showed an advantage of emotional over neutral information (e.g., D'Argembeau & Van der Linden, 2004, Experiment 4; D'Argembeau & Van der Linden, 2005; Rimmele et al., 2011, Experiments 2 and 3; Schmidt, Patnaik, & Kensinger, 2011; Yick et al., 2015). Nonetheless, source memory impairments were observed in the context of negative stimuli (Maddock & Frein, 2009; Mitchell, Mather, Johnson, Raye, & Greene, 2006). Moreover, using pictures with distinct levels of arousal, Boywitt (2015) showed that stimuli with medium arousal ratings were associated with improved source memory for location, whereas stimuli with low or high arousal were associated with the worst source

memory performance. These findings indicate that differences in stimulus arousal may account for the discrepant findings reported before.

The color and spatio-temporal features discussed so far require the discrimination of sources that can be deemed as external to the subject according to the source monitoring framework (Johnson et al., 1993). Nonetheless, some situations require discriminating between internal and external sources, such as when one has to distinguish if something was imagined or seen (i.e., reality monitoring; Johnson, 1988). Research probing emotion effects on reality monitoring is also characterized by mixed findings. For example, when discriminating between seen and imagined events, negative and high arousing stimuli were associated with improved performance (e.g., Kensinger & Schacter, 2006b), but impairment effects were also shown for negative and positive stimuli (e.g., Cook, Hicks, & Marsh, 2007). More consistently, an emotion-related impairment effect was observed when participants had to distinguish whether a given stimulus was generated by themselves or by others (e.g., experimenter; another participant; Le Bigot, Knutsen, & Gil, 2018; McKague, McAnally, Skovron, Bendall, & Jackson, 2012).

Other situations require discriminating between internal sources. In this context, memory for the encoding tasks used in the study phase has been commonly tested (e.g., self-referential *vs.* non-self-referential task), showing no effects of emotion (Ferré, Comesaña, & Guasch, 2019, Experiment 1; Kensinger & Schacter, 2006a; Sharot & Yonelinas, 2008) or deleterious effects (Cook et al., 2007; Ferré et al., 2019, Experiments 2 and 3; X. Mao, You, Li, & Guo, 2015), especially in the case of negative items (Newsome, Dulas, & Duarte, 2012; Otani, Jaffa, Libkuman, Goernert, & Kato, 2012; Otani, Libkuman, et al., 2012). Taken together, the existing studies show that emotion effects on source memory are modulated by the type of source monitoring task (external, internal or reality monitoring; Johnson et al., 1993), as well as by the type of source features assessed in the memory test.

Theoretical approaches to the effects of emotion on source memory

Several theoretical frameworks have offered possible explanations for the mixed results reported above. Most of these proposals consider arousal as the main driving factor of emotion-related effects on memory: Yerkes-Dodson law (Yerkes & Dodson, 1908); Easterbrook's cue-utilization hypothesis (Easterbrook, 1959); priority-binding mechanism (Hadley & MacKay, 2006; MacKay et al., 2004); object-based binding theory (Mather, 2007); arousal-biased competition (ABC) theory (Mather & Sutherland, 2011). Specifically, the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959), one of the most discussed approaches, states

that increased arousal results in the narrowing of attentional resources, which leads to prioritization of salient information whereas peripheral information tends to be disregarded. Consequently, there is a central/peripheral trade-off in memory, which is enhanced for the central details of an episode, but impaired for more secondary details. Several studies have supported this view by showing that central features are more accurately remembered in the context of emotional (*vs.* neutral) events, whereas peripheral features are more accurately remembered in the context of neutral (*vs.* emotional) events (e.g., Burke, Heuer, & Reisberg, 1992; Christianson, 1984, 1992; Christianson & Loftus, 1987, 1991; Kensinger, Garoff-Eaton, & Schacter, 2007). In this context, emotion effects on source memory will depend on whether a detail is central or peripheral, i.e. if it is under the focus of attention or not. The proposal of Yerkes and Dodson (1908), in its turn, posits that increased arousal facilitates attention, memory and other cognitive processes to a certain point. When such point is surpassed, the effects of arousal are no longer beneficial, but rather detrimental. Accordingly, these authors postulated an inverted-U relationship between arousal and performance, the Yerkes-Dodson law: moderate levels of arousal are expected to lead to optimal performance, whereas lower or higher levels of arousal are likely associated with poorer performance. By testing source memory for the location of pictures with different levels of arousal, Boywitt (2015) provided a good example of this inverted-U relationship.

In the context of binding approaches, one of the proposals posits that arousal preferentially activates a priority-binding mechanism that generates associative links between the stimulus and its episodic context, such as when and where the stimulus occurred (Hadley & MacKay, 2006; MacKay et al., 2004). According to this account, source features are more likely to be remembered when stimuli are more arousing. Nonetheless, interference effects that propagate from emotional to neutral stimuli may also occur, particularly when these two stimulus categories are mixed in the same list and presented at fast rates (200 ms/stimulus; Hadley & MacKay, 2006). Under these circumstances highly arousing stimuli can disrupt the binding to adjacent neutral stimuli. The object-based binding theory (Mather, 2007) also states that arousal may enhance binding between an item and some but not all contextual features. Thus, this proposal goes further in the specification of which source details are more likely to be remembered or forgotten. Specifically, arousal grants a memory advantage for source features that are an integral/intrinsic part of the arousing item, i.e., when they are perceived as a coherent and unified entity (e.g., stimulus color; location on the screen). On the contrary, when the source features are extrinsic to the arousing item, i.e., when they can be perceived as distinct elements thereby requiring associations between them (e.g., background color or shape surrounding

the stimulus; encoding task), arousal is more likely to exert a deleterious or no effect on source memory. Taken together, arousal tends to benefit within-item binding whereas it hampers or has no impact on between-item binding. Also, it is possible to draw a parallel between the effects of arousal on intrinsic/extrinsic features and the effects of arousal on central/peripheral details theorized by the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959).

The ABC theory (Mather & Sutherland, 2011) emphasizes the concept of stimulus priority. As information competes for limited attentional resources, stimuli that are perceived as more relevant given the current goals and demands of the subject will be prioritized. Therefore, arousal might facilitate memory for source features that are tagged as high priority, while hampering the memory for features deemed as low priority. Goals may range from idiosyncratic and situational goals to universal goals that are shared by many persons (see Levine & Edelstein, 2009). In this proposal, intrinsic and extrinsic features are similarly relevant. However and contrary to the object-based binding theory (Mather, 2007), the ABC theory postulates that memory for extrinsic features might be enhanced if they are goal-relevant. Accordingly, this account accommodates empirical evidence showing, for instance, that emotion may benefit memory for extrinsic features (e.g., Doerksen & Shimamura, 2001, Experiment 2; Kensinger, O'Brien, Swanberg, Garoff-Eaton, & Schacter, 2007; Kensinger & Schacter, 2006b) or even impair or exert no effect in the case of intrinsic features (e.g., Cook et al., 2007; Dougal et al., 2007; Ferré et al., 2019; Koenig & Mecklinger, 2008; Le Bigot et al., 2018; Minor & Herzmann, 2019; Wang & Fu, 2011).

The theoretical approaches discussed so far emphasized the role of arousal. Notwithstanding, stimulus valence may also affect memory performance, being commonly studied together with arousal. In this context, the contribution of stimulus valence has also been acknowledged (e.g., Kensinger, 2004; Kensinger & Corkin, 2004; Levine & Edelstein, 2009; Steinmetz & Kensinger, 2009). According to the ABC theory (Mather & Sutherland, 2011), source features associated with neutral or emotional stimuli with low arousal might also be prioritized during encoding and better remembered given the current goals of the subject (Levine & Edelstein, 2009; Ochsner, 2000). Specifically, the neurocognitive mechanisms by which valence and arousal modulate memory are hypothesized to be different (Kensinger, 2004; Kensinger & Corkin, 2004): whereas arousal increases the likelihood that a stimulus is attended during encoding and stored in memory through consolidation processes, implicating the activation of the amygdalar-hippocampal network, valence is more likely to engage controlled encoding and elaborative processes associated with the activation of the hippocampus and prefrontal brain regions. Some authors have also pinpointed differences in the encoding

of distinct valence categories. Accordingly, negative stimuli are thought to more strongly recruit occipito-temporal regions associated with sensory processes, whereas frontal regions typically associated with elaborative processes are more strongly activated by positive stimuli (Steinmetz & Kensinger, 2009). Even though theoretical proposals for arousal may accommodate valence-related effects, the discussion of such effects in the context of source memory, particularly considering negative and positive arousing information, has received less attention.

Candidate factors to moderate the effects of emotion on source memory

The complex relationship between emotion and source memory may be moderated by other factors accounting for the mixed results reported in the field. In the current meta-analysis, we considered stimulus-related factors (e.g., type of study material – words *vs.* pictures; sensory modality of presentation – visual *vs.* auditory; mean valence and/or arousal ratings; number of study items), as well as task-related factors such as the type of source memory task (Johnson et al., 1993), type of source memory feature tested during retrieval, type of source memory feature following the object-based binding theory (Mather, 2007), stimulus encoding time, encoding intentionality, list composition, delay interval between study and test, type of source memory test, number of study-test cycles, number of source memory features tested, total number of new items. The identification of these factors was based on prior meta-analyses on item memory (Galli, Vadillo, Sirota, Feurra, & Medvedeva, 2019; Shields, Sazma, McCullough, & Yonelinas, 2017), source memory (Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995), and on the effects of emotion on other memory processes (Hostler, Wood, & Armitage, 2018; Lipinska, Stuart, Thomas, Baldwin, & Bolinger, 2019; Murphy & Isaacowitz, 2008). Additionally, some of the factors were identified following the source monitoring framework (Johnson et al., 1993).

Considering stimulus-related factors such as stimulus type (e.g., words *vs.* pictures), pictures were shown to more effectively communicate the emotional significance of an event compared to words, being more likely to be remembered (e.g., Paivio, Rogers, & Smythe, 1968), and more likely to elicit arousal-related changes in the viewer (Citron, 2012; Herbert, Kissler, Junghöfer, Peyk, & Rockstroh, 2006; Hinojosa, Carretié, Valcárcel, Méndez-Bértolo, & Pozo, 2009; Mathews, Ridgeway, & Holmes, 2013). Furthermore, auditory information appears to be associated with worse memory performance when compared to visual information (e.g., Cohen, Horowitz, & Wolfe, 2009). Considering that source memory tasks commonly involve the manipulation of two source features, each one associated with multiple stimuli (Hou, Grilli, & Glisky,

2019), the number of conflicting cues associated with the same source may also increase proportionally to the number of stimuli, increasing difficulty during source memory decisions.

Considering that differences in arousal might be associated with different outcomes in source memory performance (e.g., Boywitt, 2015; Easterbrook, 1959; Yerkes & Dodson, 1908), and that stimulus arousal and valence may influence source memory through different cognitive and neural mechanisms (Kensinger, 2004; Kensinger & Corkin, 2004), both valence and arousal of the study items should be acknowledged. In the current meta-analysis, the role of these two factors was addressed in three specific ways. First, we considered studies that manipulated stimulus valence by including negative, positive and/or neutral items, as well as studies that focused on stimulus arousal by including items with low, medium or high arousal levels. Second, whenever possible, we extracted data regarding the mean valence and arousal ratings of the study materials to be used as moderators in relation to the effects of emotion on source memory. Third, considering that valence and arousal are usually studied together as negative and positive items are more arousing than neutral ones, this renders difficult the task of disentangling their influence on memory. Thus, as previously suggested by Hostler and colleagues (2018) and depending on the emotion manipulation adopted in each experiment, we coded the experiments based on whether they controlled stimulus arousal across valence categories or not, or on whether they controlled stimulus valence across distinct levels of arousal or not.

The delay interval between encoding and retrieval, list composition, and encoding intentionality are three important task-related factors to consider. Previous studies revealed that the EEM effect is more noticeable and enhanced with the passage of time (Mitchell et al., 2006; Schaefer, Pottage, & Rickart, 2011; Sharot & Phelps, 2004; Sharot & Yonelinas, 2008; Yick et al., 2015; Yonelinas & Ritchey, 2015). These findings support the notion that emotional information is more likely to be consolidated into a more stable memory trace (Hamann, 2001; Kensinger, 2004). Wang and Fu (2011) tested the effects of emotion on source memory for font color using eight different delay intervals (from immediate to two weeks of delay). Emotion failed to modulate source memory and, more importantly, this lack of effect was stable across delay intervals (see also Sharot & Yonelinas, 2008). This finding suggests, at least, that the role of time may not be the same for item and source memory.

Concerning the effect of list composition, prior evidence showed that the EEM effect is dampened when emotional and neutral stimuli are presented in blocked compared to mixed lists (e.g., Barnacle, Montaldi, Talmi, & Sommer, 2016; Cook et al., 2007; Dewhurst & Parry, 2000; Talmi, Luk, McGarry, & Moscovitch,

2007). This suggests that distinctiveness is one of the variables accounting for the EEM effect since emotional information becomes particularly salient relative to neutral information (Talmi, 2013; Talmi & McGarry, 2012). Even though most studies tend to adopt a mixed approach in source memory research, Cook and collaborators (2007; Experiments 2 and 3) examined reality monitoring performance (seen *vs.* heard) for emotional and neutral words in both blocked and mixed conditions. Irrespective of list composition, the authors found that source memory performance was improved for neutral when compared to negative and positive words.

Encoding intentionality, i.e. if participants are instructed to memorize or not the information for a subsequent memory test, has also been a factor considered in source memory research. D'Argembeau and Van der Linden (2004) ran several experiments to test whether the intentional or incidental encoding of source features affected the memory for color and location of emotional and neutral words. They found that memory for stimulus color was improved for emotional compared to neutral words in the incidental learning condition only, whereas memory for stimulus location was also improved for emotional relative to neutral words irrespective of encoding instruction. This finding suggests that encoding intentionality modulates emotion effects on source memory as a function of the source features being tested. Furthermore, the study of D'Argembeau and Van der Linden revealed that the difference between emotional and neutral words is weakened by intentional learning conditions as participants are more likely to engage in effortful encoding. This could be reflected in similar allocation of attentional resources to both emotional and neutral stimuli or in the implementation of specific encoding strategies that strengthen the memory trace for both types of stimuli.

Other task-related factors such as stimulus encoding time, number of study-test cycles, number of source memory features, total number of new items included in the memory test, and type of source memory test may directly affect memory load. Specifically, when participants are given more time to study each item, deeper encoding strategies may be implemented, thereby increasing the likelihood that the item and related details are recalled later (e.g., Murray & Kensinger, 2012). Moreover, as the number of manipulated source features increases (each one being mapped onto different stimuli), the complexity of source memory decisions will possibly increase and affect performance. When studies include more than one study-test cycle (e.g., Dougal et al., 2007; Jia, Gao, Cui, & Guo, 2018; X. Mao et al., 2015; Pereira et al., 2019), they usually intend to decrease memory load and, consequently, increase the prospect of accurate source memory responses. The number of new items included in a memory test may affect the strategy adopted by the

participants and, thereby, their performance (e.g., Vilberg & Rugg, 2009). The type of memory test (e.g., free recall; cued recall; recognition) may also impact upon the source memory decision. Free recall tests are the most challenging as they heavily rely on spontaneous retrieval processes (e.g., Lipinska et al., 2019; Rhodes, Greene, & Naveh-Benjamin, 2019). Thus, if an item is not spontaneously recalled, the associated source features are also less likely to be retrieved. Recognition tests allow participants to check if a particular type of information was presented before or not, which may be used as a cue to recall details from the study episode. In this regard, the moment when participants are required to make a source memory decision is another factor to be considered. The memory test can focus directly on source memory or it can engage other types of memory decisions, such as old/new recognition, remember/know judgments or confidence ratings.

Finally, according to the source monitoring framework (Johnson et al., 1993), the type of source monitoring task (external, internal or reality-monitoring) may affect memory performance. While external source memory decisions are expected to rely on perceptual, spatio-temporal, affective, and semantic features, internal source memory decisions are expected to rely more on cognitive operations (e.g., thoughts; mental images; strategies; feelings) that took place during the encoding episode (Ferguson, Hastroudi, & Johnson, 1992; Johnson, 1988; Johnson & Raye, 1981; Raye & Johnson, 1980). Reality-monitoring tasks that require internal-external source features discrimination are hypothesized to be easier when compared to internal and external source memory tasks, whose source features tend to share similar characteristics (Johnson & Foley, 1984; Raye & Johnson, 1980). Thereby, it is possible to speculate that emotion effects may also vary as a function of the source monitoring task. Considering differences between intrinsic/central and extrinsic/peripheral features postulated by theoretical frameworks such as the object-based binding theory (Mather, 2007) and the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959), the current work also involved categorizing each experiment following this theoretical distinction. Moreover, considering previous studies documenting that the relation between emotion and source memory differs as a function of the source features manipulated during encoding (e.g., Boywitt, 2015; D'Argembeau & Van der Linden, 2004), each experiment was also categorized according to the type of source feature tested.

The current meta-analysis

The relationship between emotion and source memory remains to be specified. The existing studies yielded mixed evidence, which may be related to methodological heterogeneity, including differences in the variables that were manipulated. The main goal of the current study was to provide an up-to-date quantitative

systematic analysis of behavioral studies probing the influence of emotion on source memory accuracy. Furthermore, we aimed to specify the contribution of the different factors described above, which could serve as potential moderators of the dynamic relationship between emotion and source memory.

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009) guidelines were adopted in the current meta-analytic review.

Literature search strategy

An initial database search was conducted on Web of Science, Scopus, Pubmed, Psychology and Behavioral Sciences Collection (through EBSCOhost), and PsycInfo (through ProQuest). The keywords are provided in Table S1 (see Supplementary Material). We also searched for relevant citations in previous reviews discussing the role of emotion on source memory (Bowen, Kark, & Kensinger, 2018; Kensinger, 2007, 2009; Kensinger & Kark, 2018; Kensinger & Schacter, 2016; Mather, 2007; Mather & Sutherland, 2011) and in the list of references of articles subjected to full-text reading. The end search date was the 1st of October of 2019.

Eligibility criteria

During the abstract and full-text screening phase, several eligibility criteria guided the selection of studies. Specifically, full-length articles reporting empirical data and published in peer-reviewed journals were considered, excluding other works such as conference abstracts/proceedings, editorials, book chapters, narrative/qualitative/quantitative reviews, patents. The articles could be written in English, Portuguese, Spanish or French. They should include at least one sample of healthy adults. Non-human animal studies and studies reporting data exclusively for patients, children, adolescents, and elderly samples were not considered. This decision was supported by evidence demonstrating that emotional processing and episodic memory are significantly modulated by age (e.g., childhood and adolescence: Ghetti & Bunge, 2012; Ladouceur, 2012; elderly: Davidson et al., 2006; Nashiro & Mather, 2011a, 2011b; Newsome et al., 2012; Spencer & Raz, 1995; Tyng et al., 2017) and by clinical conditions (e.g., depression: Degl'Innocenti & Bäckman, 1999; Urban, Charles, Levine, & Almeida, 2018; fibromyalgia: Robin, Cébron, Letellier, & Nizard, 2018; schizophrenia: Fairfield et al., 2016; Herbener, 2008). By restraining the sample to healthy young and

middle-age adults, we expected to cover most of the studies focused on emotion and source memory, while reducing the sources of variability in data analysis. In the context of source memory, few studies with affective manipulations have been conducted with children, adolescents, elderly and clinical groups. Furthermore, pharmacological (including the intake of substances such as alcohol) and non-pharmacological interventions (e.g., mnemonic strategies to boost memory) were excluded. Articles were included if the experimental procedure involved a source memory task, as well as a manipulation the emotional properties of the study stimuli, namely valence and/or arousal properties. Thus, studies including short-term memory/working memory paradigms in which participants are required to retain short lists of items for just a few seconds with immediate recall or recognition were not considered (e.g., Borg, Leroy, Favre, Laurent, & Thomas-Antérion, 2011; Mitchell et al., 2006). In a similar fashion, studies testing other type of emotional manipulations such as mood induction and emotional encoding contexts (e.g., presentation of task-irrelevant emotional stimuli during the encoding of target stimuli) were not considered. Also, studies using specific stimulus categories such as taboo and erotic materials (e.g., Hadley & MacKay, 2006; MacKay, Johnson, Graham, & Burke, 2015), spiders or snakes, were not included (e.g., Bell, Mieth, & Buchner, 2017). Furthermore, only studies that provided the necessary information (e.g., *M*; *SD*; *SE*; *N*) to compute effect sizes in the published version or by request to the authors were considered.

Study selection

Figure 1 presents a schematic illustration of the search strategy adopted in the current study. After removal of duplicates, two independent authors, DRP and ACTS, screened each record by title and abstract following the eligibility criteria. The disagreements were discussed between the two authors until a consensus was reached. If any doubts persisted, records were included in the full-text reading phase. Then the full-text screening phase was ensued. The interrater reliability obtained at this stage as measured by the Cohen's kappa was 0.61, which was indicative of a moderate level of agreement (McHugh, 2012). Once again, the lack of consensus was resolved through discussion. The reasons for exclusion can be consulted in Figure 1.

Coding procedure (main and moderator variables)

The main outcome analyzed in the current meta-analysis was the performance in the source memory task reflected in response accuracy. When studies included data for both correct source responses and errors, priority was given to correct responses. Of note, when errors were the only outcome available, means were

multiplied by "- 1", following previous studies (e.g., Sala, Aksayli, Tatlidil, Gondo, & Gobet, 2019). Means, standard deviations/standard errors of the mean, and sample sizes were extracted for each experiment and emotion-related condition. In the case of figures, we used the web-based tool WebPlotDigitizer (https://apps.automeris.io/wpd/; Rohatgi, 2019) to extract the numerical data as this method appears to have a reduced influence on the overall meta-analytic results (Lipinska et al., 2019). If the sources available in the articles were insufficient to obtain the data, authors were contacted and asked to provide the missing information (means and standard deviations).

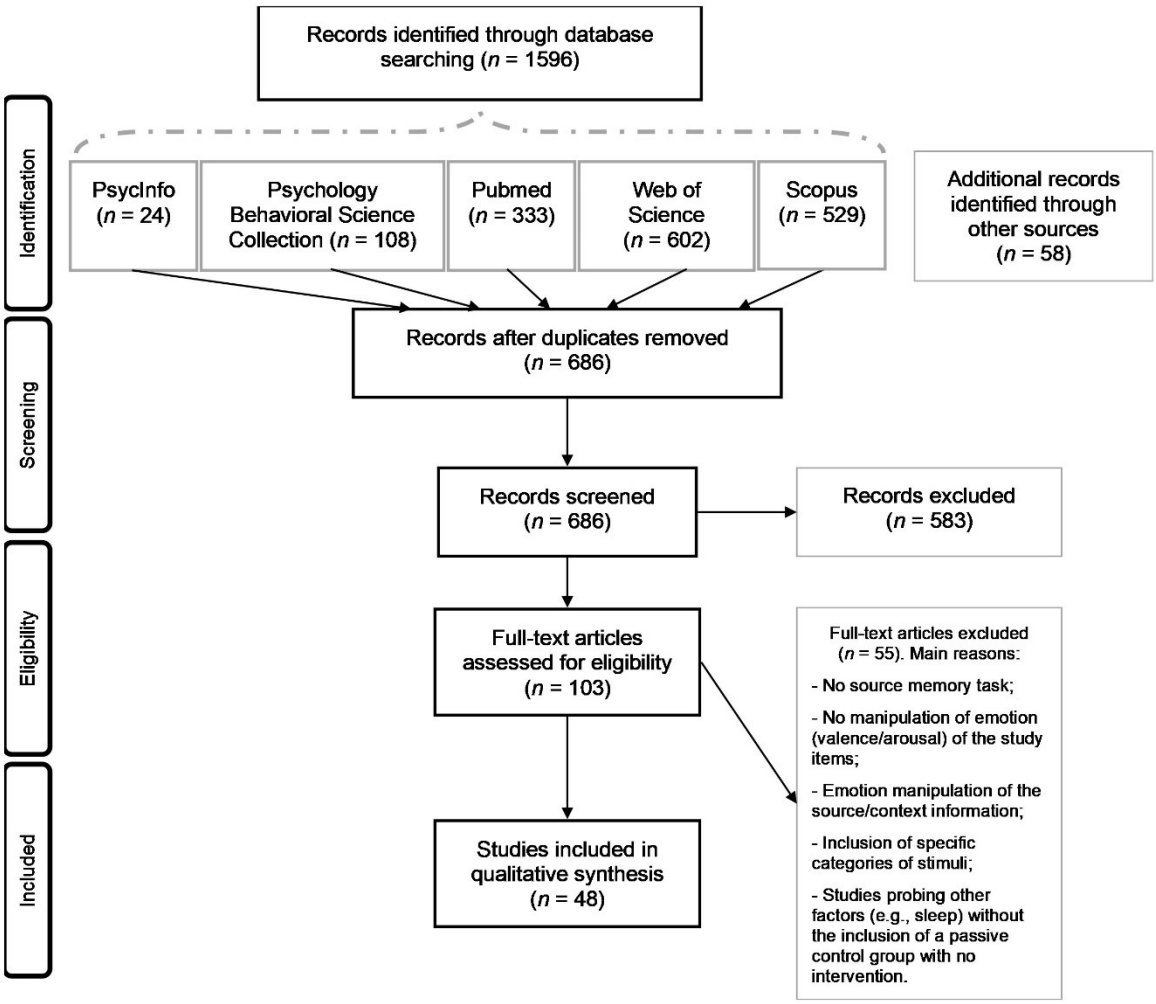


Figure 1. Flow diagram of the search strategy according to the PRISMA guidelines (Moher et al., 2009).

Table 1

Summary of the Moderator Variables Considered in this Study

Moderator variables	
Sociodemographic	Total number of females in the final sample.
	Mean age of the participants.
Materials	Type of study material: words, pictures, both.
	Sensory modality of presentation: visual, auditory, both.
	Mean valence and/or arousal ratings of the stimuli considering each emotional condition.
	The stimuli were controlled for arousal or valence categories? Yes or no.
	Total number of study items.
Task-related conditions	Type of SM task (Johnson et al., 1993): external, internal, reality monitoring.
	Type of SM feature manipulated during the encoding phase and evaluated during the test phase: temporal, spatial (location), stimulus color, background (color; shape; scene), identity of whom generated the stimuli (e.g., self-generated; experimenter-generated; partner-generated; sex of the speaker), modality of stimulus presentation (e.g., seen; heard; imagined), encoding task (e.g., remember or forgot instruction; common or indoor question; scene or people judgment).
	Type of SM feature according to the object-based binding theory (Mather, 2007): extrinsic, intrinsic.
	Stimulus encoding time (in seconds).
	Participants were instructed to overtly memorize the stimuli and/or they knew their memory would be tested at some point in the future: intentional, incidental.
	List composition: mixed, pure.
	Retention interval between the study and the test phase (in minutes).
	Type of SM test: free recall, cued recall, recognition.
	The memory test focused on SM features or evaluated first item memory and then source memory (e.g., old/new recognition; remember/know judgments): direct, indirect.
	Number of study-test cycles.
	Number of SM features tested.
	Total number of new items included in the test phase.

Note. SM = Source Memory.

Information about the selected moderator variables was also considered (see Table 1 for an overview). Additionally, other relevant methodological details were extracted, namely: if the emotional manipulation was achieved between- or within-participants; average years of formal education of the participants; source of the study materials; total number of stimuli used in the experiment; number of stimuli per emotional condition; instruction(s) provided during encoding; time given to the participants during the test phase; how the source memory accuracy was computed. All the information was coded by the first author (see Supplementary Tables S5-S8).

Quality assessment

To assess the methodological quality of the studies included in the review, a list of 12 items based on the Cochrane Collaboration's tool for assessing risk of bias (Higgins & Altman, 2008) and on the study of Thompson et al. (2017) was adapted. The complete list of items is presented in Supplementary Table S9. Each item was rated from zero to two ("0" = not verified; "1" = verified; "2" = unclear, i.e., not enough information to evaluate the item; thus, "1" was the desired score for each item). To obtain a qualitative index for each experiment, the total number of items scored as "1" was divided by the total number of items, and then categorized following Taneja, Olausson, Trulsson, Svensson, and Baad-Hansen (2019): low (≤ 33%); moderate (≥ 34% and ≤ 66%); high (≥ 67%).

Meta-analytic procedure

Since some studies probed valence and arousal effects whereas others focused solely on arousal effects, six separate comparisons were conducted: three comparing valence categories (negative *vs.* neutral; positive *vs.* neutral; negative *vs.* positive) and three comparing different levels of arousal (high *vs.* low; medium *vs.* low; high *vs.* medium). The manipulation of emotion in all studies included in the current meta-analysis followed a within-subjects design. Accordingly, the standardized mean difference was computed based on the *d* unbiased (d_{unb}) estimate (Cumming, 2012) and used as a measure of effect size. This measure is a correction to the Cohen's *d* proposed by Hedges (1981) that attenuates the overestimation bias in studies with small sample sizes. The interpretation is similar to the Cohen's *d*, i.e., 0.8, 0.5, 0.2 indicates a large, moderate, or small effect size, respectively. The sample variance of the effect size was computed according to Borenstein, Hedged, Higgins, and Rothstein (2009; see Supplementary Table S2). However, as the formula requires a correlation coefficient between experimental conditions, which is seldom reported in studies

(Dunlap, Cortina, Vaslow, & Burke, 1996), a correlation of .5 was assumed by default following prior studies (e.g., Galli et al., 2019). Nonetheless, a sensitivity analysis was conducted with the values .3, .5, .7, following the recommendation of Borenstein and colleagues (2009). All formulas and results of the sensitivity analysis are presented in Supplementary Tables S2 and S4, respectively.

Furthermore, since response accuracy data could be extracted for the same sample of participants from distinct experimental conditions, this could represent a problem to conventional meta-analytic procedures by assuming that effect sizes are independent from each other. One way to mitigate this problem is to average effect sizes and sample variances across conditions to obtain one estimate per sample (e.g., Murphy & Isaacowitz, 2008), or to select only one estimate per study (see Assink & Wibbelink, 2016 for other examples). However, such approaches may also lead to loss of information and to biases in the estimation of true effect sizes (Boreinstein et al., 2009). As an alternative, multilevel models account for dependency among multiple observations extracted from a common sample of participants. Therefore, this approach was adopted, in accordance with previous meta-analyses (e.g., Galli et al., 2019; Schweizer et al., 2019; Teixeira-Santos et al., 2019). It is worth noting that a positive effect of negative or positive valence relative to neutral valence was taken as a boosting effect of emotion on source memory, while a negative effect was interpreted as a deleterious effect of emotion on source memory. A similar train of thought was applied to the comparison between negative and positive valence, as well as to the comparisons across levels of arousal.

The analyses reported here were conducted with the metafor package (version 2.1-0; Viechtbauer, 2010) in RStudio (version 1.1.419; RStudio Team, 2018). Databases and script are available at Mendeley Data. First, an influential outcomes analysis was conducted with the influence() function of metafor to identify outliers. Then, outliers were removed from the initial database following previous studies (e.g., Teixeira-Santos et al., 2019). After this analysis, seven influential outcomes were excluded (see Supplementary Table S3). Second, the influence of publication bias was assessed using the Egger's test (Egger, Smith, Schneider, & Minder, 1997) and the rank correlation test (Begg & Mazumdar, 1994) for funnel plot asymmetry, the trim-and-fill method (Duval & Tweedie, 2000), and the Henmi and Copas method (Henmi & Copas, 2010). Nonetheless, these analyses were not performed for two of the comparisons reported here (medium *vs.* low; high *vs.* medium), as tests based on funnel plot asymmetry are not recommended when the number of studies is small (less than 10) and the heterogeneity is large (Ioannidis & Trikalinos, 2007; Sterne et al., 2011). Of note, both the influential outcomes and the publication bias analyses were based on a standard meta-analytic method. Specifically, the rma() function was employed.

Finally, third-level meta-analyses (study sample as the cluster level) were computed using the function rma.mv() from metafor. Of note, a random effect method was selected based on the assumption that the parameters to be estimated are likely to vary from study to study in the population, and that the consequences of applying this method to fixed effect data are more lenient than the other way around (Field & Gillett, 2010; Quintana, 2015). Furthermore, to examine the role of such variables (see Table 1), moderation analyses were conducted keeping the same multilevel approach, in which moderators were included one at a time. When mixed-effects models were statistically significant and in the specific case of categorical variables, sub-group analyses were also computed to supplement the results. Nonetheless, due to the small number of experiments (less than 10) included in two of the three arousal-based analyses (i.e., medium *vs.* low; high *vs.* medium), moderation analyses were not reported for these cases (e.g., Galli et al., 2019; Mancuso, Ilieva, Hamilton, & Farah, 2016). Also, moderators that presented ranges of values, which happened in some instances for valence/arousal ratings of the stimuli, the delay interval between study and test, or the age of the participants, were replaced by the average between minimum and maximum value. The idea was to retain as much information as possible and to attenuate naturally occurring imbalances across moderator levels, which can be problematic (see Field & Gillett, 2010).

Results

The main results of the search strategy are shown in Figure 1. From a total of 1654 records identified through database search and additional sources, 48 studies published between 1997 and 2019 were included in the meta-analysis. These 48 studies comprised a total of 79 individual experiments: 65 probing valence and arousal effects, and 14 focusing on arousal effects only. In the following subsections, we start by reporting the main characteristics and results of the quality assessment considering the 48 studies. Then, results regarding the publication and reporting bias, as well as the main findings are separately described for each comparison.

Characteristics of the studies included

A total of 2809 participants (between 14 and 186 participants per sample; 1802 females) with mean age of 22.42 years and 13.99 years of formal education were included in the pool of studies. Of the 80 experiments, 43 used an external source memory task, 22 used an internal source memory task, whereas 15 used a reality monitoring task. On average, these experiments presented 110 items during the encoding

phase and gave participants four seconds to study each item, even though some of the experiments (n = 10) implemented a self-paced strategy. Most of the experiments (n = 72) presented visual stimuli, three used the auditory modality, and five used both visual and auditory modalities. Also, 38 experiments used pictures as stimuli, 34 used words, seven combined both types of stimuli, and one experiment used faces. Only 15 experiments controlled the arousal/valence ratings of the stimuli across all valence/arousal categories under study, respectively. Emotional and neutral stimuli were intermixed during study and test phases (i.e., mixed list composition) in the vast majority of the experiments, and only two experiments presented stimuli in separate lists per valence (i.e., blocked design). In 47 experiments, participants were not aware that their memory would be tested (i.e., incidental encoding), whereas in 32 experiments participants were overtly instructed to memorize the stimuli (i.e., intentional encoding). The source memory feature most often examined was the encoding task (n = 22), followed by spatial location (n = 20), modality of presentation (n = 11), temporal features (n = 9), stimulus color (n = 9), stimulus background (n = 7), and identity of who presented or generated the stimulus (n = 5). The number of source memory features manipulated during encoding and testing ranged between 2 and 56. Nonetheless, testing two features was the most frequent approach (n = 45). Moreover, whereas 42 experiments manipulated extrinsic features, 41 manipulated intrinsic features. Most of the experiments included only one study-test cycle; only seven included more than one cycle. The delay interval between study and test ranged between immediate to two weeks; participants' memory was commonly assessed by using recognition tests (n = 72; 7 free recall; 1 cued recall). The memory test had, on average, 63 new items. Finally, source memory was directly tested in 42 experiments, yet indirectly tested in 38 experiments.

Quality assessment

For each experiment, the score attributed to each item and the final score are presented in Supplementary Table S9. Following the qualitative assessment described above (Taneja et al., 2019), the methodological quality ranged between 50% and 91.67% (M = 76.46%, SD = 11.89): most experiments (n = 46) were characterized by high methodological quality, whereas 28 were characterized by moderate methodological quality.

Publication bias

The funnel plots for each meta-analysis are displayed in Figure 2. For two of the four comparisons (positive *vs*. neutral; negative *vs*. positive), the assessment of publication bias indicated no major bias (positive *vs*. neutral: $z = 0.65$, $p = .517$; Kendall's $\tau = 0.08$, $p = .319$; negative *vs*. positive: $z = -0.25$, $p = .799$; Kendall's $\tau = -0.06$, $p = .464$; no missing studies were imputed by the trim-and-fill method in the case of negative *vs*. positive, yet 11 missing studies were imputed to the left side in the case of positive *vs*. neutral; in both cases, the estimations derived from the Henmi and Copas method were similar to the estimations provided by the standard random effects meta-analyses). However, regarding the negative *vs*. neutral comparison, the tests yielded evidence of publication bias ($z = 1.81$, $p = .071$; Kendall's $\tau = 0.16$, $p = .028$). In fact, when using the trim-and-fill method as a correction for publication bias, a total of 12 missing studies were imputed to the left side of the funnel plot (see Supplementary Figure S1). With respect to the high *vs*. low arousal, evidence of publication bias was also found ($z = 2.43$, $p = .015$; Kendall's $\tau = 0.28$, $p = .091$). However, the trim-and-fill method did not impute any missing studies to either side of the funnel plot (see Supplementary Figure S1), indicating that the meta-analytic results would remain unaltered. Since heterogeneity may influence the results of publication bias (Ioannidis & Trikalinos, 2007), caution is warranted regarding the meta-analytical results and their interpretations, especially in the case of the negative *vs*. neutral, high *vs*. medium, medium *vs*. low comparisons in the current work (these last two for containing few studies).

Main findings on source memory accuracy

An overview of the main findings for each meta-analysis is provided in Figure 3. The forest plots generated during the meta-analytic analyses are shown in Supplementary Figure S2.

Valence-based effects

Negative *vs*. neutral valence.

There was no statistically significant difference between negative and neutral stimuli considering their influence on source memory performance, $d_{unb} = -0.11$, $SE = 0.07$, $z = -1.63$, $p = .103$, 95% CI [- 0.25, 0.02] (see Supplementary Figure S2A and Table S4). However, the heterogeneity Q test was statistically significant, $Q(91) = 489.59$, $p < .001$, indicating that the true effect sizes vary from study to study (Borenstein et al.,

2009; Viechtbauer, 2010). Such heterogeneity might be partially explained by study-specific factors. Thereby, moderator and subgroup analyses may be useful in this regard (see Table 2).

Considering the impact of sociodemographic-related moderators, namely the number of female participants in the sample and the mean age of participants, only the first emerged as a statistically significant moderator, $Q_M(1) = 8.28$, $p = .004$. Specifically, as the number of female participants increases, the source memory performance for negative stimuli decreases when compared to neutral stimuli, $z = -2.88$, $p = .004$, suggesting a greater impairment effect of negative stimuli on source memory.

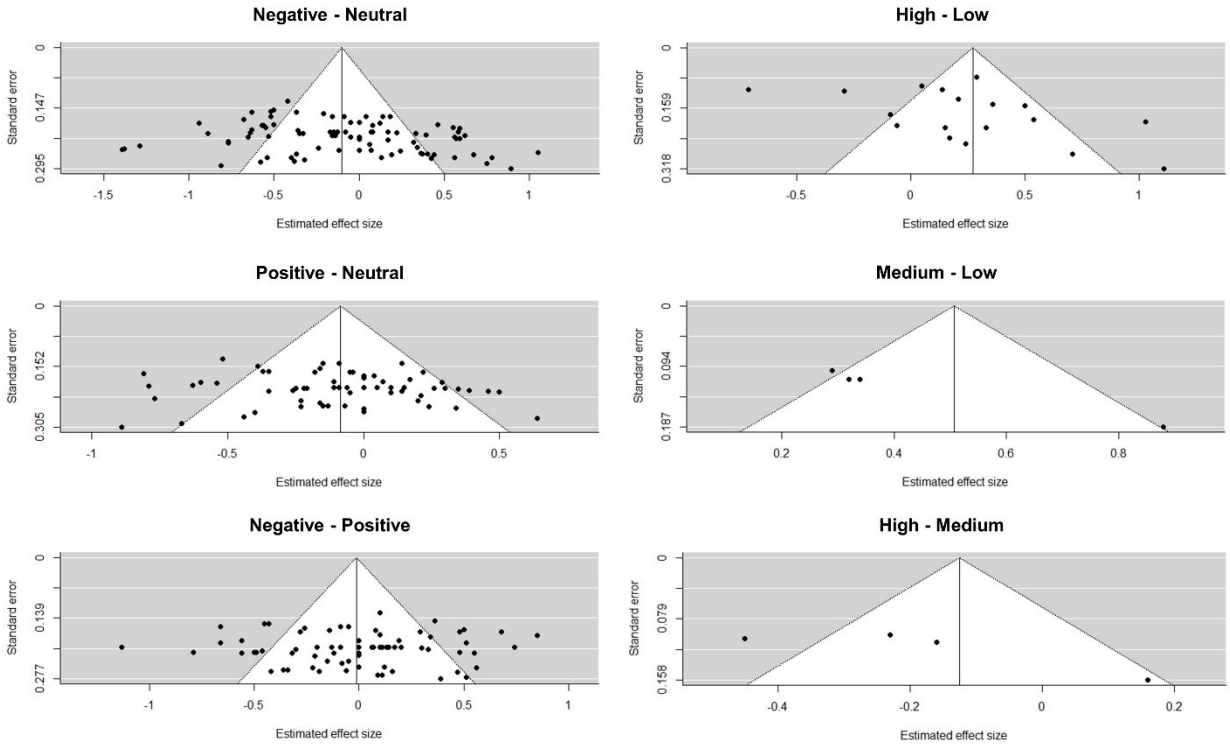


Figure 2. Funnel plot graphs for each analysis (valence and arousal analyses on the left side, and arousal analyses on the right side).

The type of material used in the studies ($Q_M(2) = 7.88$, $p = .019$) also emerged as a statistically significant moderator. Particularly, the mean effect for pictures (- 0.19) and for words (- 0.16) was lower than the mean effect of including both types of stimuli (0.46; $z = -2.77$, $p = .006$; $z = -2.61$, $p = .009$, respectively). Exploratory subgroup analyses showed that whereas the effect of negative relative to neutral information on source memory remained non-significant for pictures ($z = -1.69$, $p = .091$) and words ($z = -1.88$, $p = .060$),

the same was not verified for studies using both types of stimuli ($z = 5.25$, $p < .001$). In this case, the source memory performance was improved for negative relative to neutral information ($d_{unb} = 0.47$, $SE = 0.09$).

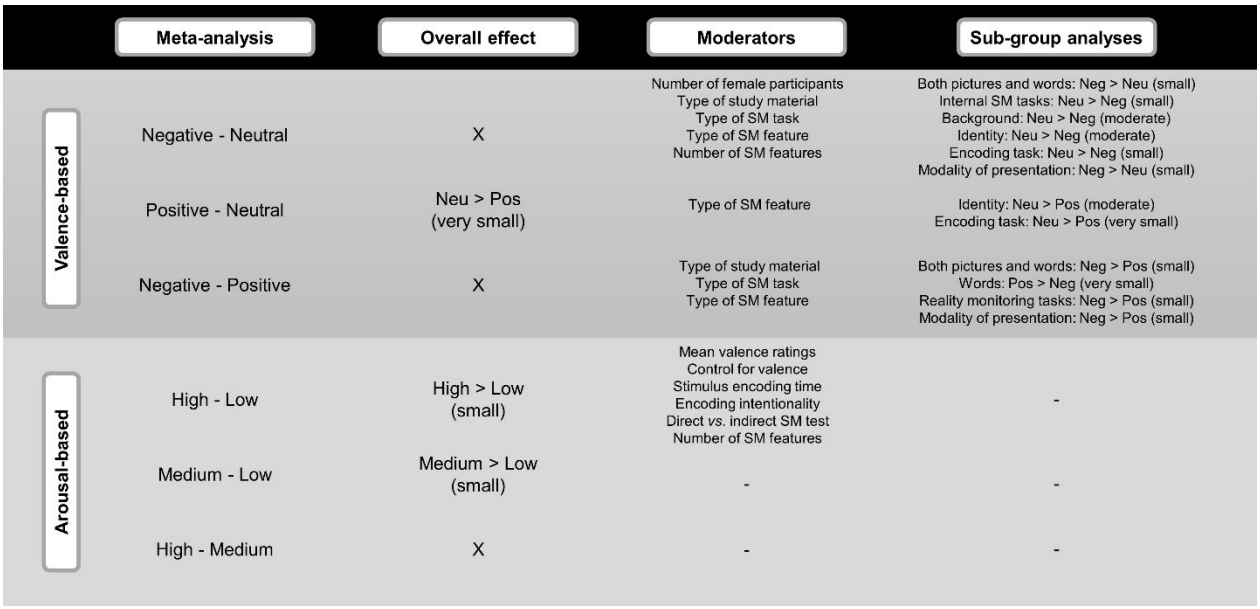


Figure 3. Summary of the main findings obtained for each analysis. Of note, for the moderator and sub-group analyses, only the statistically significant results were presented. Neg = Negative; Neu = Neutral; Pos = Positive; SM = Source Memory.

When probing task-related conditions, the type of SM task ($Q_M(2) = 6.43$, $p = .040$) and the type of source memory feature ($Q_M(6) = 35.06$, $p < .001$) also appeared as moderators of the overall effect of negative emotion on source memory. With respect to the type of task, even though the mean effect of internal source memory tasks (- 0.26) was not significantly lower than the mean effect of external source memory tasks (- 0.14; $z = - 0.77$, $p = .443$), the mean effect of reality monitoring tasks (0.24) was significantly higher than of external source memory tasks ($z = 2.07$, $p = .039$). Furthermore, subgroup analyses revealed no significant changes in the overall effect for external source memory ($z = - 1.28$, $p = .202$) and reality monitoring tasks ($z = 1.44$, $p = .150$). Notwithstanding, in the case of internal source memory, memory performance was improved for neutral compared to negative stimuli ($z = - 3.73$, $p < .001$), even though the effect magnitude was small ($d_{unb} = - 0.25$, $SE = 0.07$).

Regarding the type of source memory feature, the mean effects of color (- 0.12), encoding task (- 0.26), location (- 0.01), modality of presentation (0.49), and temporal features (0.12) were larger than the mean effect of background (- 0.66; see Table 2). No difference was observed between the latter and the mean effect

of identity (- 0.56; z = 0.32, p = .751). Subgroup analyses demonstrated worse source memory for negative relative to neutral stimuli in the subset of studies that manipulated background features (d_{unb} = - 0.67, SE = 0.22, z = - 3.03, p = .002), and an effect of a similar magnitude was found for identity (d_{unb} = - 0.58, SE = 0.15, z = - 3.83, p < .001). This pattern was also observed for encoding task, even though the magnitude was small (d_{unb} = - 0.25, SE = 0.07, z = - 3.73, p < .001). In contrast, a beneficial effect of negative information on source memory was obtained in the subset of studies that manipulated modality of presentation (d_{unb} = 0.49, SE = 0.07, z = - 6.60, p < .001). Also, no effect of negative stimuli on source memory was found for color (z = - 0.53, p = .600), location (z = - 0.11, p = .910), and temporal features (z = 1.27, p = .204).

The number of source memory features manipulated during encoding and tested during retrieval also emerged as a statistically significant moderator, $Q_M(1) = 16.84$, p < .001. As the number of source memory features increased, the source memory accuracy for negative stimuli decreased relative to neutral stimuli (z = - 4.10, p < .001). No other factors emerged as statistically significant moderators (see Table 2).

Positive *vs.* neutral valence.

The overall effect of positive information on source memory accuracy was statistically significant, yet its magnitude was negligible and similar to what was reported for the negative *vs.* neutral meta-analysis, d_{unb} = - 0.11, SE = 0.05, z = - 2.36, p = .018, 95% CI [- 0.20, - 0.02] (see Supplementary Figure S2B and Table S4). Once again, moderator and subgroup analyses may clarify the effects of positive information on source memory.

In this regard, only task-related factors, namely the type of source memory feature ($Q_M(6) = 25.14$, p < .001), emerged as candidate moderators of the overall effect of positive stimuli on source memory. The mean effect of modality of presentation (0.17; z = 1.98, p = .048) was higher than the mean effect of background (- 0.24). No other levels of this moderator reached the threshold of statistical significance (see Table 2). Subgroup analyses showed that the manipulation of the encoding task and of the identity led to worse source memory for positive relative to neutral information (d_{unb} = - 0.18, SE = 0.06, z = - 3.06, p = .002; d_{unb} = - 0.60, SE = 0.14, z = - 4.45, p < .001, respectively). No other statistically significant effects emerged during these analyses (background: z = - 1.30, p = .194; color: z = 0.62, p = .538; location: z = 0.35, p = .730; modality of presentation: z = 1.38, p = .168; temporal: z = 0.57, p = .571). Furthermore, no other moderator analyses were statistically significant (see Table 2).

Negative *vs.* positive valence.

When comparing these two valence categories, the overall effect was not statistically significant, d_{unb} = - 0.01, SE = 0.05, z = - 0.29, p = .775, 95% CI [- 0.10, 0.08], $Q(74)$ = 268.93, p < .001 (see Supplementary Figure S2C and Table S4). The type of study material ($Q_M(2)$ = 24.72, p < .001), type of source memory task ($Q_M(2)$ = 16.68, p < .001), and type of source memory feature ($Q_M(6)$ = 25.81, p < .001) influenced the effect of stimulus valence on source memory.

Similarly to the negative *vs.* neutral comparison, the mean effect for both pictures (0.01) and words (- 0.15) was lower than the mean effect of including both types of stimuli (0.48; z = - 3.60, p < .001; z = - 4.94, p < .001, respectively). Subgroup analyses revealed that for studies using words, negative stimuli hampered source memory performance when compared to positive stimuli; however, the magnitude of the effect was very small (d_{unb} = - 0.15, SE = 0.06, z = - 2.31, p = .021). In contrast, for studies that used both pictures and words, the effect was in the opposite direction and close to moderate in magnitude (d_{unb} = 0.49, SE = 0.11, z = 4.34, p < .001). Additionally, the overall effect for the subset of studies that used pictures did not deviate from zero (d_{unb} = - 0.03, SE = 0.06, z = - 0.44, p = .657).

With respect to the type of source memory task, the mean effect for reality monitoring tasks (0.38) was higher than the mean effect for external tasks (- 0.06; z = 3.63, p < .001). The difference in the mean effect for external and internal source memory tasks was non-significant (z = - 0.61, p = .542). Moreover, the results from the subgroup analyses showed that whereas the overall effect for internal and external source memory remained non-significant (z = - 1.65, p = .099; z = - 1.43, p = .154, respectively), the overall effect concerning reality monitoring was statistically significant, indicating improved source memory accuracy for negative compared to positive stimuli (d_{unb} = 0.38, SE = 0.12, z = 3.23, p = .001).

Even though the moderator analysis focusing on the type of source memory feature was statistically significant, none of the comparisons between the mean effect for background *vs.* other features reached statistical significance (see Table 2). The subgroup analyses revealed that only studies focusing on modality of presentation yielded a significant overall effect, once again showing better source memory performance for negative information (d_{unb} = 0.49, SE = 0.11, z = 4.34, p < .001). The other subgroup analyses remained non-significant (background: z = 1.25, p = .212; color: z = - 1.61, p = .107; encoding task: z = - 1.65, p = .099; identity: z = 0.28, p = .778; location: z = - 1.71, p = .087; temporal features: z = 0.12, p = .903). No other factors emerged as moderators of the effect of emotion on source memory (see Table 2).

Arousal-based effects

High *vs.* low arousal.

Considering studies that manipulated stimulus arousal, a statistically significant (but small) effect of arousal on source memory was observed, indicating improved performance for stimuli with high *vs.* low arousal, d_{unb} = 0.27, *SE* = 0.11, z = 2.42, p = .016, 95% CI [0.05, 0.48], $Q(19)$ = 150.35, p < .001 (see Supplementary Figure S2D and Table S4). Different factors were found to moderate the effect of stimulus arousal on source memory, namely: mean valence ratings of both low and high arousal stimuli ($Q_M(1)$ = 9.43, p = .002; $Q_M(1)$ = 24.20, p < .001, respectively), control for valence across arousal levels ($Q_M(1)$ = 6.06, p = .014), stimulus encoding time ($Q_M(1)$ = 9.35, p = .002), encoding intentionality ($Q_M(1)$ = 5.59, p = .018), direct *vs.* indirect source memory test ($Q_M(1)$ = 6.79, p = .009), and number of source memory features tested ($Q_M(1)$ = 5.37, p = .021). No other factors, including those that were consistently found in the analyses described above, emerged as significant moderators (see Table 3).

Considering factors related to stimulus properties, the moderator analyses revealed that an increase in the mean valence ratings of low arousing stimuli was reflected in an increase of the effect described earlier (3.53; z = 3.07, p = .002), an increase in the mean valence ratings of high arousing stimuli resulted in the opposite pattern (- 0.31; z = - 3.64, p < .001). Furthermore, the mean effect was lower when valence was controlled across stimuli differing in arousal (- 0.20) than when no control was ensured (0.41; z = - 2.46, p = .014).

With respect to task-related characteristics, the moderator analysis revealed that the effect of arousal on source memory was diminished the longer the time provided for stimulus encoding (- 0.22; z = - 3.06, p = .002). Similarly, the mean effect was lower when intentional learning (- 0.06) was used when compared to incidental learning (0.41; z = - 2.36, p = .018). The mean effect was also lower when an indirect source memory test (- 0.05) was used compared to a direct source memory test (0.44; z = - 2.61, p = .009). In contrast, as the number of source memory features under study increased, there was also an enhanced effect of high arousal (*vs.* low arousal) on source memory accuracy (0.07; z = 2.32, p = .021).

Medium *vs.* low arousal.

Similarly to the high *vs.* low arousal comparison, source memory performance was improved for stimuli with medium compared to low arousal levels, d_{unb} = 0.49, *SE* = 0.18, z = 2.65, p = .008, 95% CI [0.13, 0.85],

$Q(4) = 15.04$, $p = .005$ (see Supplementary Figure S2E and Table S4). Due to the reduced number of experiments, no moderator or sub-group analyses were conducted in this case.

High *vs.* medium arousal.

The overall effect did not diverge from zero, $d_{unb} = -0.12$, $SE = 0.15$, $z = -0.82$, $p = .414$, 95% CI [-0.41, 0.17], $Q(4) = 16.16$, $p = .003$ (see Supplementary Figure S2F and Table S4). For the same reason described for the high *vs.* medium arousal comparison, no moderator or sub-group analyses were pursued.

Discussion

The current study aimed to provide an up-to-date meta-analytic review of behavioral studies probing the effects of emotional information on source memory. More specifically, we included studies that manipulated stimulus valence and arousal, and studies focusing on arousal effects only. With respect to valence-based studies, the meta-analytic results showed no reliable effect of stimulus valence on source memory. The positive *vs.* neutral comparison suggested an impairment effect of positive (*vs.* neutral) information on source memory. However, its magnitude was very small (< .2). In the case of arousal-based studies, source memory performance was improved for both high and medium compared to low arousing stimuli. Nonetheless, the effects were also small in magnitude. Furthermore, the comparison between high and medium arousing information did not diverge from zero. Even though the current results may reflect the mixed findings (see Pereira et al., 2019) and methodological heterogeneity that characterize the literature in the field (see Supplementary Tables S5-S8), it also underscores the need to consider the role of moderator variables (see Figure 3).

In the case of studies that manipulated valence, two factors appeared more consistently as moderators: the type of source memory task and the type of source memory feature. Specifically, in studies that used internal source memory tasks, accuracy was improved for neutral compared to negative stimuli (despite a small effect magnitude). Since manipulation of the encoding task is the most frequently used in internal source memory tasks, there was a clear overlap of studies in these two categories. So, it was not surprising that source memory performance for this type of feature was improved for neutral as compared to both negative and positive stimuli. Nonetheless, the magnitude of the effect ranged between trivial and small.

Table 2

Summary of the Results of the Moderator Analyses for the Effects Based on Valence and Arousal

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UP]	Q_E, *p*	Q_M, *p*
Neg-Neu									
N of females	86	49	- 0.02	0.01	- 2.88	.004	- 0.03, - 0.01	375.44, < .001	8.28, .004
Mean age	89	50	0.001	0.01	0.08	.933	- 0.02, 0.02	473.35, < .001	0.007, .933
Type of study material	92	52						440.34, < .001	7.88, .019
Pictures			- 0.65	0.24	- 2.77	.006	- 1.11, - 0.19		
Words			- 0.62	0.24	- 2.61	.009	-1.08, - 0.15		
Both (reference)			0.46	0.22	2.14	.033	0.04, 0.89		
Modality of presentation	92	52						482.95, < .001	2.60, .273
Visual			0.35	0.51	0.69	.491	- 0.65, 1.35		
Auditory (reference)			- 0.49	0.51	- 0.96	.335	- 1.48, 0.50		
Both			0.72	0.56	1.28	.200	- 0.38, 1.83		
Mean valence/arousal ratings									
Neg – mean valence	77	43	0.05	0.03	1.80	.072	- 0.01, 0.11	374.96, < .001	3.23, .072
Neu – mean valence	82	45	0.04	0.03	1.44	.151	- 0.02, 0.10	406.48, < .001	2.07, .151
Neg – mean arousal	76	44	- 0.07	0.06	- 1.16	.244	- 0.19, 0.05	369.19, < .001	1.36, .244
Neu – mean arousal	79	44	- 0.08	0.07	- 1.15	.252	- 0.22, 0.06	403.26, < .001	1.31, .252
Control of arousal	86	50						479.48, < .001	1.81, .179
Yes			- 0.17	0.12	- 1.35	.179	- 0.41, 0.08		
No (reference)			- 0.08	0.08	- 1.09	.275	- 0.23, 0.07		
N of study items	92	52	0.001	0.001	1.00	.318	- 0.001, 0.002	472.74, < .001	1.00, .318
Type of SM task	92	52						465.63, < .001	6.43, .040
External (reference)			- 0.14	0.09	- 1.53	.126	- 0.33, 0.04		
Internal			- 0.12	0.15	- 0.77	.443	- 0.41, 0.18		
Reality			0.39	0.19	2.10	.039	0.02, 0.75		
Type of SM feature	92	52						307.87, < .001	35.06, < .001
Color			0.54	0.23	2.38	.017	0.10, 0.98		
Encoding task			0.40	0.19	2.13	.033	0.03, 0.78		
Identity			0.10	0.32	0.32	.751	- 0.52, 0.73		
Location			0.65	0.21	3.16	.002	0.25, 1.05		
Modality of presentation			1.15	0.22	5.11	< .001	0.71, 1.59		
Temporal			0.78	0.20	3.86	< .001	0.39, 1.18		
Background (reference)			- 0.66	0.16	- 4.04	< .001	- 0.98, - 0.34		

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UP]	Q_E, *p*	Q_M, *p*
Extrinsic *vs.* intrinsic feature	92	52						481.74, < .001	0.96, .327
Extrinsic (reference)			- 0.18	0.10	- 1.86	.063	- 0.36, 0.01		
Intrinsic			0.14	0.14	0.98	.327	- 0.14, 0.41		
Encoding time (seconds)	75	41	- 0.02	0.06	- 0.33	.743	- 0.13, 0.09	397.64, < .001	0.11, .743
Encoding intentionality	90	51						475.72, < .001	0.37, .543
Intentional			0.08	0.13	0.61	.543	- 0.18, 0.34		
Incidental (reference)			- 0.15	0.10	- 1.50	.133	- 0.34, 0.05		
List composition	92	52						489.59, < .001	0.003, .959
Mixed (reference)			- 0.11	0.07	- 1.59	.112	- 0.25, 0.03		
Pure			- 0.03	0.53	- 0.05	.959	- 1.06, 1.01		
Retention interval	84	45	1.77 x 10^{-5}	1.16 x 10^{-5}	1.52	.129	- 5.13 x 10^{-5}, 4.04 x 10^{-5}	428.74, < .001	2.31, .129
Type of SM test	92	52						462.79, < .001	1.49, .476
Free recall			- 0.62	0.53	- 1.17	.242	- 1.66, 0.42		
Cued recall (reference)			0.34	0.48	0.70	.485	- 0.61, 1.28		
Recognition			- 0.44	0.49	- 0.90	.366	- 1.40, 0.51		
Direct *vs.* indirect SM test	92	52						489.32, < .001	0.48, .491
Direct (reference)			- 0.07	0.10	- 0.71	.475	- 0.26, 0.12		
Indirect			- 0.10	0.14	- 0.69	.491	- 0.37, 0.18		
N of study-test cycles	92	52	- 0.02	0.04	- 0.58	.559	- 0.10, 0.05	488.06, < .001	0.34, .559
N of SM features	91	51	- 0.02	0.01	- 4.10	< .001	- 0.04, - 0.01	357.69, < .001	16.84, < .001
N of new items	92	52	0.001	0.001	1.53	.126	- 0.0003, 0.003	459.88, < .001	2.34, .126
Pos-Neu									
N of females	67	34	- 0.006	0.004	- 1.56	.118	- 0.01, 0.002	147.30, < .001	2.45, .118
Mean age	64	32	- 0.01	0.01	- 1.59	.111	- 0.02, 0.002	147.75, < .001	2.54, .111
Type of study material	67	34						158.53, < .001	3.10, .214
Pictures			- 0.26	0.19	- 1.37	.172	- 0.62, 0.11		
Words			- 0.32	0.18	- 1.73	.083	- 0.68, 0.04		
Both (reference)			0.16	0.17	0.93	.352	- 0.18, 0.50		
Modality of presentation	67	34						149.67, < .001	4.55, .103
Visual			0.54	0.29	1.85	.065	- 0.03, 1.11		
Auditory (reference)			- 0.62	0.29	- 2.16	.031	- 1.19, - 0.06		
Both			0.32	0.34	0.94	.346	- 0.35, 0.99		

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UP]	Q_E, *p*	Q_M, *p*
Mean valence/arousal ratings									
Pos – mean valence	55	28	- 0.01	0.02	- 0.45	.650	- 0.06, 0.03	141.51, < .001	0.21, .650
Neu – mean valence	59	29	- 0.003	0.02	- 0.14	.891	- 0.05, 0.05	152.17, < .001	0.02, .891
Pos – mean arousal	51	27	0.02	0.05	0.36	.720	- 0.08, 0.12	128.42, < .001	0.13, .720
Neu – mean arousal	54	27	- 0.01	0.06	- 0.11	.912	- 0.13, 0.12	132.86, < .001	0.01, .912
Control of arousal	59	31						151.27, < .001	1.59, .208
Yes			- 0.13	0.10	- 1.26	.208	- 0.33, 0.07		
No (reference)			- 0.06	0.06	- 1.07	.287	- 0.17, 0.05		
N of study items	67	34	0.0002	0.0004	0.57	.566	- 0.0006, 0.001	162.87, < .0001	0.33, .566
Type of SM task	67	34						149.18, < .001	3.54, .171
External (reference)			- 0.02	0.07	- 0.24	.812	- 0.15, 0.11		
Internal			- 0.17	0.10	- 1.73	.084	- 0.35, 0.02		
Reality			- 0.18	0.14	- 1.26	.206	- 0.45, 0.10		
Type of SM feature	67	34						117.13, < .001	25.14, < .001
Color			0.27	0.17	1.54	.123	- 0.07, 0.60		
Encoding task			0.06	0.16	0.41	.683	- 0.24, 0.37		
Identity			- 0.36	0.21	- 1.74	.081	- 0.76, 0.05		
Location			0.27	0.17	1.57	.117	- 0.07, 0.61		
Modality of presentation			0.41	0.21	1.98	.048	0.004, 0.81		
Temporal			0.33	0.17	1.95	.052	- 0.002, 0.67		
Background (reference)			- 0.24	0.14	- 1.68	.093	- 0.52, 0.04		
Extrinsic *vs.* intrinsic feature	67	34						164.67, < .001	0.04, .834
Extrinsic (reference)			- 0.12	0.07	- 1.83	.067	- 0.25, 0.01		
Intrinsic			0.02	0.10	0.21	.834	- 0.17, 0.21		
Encoding time (seconds)	54	25	- 0.01	0.03	- 0.16	.870	- 0.07, 0.06	126.51, < .001	0.03, .870
Encoding intentionality	65	33						157.22, < .001	0.004, .947
Intentional			0.006	0.10	0.07	.947	- 0.18, 0.20		
Incidental (reference)			- 0.10	0.08	- 1.36	.175	- 0.25, 0.05		
List composition	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Retention interval	58	27	7.04×10^{-6}	1.22×10^{-5}	0.58	.563	-1.68×10^{-5}, 3.09×10^{-5}	143.31, < .0001	0.34, .563
Type of SM test	67	34						141.63, < .001	5.91, .052
Free recall			- 0.49	0.24	- 2.06	.039	- 0.95, - 0.02		
Cued recall (reference)			0.17	0.21	0.81	.421	- 0.24, 0.58		

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UP]	Q_E, *p*	Q_M, *p*
Type of SM test									
Recognition			- 0.25	0.22	- 1.15	.250	- 0.67, 0.18		
Direct *vs.* indirect SM test	67	34						160.45, < .001	1.34, .247
Direct (reference)			- 0.16	0.06	- 2.54	.011	- 0.28, - 0.04		
Indirect			0.11	0.09	1.16	.247	- 0.08, 0.29		
N of study-test cycles	67	34	- 0.01	0.03	- 0.27	.788	- 0.07, 0.05	165.34, < .001	0.07, .788
N of SM features	66	33	0.003	0.01	0.22	.824	- 0.02, 0.03	162.60, < .001	0.05, .824
N of new items	67	34	0.0006	0.001	1.13	.258	- 0.0004, 0.002	157.67, < .0001	1.28, .258
Neg-Pos									
N of females	75	41	- 0.002	0.01	- 0.40	.688	- 0.01, 0.01	268.87, < .001	0.16, .688
Mean age	68	36	0.01	0.01	0.87	.384	- 0.01, 0.02	204.39, < .001	0.76, .384
Type of study material	75	41						188.62, < .001	24.72, < .001
Pictures			- 0.47	0.13	- 3.60	< .001	- 0.72, - 0.21		
Words			- 0.63	0.13	- 4.94	< .001	- 0.88, - 0.38		
Both (reference)			0.48	0.11	4.20	< .001	0.26, 0.70		
Modality of presentation	75	41						261.60, < .001	2.09, .352
Visual			- 0.08	0.24	- 0.34	.732	- 0.56, 0.39		
Auditory (reference)			0.06	0.24	0.25	.804	- 0.41, 0.53		
Both			0.45	0.45	1.01	.311	- 0.42, 1.32		
Mean valence/arousal ratings									
Neg – mean valence	56	30	0.02	0.02	0.94	.350	- 0.02, 0.06	160.15, < .001	0.87, .350
Pos – mean valence	56	30	- 0.01	0.02	- 0.27	.786	- 0.04, 0.03	160.96, < .001	0.07, .786
Neg – mean arousal	50	27	- 0.09	0.05	- 1.72	.085	- 0.19, 0.01	114.73, < .001	2.96, .085
Pos – mean arousal	50	27	- 0.08	0.05	- 1.81	.070	- 0.17, 0.01	113.63, < .001	3.28, .070
Control of arousal	63	35						191.34, < .001	0.78, .377
Yes			- 0.20	0.22	- 0.88	.377	- 0.63, 0.24		
No (reference)			0.12	0.22	0.54	.587	- 0.31, 0.54		
N of study items	75	41	0.001	0.0004	1.41	.160	- 0.0002, 0.001	255.60, < .001	1.98, .160
Type of SM task	75	41						207.44, < .001	16.68, < .001
External (reference)			- 0.06	0.06	- 0.97	.335	- 0.17, 0.06		
Internal			- 0.05	0.09	- 0.61	.542	- 0.23, 0.12		
Reality			0.44	0.12	3.63	< .001	0.20, 0.67		
Type of SM feature	75	41						184.23, < .001	26.49, < .001
Color			- 0.34	0.19	- 1.78	.075	- 0.71, 0.03		

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UP]	Q_E, *p*	Q_M, *p*
Type of SM feature									
Encoding task			- 0.25	0.17	- 1.43	.152	- 0.58, 0.09		
Identity			- 0.08	0.26	- 0.29	.770	- 0.59, 0.44		
Location			- 0.25	0.19	- 1.35	.178	- 0.62, 0.12		
Modality of presentation			0.34	0.20	1.74	.082	- 0.04, 0.73		
Temporal			- 0.05	0.19	- 0.27	.791	- 0.43, 0.33		
Background (reference)			0.14	0.16	0.86	.390	- 0.18, 0.45		
Extrinsic *vs.* intrinsic feature	75	41						259.33, < .001	1.31, .253
Extrinsic (reference)			0.03	0.06	0.52	.603	- 0.09, 0.15		
Intrinsic			- 0.11	0.10	- 1.14	.253	- 0.29, 0.08		
Encoding time (seconds)	61	31	- 0.004	0.01	- 0.71	.478	- 0.01, 0.01	159.84, < .001	0.50, 478
Encoding intentionality	71	39						260.83, < .001	0.34, .561
Intentional			- 0.06	0.10	- 0.58	.561	- 0.25, 0.14		
Incidental (reference)			0.02	0.07	0.20	.844	- 0.13, 0.16		
List composition	75	41						267.83, < .001	0.34, .560
Mixed (reference)			- 0.01	0.05	- 0.19	.853	- 0.10, 0.08		
Pure			- 0.17	0.28	- 0.58	.560	- 0.72, 0.39		
Retention interval	64	32	1.15 x 10^{-5}	1.30 x 10^{-5}	0.88	.378	- 1.40 x 10^{-5}, 3.69 x 10^{-5}	165.95, < .0001	0.78, .378
Type of SM test	75	41						265.69, < .001	0.79, .673
Free recall			- 0.17	0.27	- 0.65	.519	- 0.69, 0.35		
Cued recall (reference)			0.17	0.21	0.79	.428	- 0.25, 0.59		
Recognition			- 0.20	0.22	- 0.89	.376	- 0.63, 0.24		
Direct *vs.* indirect SM test	75	41						268.06, < .001	0.07, .785
Direct (reference)			- 0.0002	0.07	- 0.003	.998	- 0.13, 0.13		
Indirect			- 0.03	0.09	- 0.27	.785	- 0.21, 0.16		
N of study-test cycles	75	41	- 0.01	0.04	- 0.16	.875	- 0.08, 0.06	268.85, < .001	0.03, .875
N of SM features	74	40	- 0.02	0.02	- 1.43	.154	- 0.05, 0.01	239.34, < .001	2.03, .154
N of new items	75	41	0.0001	0.001	0.09	.926	- 0.001, 0.001	268.84, < .0001	0.01, .926

Note. HB = Higher Bound; LB = Lower Bound; n/a = Not Applicable; Neg = Negative; Neu = Neutral; Pos = Positive; SM = Source Memory.

Table 3

Summary of the Results of the Moderator Analyses for the Effects Based on Arousal

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UP]	Q_E, *p*	Q_M, *p*
High-Low									
N of females	17	11	- 0.01	0.003	- 1.40	.163	- 0.01, 0.002	130.47, < .001	1.95, .163
Mean age	16	12	- 0.01	0.05	- 0.24	.810	- 0.11, 0.09	45.80, < .001	0.06, .810
Type of study material	20	14						149.99, < .001	0.09, .957
Pictures			0.05	0.45	0.10	.919	- 0.84, 0.93		
Words			- 0.04	0.51	- 0.08	.934	- 1.04, 0.95		
Both (reference)			0.24	0.43	0.56	.575	- 0.60, 1.08		
Modality of presentation	20	14						146.08, < .001	0.40, .525
Visual			- 0.29	0.46	- 0.64	.525	- 1.20, 0.61		
Auditory (reference)			0.54	0.45	1.21	.227	- 0.34, 1.42		
Both	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Mean valence/arousal ratings									
Low arousal – mean arousal	14	8	- 0.07	0.25	- 0.29	.771	- 0.56, 0.41	135.24, < .001	0.09, .771
High arousal – mean arousal	14	8	0.09	0.13	0.68	.497	- 0.16, 0.33	135.57, < .001	0.46, .497
Low arousal – mean valence	12	7	3.53	1.15	3.07	.002	1.28, 5.78	66.12, < .001	9.43, .002
High arousal – mean valence	12	7	- 0.31	0.09	- 3.64	< .001	- 0.48, - 0.15	62.10, < .001	13.28, < .001
Control of valence	17	12						83.51, < .001	6.06, .014
Yes			- 0.61	0.25	- 2.46	.014	- 1.10, - 0.13		
No (reference)			0.41	0.12	3.45	< .001	0.18, 0.65		
N of study items	20	14	3.57 x 10^{-7}	0.002	0.0002	.9998	- 0.003, 0.003	147.36, < .001	5.12 x 10^{-8}, .9998
Type of SM task	20	14						150.11, < .001	0.004, .949
External (reference)			0.27	0.12	2.26	.024	0.04, 0.50		
Internal	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Reality			- 0.03	0.43	- 0.06	.949	- 0.87, 0.82		
Type of SM feature	20	14						98.15, < .001	4.57, .334
Color			- 0.33	0.35	- 0.93	.350	- 1.01, 0.36		
Encoding task	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Identity			0.87	0.55	1.57	.116	- 0.21, 1.94		
Location			0.70	0.37	1.87	.061	- 0.03, 1.42		
Modality of presentation			0.57	0.52	1.09	.275	- 0.45, 1.58		
Temporal	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UP]	Q_E, *p*	Q_M, *p*
Type of SM feature									
Background (reference)			- 0.33	0.35	- 0.93	.350	- 1.01, 0.36		
Extrinsic *vs.* intrinsic feature	20	14						111.63, < .001	2.10, .148
Extrinsic (reference)			- 0.07	0.25	- 0.28	.780	- 0.57, 0.43		
Intrinsic			0.40	0.28	1.45	.148	- 0.14, 0.95		
Encoding time (seconds)	18	12	- 0.22	0.07	- 3.06	.002	- 0.36, - 0.08	76.92, < .001	9.35, .002
Encoding intentionality	20	14						92.22, < .001	5.59, .018
Intentional			- 0.47	0.20	- 2.36	.018	- 0.86, - 0.08		
Incidental (reference)			0.41	0.11	3.68	< .001	0.19, 0.63		
List composition	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Retention interval	20	14	- 0.00002	0.002	- 0.07	.943	- 0.0004, 0.0004	150.16, < .001	0.005, .943
Type of SM test	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Direct *vs.* indirect SM test	20	14						87.94, < .001	6.79, .009
Direct (reference)			0.44	0.11	3.89	< .001	0.22, 0.66		
Indirect			- 0.49	0.19	- 2.61	.009	- 0.85, - 0.12		
N of study-test cycles	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
N of SM features	20	14	0.07	0.03	2.32	.021	0.01, 0.14	96.75, < .001	5.37, .021
N of new items	20	14	- 0.002	0.003	- 0.75	.453	- 0.008, 0.003	138.07, < .001	0.56, .453

Note. HB = Higher Bound; LB = Lower Bound; n/a = Not Applicable; SM = Source Memory.

Taken together, these findings seem to reflect what has been reported in the literature: an impairment (Cook et al., 2007; Ferré et al., 2019, Experiments 2 and 3; X. Mao et al., 2015; Newsome et al., 2012; Otani, Jaffa, et al., 2012; Otani, Libkuman, et al., 2012) or a lack of effect of emotional stimuli on internal source memory (Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006a; Pereira et al., 2019, Experiment 1; Sharot & Yonelinas, 2008).

In the two comparisons testing negative stimuli, we also observed that the mean effect for reality monitoring tasks was larger when compared to the mean effect of external source memory tasks. Furthermore, there was a small source memory advantage for negative as compared to positive stimuli, which contrasts with overall effects, as well as with effects for external and internal source memory tasks. A closer look into the characteristics of the reality monitoring studies indicated that there were two main manipulations, one based on the identity of who provided the information (Le Bigot et al., 2018; McKague et al., 2012; Morrison & Haddock, 1997), and another, more frequent, based on the modality of presentation involving the discrimination between seen and imagined information (Kensinger, O'Brien, et al., 2007; Kensinger & Schacter, 2006b; Mathews et al., 2013; Robin et al., 2018). Even though visual imagery during encoding is a strategy that benefits memory (e.g., Oliver, Bays, & Zabrucky, 2016), being recommended as a compensatory strategy in clinical settings (e.g., Cicerone et al., 2019), it remains to be clarified, for instance, why these findings were observed in the case of negative but not positive stimuli. In this regard, previous studies demonstrated that when participants receive a visual imagery instruction (*vs.* a verbal processing instruction to focus on stimulus meaning) to process a negative event, they are more likely to report increased self-report state anxiety, whereas the opposite seems to occur when the event is more positive (Holmes & Mathews, 2005, 2010; Holmes, Mathews, Mackintosh, & Dalgleish, 2008). Furthermore, descriptions of the generated images were more likely to incorporate a personal account of the events, including more specific and emotion-related details (Holmes et al., 2008). Taken together, these factors might provide a plausible explanation for the meta-analytic findings regarding reality monitoring, as well as for the fact that the memory discrimination between imagined and seen events seems to be facilitated by the presence of negative information (Kensinger, O'Brien, et al., 2007; Kensinger & Schacter, 2006b; Mathews et al., 2013).

The same explanation can be applied to the comparisons that included negative stimuli, namely: the source memory benefit for negative (*vs.* both neutral and positive) stimuli in studies that manipulated modality of presentation; the moderator effects found for the type of study material. Regarding the latter, the

mean effect associated with pictures and words was found to be smaller than the mean effect of using both type of stimuli. Additionally, subgroup analyses demonstrated that for studies using both type of stimuli, source memory accuracy was improved for negative relative to both neutral and positive stimuli. As these studies overlap with the reality monitoring studies that implemented a manipulation of modality of presentation (i.e., seen *vs.* imagine), it is plausible that this effect also emerges from the interaction between negative valence and visual imagery. However, not all reality monitoring discriminations are benefited by negative valence. In fact, subgroup analyses indicated that studies whose source manipulation was based on the identity of who provided the information reported a deleterious effect of emotion, both negative and positive, on source memory (Le Bigot et al., 2018; Morrison & Haddock, 1997). Such enhancement and impairment effects in the context of reality monitoring may be accounted for by the ABC theory (Mather & Sutherland, 2011). Specifically, when participants were prompted to mentally visualize a specific item, especially a negative one, it is likely that they were more aroused and added more self-related information to the mental image (Holmes & Mathews, 2005; Holmes et al., 2008). Thereby, items and respective mental images were likely tagged as highly relevant during encoding and later used as diagnostic cues to discriminate between sources, thus improving source memory. In contrast, when identity was implemented as a source manipulation, the presence of emotion most likely led to the prioritization of other features over identity, thus resulting in worse source memory performance.

Stimulus background was another type of source memory feature that, in the subgroup analysis, was related to worse source memory performance for negative compared to neutral stimuli. Such effect was expected given the converging evidence from studies that manipulated this feature (Jia et al., 2018; MacKenzie et al., 2015; W. B. Mao, An, & Yang, 2017; Rimmele et al., 2011). Moreover, this finding dovetails with the object-based binding theory (Mather, 2007), as well as the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959), which propose that stimulus background represents an extrinsic/peripheric detail. It is worth noting, however, that the factor intrinsic (central) *vs.* extrinsic (peripheric) did not emerge as a statistically significant moderator in any of the analyses. Furthermore, for specific source features such as stimulus color and spatio-temporal details, which as are usually regarded as intrinsic and expected to be better remembered in the presence of emotion, the current meta-analyses failed to corroborate such beneficial effect. Instead, subgroup analyses indicated that source memory performance was similar for emotional and neutral stimuli in the case of stimulus color, location, and temporal features. Once again, these findings may result from the discrepant results reported in the literature, including impairment (color:

MacKenzie et al., 2015, Experiment 2; location: Maddock & Frein, 2009; temporal: Aupée, 2007; Maddock & Frein, 2009, Experiment 3), enhancement (color: Kensinger & Corkin, 2003; location: Novak & Mather, 2009; Rimmele et al., 2011, Experiment 2; Schmidt et al., 2011; temporal: D'Argembeau & Van der Linden, 2005; Rimmele et al., 2011, Experiment 3; Schmidt et al., 2011, Experiment 1; Yick et al., 2015), and null findings (color: Dougal et al., 2007; Wang & Fu, 2011; location: Koenig & Mecklinger, 2008; temporal: Koenig & Mecklinger, 2008; Minor & Herzmann, 2019; Schmidt et al., 2011, Experiment 2).

Together with the intrinsic *vs*. extrinsic factor, other variables such as delay time and list composition did not alter the overall effects. Specifically, studies that tested distinct delay times showed that the effect of emotional stimuli on source memory was somehow stable across time (Sharot & Yonelinas, 2008; Wang & Fu, 2011). It should be noted, however, that most studies included in the current meta-analytic work relied on immediate recall or a short delay (see Supplementary Tables S5-S8). Therefore, studies with longer delays and that systematically manipulate different delay intervals are necessary, similarly to what has been done for item memory (e.g., Sharot & Phelps, 2004; Wirkner et al., 2018; Yonelinas & Ritchey, 2015). Moreover, few studies have probed the role of list composition (e.g., Cook et al., 2007), and most experiments intermingled emotional and neutral stimuli within the same list (see Supplementary Tables S5-S8). Indeed, for the positive *vs*. neutral and high *vs*. low arousal comparisons, we were unable to introduce list composition as a moderator due to the lack of experiments presenting emotional and neutral stimuli in a blocked design. Thereby, more research is warranted to specify how variables such as distinctiveness (Talmi, 2013; Talmi & McGarry, 2012) may account for the relationship between emotion and source memory.

Three other factors of theoretical interest were average ratings for both stimulus arousal and valence, the control of arousal across valence types in the case of valence-based studies, and of valence in the case of arousal-based studies, as well as encoding intentionality. Whereas for valence-based studies, the mean average ratings and control of arousal across valence types did not interfere with the overall meta-analytic effects, this was not the case for arousal-based studies (high *vs*. low). Specifically, we observed that the beneficial effect of high arousal on source memory performance increased as mean valence ratings of low arousal stimuli increased, and it diminished as the mean valence ratings of high arousing stimuli increased. In this context, it is important to acknowledge that, according to the dimensional account of emotion (Bradley & Lang, 1994; Lang et al., 1993), affective ratings of valence and arousal tend to fit a U-shape distribution with stimuli with lower or higher valence scores being typically rated as more arousing than neutral ones (Lang, Bradley, & Cuthbert, 1998). Therefore, an increase in the mean valence ratings might indicate that

average ratings were within the neutral range (see Supplementary Table S7), also implying a more rigorous control of valence types across arousal levels. Whereas in the case of low arousing stimuli, an increase in mean valence ratings might have enhanced the affective differences between low and high arousing stimuli, in the case of high arousing stimuli a similar tendency might have led to the opposite result. Moreover, this pattern also dovetails with the moderator effect found for the control of valence categories across arousal levels. However, these findings should be interpreted with caution as the high *vs*. low arousal contrast included only a few experiments, and those reporting mean affective ratings were even more scarce.

With respect to encoding intentionality, previous studies suggested that the effect of emotional stimuli on source memory might be detected under incidental but not intentional learning instructions (e.g., D'Argembeau & Van der Linden, 2004; Ferré et al., 2019). Although encoding intentionality did not moderate the effect for valence-based comparisons, in the case of the high *vs*. low comparison the mean effect for incidental conditions was higher than the mean effect for intentional conditions. That is, intentional learning might mitigate emotion effects on source memory. Following the ABC proposal (Mather & Sutherland, 2011), when participants are instructed to memorize the stimuli and they know their memory will be tested at some point, they may be more likely to engage in effortful encoding and to prioritize both emotional and neutral items. As a result, source memory performance may be similar irrespective of the emotional quality of the stimuli. Notwithstanding, it is worth noting that prior studies indicate that encoding intentionality may interact with other factors such as type of source memory features, being more relevant for some but not all features (e.g., D'Argembeau & Van der Linden, 2004; Kensinger & Schacter, 2006b). As the high *vs*. low arousal meta-analysis was composed of less studies and characterized by lower variability in terms of source features (see Supplementary Table S8), this might have contributed to the current pattern of findings.

The number of source memory features manipulated during study and test also changed the magnitude of the overall effect in the cases of high *vs*. low and negative *vs*. neutral meta-analyses. Whereas in the case of the negative *vs*. neutral analysis, an increase in the number of source memory features enhanced the advantage of neutral over negative stimuli, in the case of the high *vs*. low analysis, an increase corresponded to an enhanced effect of arousal on source memory performance. Even though these results seem to be conflicting, one tentative explanation is that an increase in the number of source memory features enhances the direction of the emotion effect on source memory. Accordingly, impairment effects are stronger in the case of negative stimuli, whereas enhancement effects are stronger in the case of high arousing stimuli.

Even though different stimuli- and task-related factors were identified as potential moderators in the current meta-analysis, we also conducted exploratory analyses considering two sociodemographic-related factors: the number of female participants and the mean age of the sample. Nonetheless, we did not expect any of them to influence the emotion effect on source memory as, on the one hand, samples of children, adolescents and elderly were not included and most studies were composed by samples of young adults (see Supplementary Tables S5 and S7) and, on the other hand, no major sex effects were reported in studies that compared female and male participants (D'Argembeau & Van der Linden, 2005; Wang, 2012). Although mean age did not emerge as a moderator in any of the analyses, it was observed that as the number of females increased, the advantage of neutral over negative stimuli on source memory accuracy was enhanced. Previous studies revealed that females tend to present increased activation for negative information in brain regions involved in emotional processing (see Stevens & Hamann, 2012), and that they present an advantage in episodic memory tasks implying verbal abilities (Asperholm, Högman, Rafi, & Herlitz, 2019). However, few studies have addressed sex differences in emotional source memory. Therefore, future studies should more systematically address the role of sex when probing the relationship between emotion and source memory.

Lastly, the overall effect of stimulus arousal on source memory was modulated by stimulus encoding time in the study phase, as well as by the direct *vs.* indirect approach to evaluate source memory in the test phase. As expected, as the encoding time increased, the beneficial effect of high relative to low arousal decreased. This effect may be related to the opportunity to elaborate more on the information during encoding, hence minimizing the memory difference between low and high arousal stimuli. With respect to the direct *vs.* indirect test of source memory, the mean effect was found to be lower when an indirect (*vs.* direct) test was used. Indirect tests may grant a delayed source decision and more retrieval opportunities in contrast to direct tests, since participants are usually asked first about item memory and then about source memory. The observed effect suggests that high arousal is particularly beneficial when there is less time and less recall attempts (e.g., in direct source memory tests). Once more, due to the limited number of studies included in this meta-analysis, these results and interpretations warrant caution and further investigation.

Limitations

First, it is worth noting some factors that might have contributed to publication bias: only English databases were searched; only published studies were included; it was not possible to incorporate all the

eligible studies due to missing data (see Supplementary Tables S5-S8); publication bias tests were not performed for two of the arousal-based comparisons (medium *vs*. low; high *vs*. medium) due to the small number of experiments; there was some evidence of bias in the case of the negative *vs*. neutral comparison. However, it is worth noting that for the remaining comparisons there were few indices of publication bias. Notwithstanding, the publication bias methods used here are not without limitations. Specifically, the performance of these tests is affected when the true effect sizes are heterogeneous (Ioannidis & Trikalinos, 2007; Song, Hooper, & Loke, 2013; van Aert, Wicherts, & van Assen, 2019). Second, considering the heterogeneity among experiments, the moderator and sub-group analyses were of the upmost relevance to better understand the results. However, these analyses require a reasonable amount of experiments and moderator-related data to be robust. In the current work, there was a limited number of experiments, and incomplete information for the selected moderator variables. Thus, even though these analyses were useful to pinpoint specific variables that modulate the relationship between emotion and source memory, they need to be interpreted with caution. Third, the experiment categorization for some of the moderators was not trivial, being the distinction between intrinsic and extrinsic features a good example. Accordingly, it is not possible to rule out the putative influence of the subjectivity underpinning categorization of the studies.

Conclusion

The current study aimed to provide a systematic overview on the effects of emotional stimuli on source memory, and to identify factors that modulate the magnitude and direction of such effects. For valence-based analyses (i.e., negative *vs*. neutral; positive *vs*. neutral; negative *vs*. positive), an impairment effect of positive relative to neutral stimuli on source memory was observed. However, the magnitude of this effect was very small. Therefore, no robust effects of emotional stimuli on source memory emerged from these analyses. Notwithstanding, both the type of source memory task and the type of source memory features were the factors that more consistently affected the former effects. On the contrary, for arousal-based analyses (i.e., high *vs*. low; medium *vs*. low; high *vs*. medium), the results revealed that high and medium arousal stimuli granted a source memory advantage when compared to low arousal stimuli. No difference was observed between high and medium arousal levels. Nonetheless, caution regarding these findings is required due to the reduced number of experiments that were included in each meta-analysis.

Taken together, this study demonstrated that mixed findings are likely to be expected when probing the effects of emotional stimuli on source memory. However, it also highlights the need to consider the role

of moderator variables, such as the type of source memory task according to the source monitoring framework (Johnson et al., 1993), and the type of source memory features. Furthermore, although stimulus arousal is proposed as the main driver of emotion-related effects by most of the theoretical frameworks, the vast majority of the experiments also include a valence manipulation. Thereby, the role of valence in the context of source memory should also be addressed. In this regard, the ABC theory (Mather & Sutherland, 2011) appears to be the theoretical account that more flexibly accommodates not only arousal but also valence-related effects. Nevertheless, the distinction between low and high priority features, or between goal-relevant or goal-irrelevant features is not straightforward. Taken together, more research is needed to support theoretical claims, to determine the role of valence, and to reach a broader understanding of the effects of emotion on source memory. This includes studying the effects of emotion based not only on dimensional approaches (Bradley & Lang, 1994; Lang et al., 1993; Russel, 1980) but also on categorical approaches (Levine & Pizarro, 2004), as well as the effects derived from mood or emotional encoding contexts (e.g., Gingerich & Dodson, 2013; Palumbo, Mammarella, Di Domenico, & Fairfield, 2018; Pereira, Sampaio, & Pinheiro, 2020; Ventura-Bort et al., 2017) for which the existing evidence is still limited.

References

**References marked with an asterisk indicate studies included in the meta-analysis*

Asperholm, M., Högman, N., Rafi, J., & Herlitz, A. (2019). What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychological Bulletin*, *145*(8), 785–821. https://doi.org/10.1037/bul0000197

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

*Aupée, A. M. (2007). A detrimental effect of emotion on picture recollection. *Scandinavian Journal of Psychology*, *48*(1), 7–11. https://doi.org/10.1111/j.1467-9450.2006.00533.x

Barnacle, G. E., Montaldi, D., Talmi, D., & Sommer, T. (2016). The list-composition effect in memory for emotional and neutral pictures: Differential contribution of ventral and dorsal attention networks to successful encoding. *Neuropsychologia*, *90*, 125–135. https://doi.org/10.1016/j.neuropsychologia.2016.06.023

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias.

Biometrics, *50*(4), 1088–1101. https://doi.org/10.2307/2533446

Bell, R., Mieth, L., & Buchner, A. (2017). Emotional memory: No source memory without old-new recognition. *Emotion*, *17*(1), 120–130. https://doi.org/10.1037/emo0000211

*Blix, I., & Brennen, T. (2011). Intentional forgetting of emotional words after trauma: A study with victims of sexual assault. *Frontiers in Psychology*, *2*, 235. https://doi.org/10.3389/fpsyg.2011.00235

Borenstein, M., Hedged, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, USA: Routledge.

Borg, C., Leroy, N., Favre, E., Laurent, B., & Thomas-Antérion, C. (2011). How emotional pictures influence visuospatial binding in short-term memory in ageing and Alzheimer's disease? *Brain and Cognition*, *76*(1), 20–25. https://doi.org/10.1016/j.bandc.2011.03.008

Bowen, H. J., Kark, S. M., & Kensinger, E. A. (2018). NEVER forget: Negative emotional valence enhances recapitulation. *Psychonomic Bulletin and Review*, *25*(3), 870–891. https://doi.org/10.3758/s13423-017-1313-9

*Boywitt, C. D. (2015). Non-monotonic relationships between emotional arousal and memory for color and location. *Cognition and Emotion*, *29*(8), 1335–1349. https://doi.org/10.1080/02699931.2014.977850

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition*, *20*(3), 277–290. https://doi.org/10.3758/BF03199665

Christianson, S.-Å. (1984). The relationship between induced emotional arousal and amnesia. *Scandinavian Journal of Psychology*, *25*(2), 147–160. https://doi.org/10.1111/j.1467-9450.1984.tb01007.x

Christianson, S.-Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, *112*(2), 284–309. https://doi.org/10.1037/0033-2909.112.2.284

Christianson, S.-Å., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, *1*(4), 225–239. https://doi.org/10.1002/acp.2350010402

Christianson, S.-Å., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition and Emotion*, *5*(2), 81–108. https://doi.org/10.1080/02699939108411027

Cicerone, K. D., Goldin, Y., Ganci, K., Rosenbaum, A., Wethe, J. V., Langenbahn, D. M., ...Harley, J. P.

(2019). Evidence-based cognitive rehabilitation: systematic review of the literature from 2009 through 2014. *Archives of Physical Medicine and Rehabilitation*, *100*(8), 1515–1533. https://doi.org/10.1016/j.apmr.2019.02.011

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*(3), 211–226. https://doi.org/10.1016/j.bandl.2011.12.007

Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(14), 6008–6010. https://doi.org/10.1073/pnas.0811884106

Cook, G. I., Hicks, J. L., & Marsh, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition*, *35*(2), 222–230. https://doi.org/10.3758/BF03193443

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, USA: Routledge.

D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4*(2), 173–188. https://doi.org/10.1037/1093-4510.

*D'Argembeau, A., & Van der Linden, M. (2005). Influence of emotion on memory for temporal information. *Emotion*, *5*(4), 503–507. https://doi.org/10.1037/1528-3542.5.4.503

*Davidson, P. S. R., Mcfarland, C. P., & Glisky, E. L. (2006). Effects of emotion on item and source memory in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 306–322. https://doi.org/10.3758/CABN.6.4.306

Degl'Innocenti, A., & Bäckman, L. (1999). Source memory in major depression. *Journal of Affective Disorders*, *54*(1-2), 205–209. https://doi.org/10.1016/S0165-0327(98)00167-0

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology*, *12*(4), 541–551. https://doi.org/10.1080/095414400750050222

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1*(1), 5–11. https://doi.org/10.1037/1528-3542.1.1.5

Dolcos, F., Katsumi, Y., Weymar, M., Moore, M., Tsukiura, T., & Dolcos, S. (2017). Emerging directions in emotional episodic memory. *Frontiers in Psychology*, *8*, 1867. https://doi.org/10.3389/fpsyg.2017.01867

*Dougal, S., Phelps, E. A., & Davachi, L. (2007). The role of medial temporal lobe in item recognition and source recollection of emotional stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(3), 233–242. https://doi.org/10.3758/CABN.7.3.233

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177. https://doi.org/10.1037/1082-989X.1.2.170

*Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory*, *25*(9), 1191–1200. https://doi.org/10.1080/09658211.2017.1282517

Duval, S., & Tweedie, R. (2000). Trim and Fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183–201. https://doi.org/10.1037/h0047707

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7129), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Fairfield, B., Altamura, M., Padalino, F. A., Balzotti, A., Di Domenico, A., & Mammarella, N. (2016). False memories for affective information in schizophrenia. *Frontiers in Psychiatry*, *7*, 191. https://doi.org/10.3389/fpsyt.2016.00191

Ferguson, S. A., Hastroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging*, *7*(3), 443–452. https://doi.org/10.1037/0882-7974.7.3.443

*Ferré, P., Comesaña, M., & Guasch, M. (2019). Emotional content and source memory for language: Impairment in an incidental encoding task. *Frontiers in Psychology*, *10*, 65. https://doi.org/10.3389/fpsyg.2019.00065

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, *63*(3), 665–694. https://doi.org/10.1348/000711010X502733

Ford, J. H., & Kensinger, E. A. (2019). Age-by-emotion interactions in memory retrieval processes: An event-related potential study. *The Journals of Gerontology: Series B*, *74*(7), 1101–1110. https://doi.org/10.1093/geronb/gbx098

Ford, J. H., Morris, J. A., & Kensinger, E. A. (2014). Effects of emotion and emotional valence on the neural

correlates of episodic memory search and elaboration. *Journal of Cognitive Neuroscience*, *26*(4), 825–839. https://doi.org/10.1162/jocn_a_00529

*Gallant, S. N., Spaniol, J., & Yang, L. (2019). Age differences in cue utilization during prospective and retrospective memory monitoring. *Psychology and Aging*, *34*(4), 545–557. https://doi.org/10.1037/pag0000352

*Gallant, S. N., & Yang, L. (2014). Positivity effect in source attributions of arousal-matched emotional and non-emotional words during item-based directed forgetting. *Frontiers in Psychology*, *5*, 1334. https://doi.org/10.3389/fpsyg.2014.01334

Galli, G., Vadillo, M. A., Sirota, M., Feurra, M., & Medvedeva, A. (2019). A systematic review and meta-analysis of the effects of transcranial direct current stimulation (tDCS) on episodic memory. *Brain Stimulation*, *12*(2), 231–241. https://doi.org/10.1016/j.brs.2018.11.008

*Gallo, D. A., Korthauer, L. E., McDonough, I. M., Teshale, S., & Johnson, E. L. (2011). Age-related positivity effects and autobiographical memory detail: Evidence from a past/future source memory task. *Memory*, *19*(6), 641–652. https://doi.org/10.1080/09658211.2011.595723

Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*(4), 381–395. https://doi.org/10.1016/j.dcn.2012.05.002

Gingerich, A. C., & Dodson, C. S. (2013). Sad mood reduces inadvertent plagiarism: Effects of affective state on source monitoring in cryptomnesia. *Motivation and Emotion*, *37*(2), 355–371. https://doi.org/10.1007/s11031-012-9309-2

Hadley, C. B., & MacKay, D. G. (2006). Does emotion help or hinder immediate memory? Arousal versus priority-binding mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 79–88. https://doi.org/10.1037/0278-7393.32.1.79

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394–400. https://doi.org/10.1016/S1364-6613(00)01707-1

Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Henmi, M., & Copas, J. B. (2010). Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine*, *29*(29), 2969–2983. https://doi.org/10.1002/sim.4029

Herbener, E. S. (2008). Emotional memory in schizophrenia. *Schizophrenia Bulletin*, *34*(5), 875–887.

https://doi.org/10.1093/schbul/sbn081

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., & Rockstroh, B. (2006). Processing of emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology*, *43*(2), 197–206. https://doi.org/10.1111/j.1469-8986.2006.00385.x

Higgins, J., & Altman, D. G. (2008). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions: Cochrane book series* (pp. 187-241). Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470712184.ch8

Hinojosa, J. A., Carretié, L., Valcárcel, M. A., Méndez-Bértolo, C., & Pozo, M. A. (2009). Electrophysiological differences in the processing of affective information in words and pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(2), 173–189. https://doi.org/10.3758/CABN.9.2.173

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: A special relationship? *Emotion*, *5*(4), 489–497. https://doi.org/10.1037/1528-3542.5.4.489

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*(3), 349–362. https://doi.org/10.1016/j.cpr.2010.01.001

Holmes, E. A., Mathews, A., Mackintosh, B., & Dalgleish, T. (2008). The causal effect of mental imagery on emotion assessed using picture-word cues. *Emotion*, *8*(3), 395–409. https://doi.org/10.1037/1528-3542.8.3.395

Hostler, T. J., Wood, C., & Armitage, C. J. (2018). The influence of emotional cues on prospective memory: A systematic review with meta-analyses. *Cognition and Emotion*, *32*(8), 1578–1596. https://doi.org/10.1080/02699931.2017.1423280

Hou, M., Grilli, M. D., & Glisky, E. L. (2019). Self-reference enhances relational memory in young and older adults. *Aging, Neuropsychology, and Cognition*, *26*(1), 105–120. https://doi.org/10.1080/13825585.2017.1409333

Ioannidis, J. P. A., & Trikalinos, T. A. (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *Canadian Medical Association Journal*, *176*(8), 1091–1096. https://doi.org/10.1503/cmaj.060410

*Jia, X., Gao, C., Cui, L., & Guo, C. (2018). Does emotion arousal influence the benefit received from testing: Insights from neural correlates of retrieval mode effect. *NeuroReport*, *29*(17), 1449–1455. https://doi.org/10.1097/WNR.0000000000001130

Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of*

Experimental Psychology: General, 117(4), 390-394. https://doi.org/10.1037/0096-3445.117.4.390

Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues*, *40*(2), 33–50. https://doi.org/10.1111/j.1540-4560.1984.tb01092.x

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85. https://doi.org/10.1037/0033-295X.88.1.67

Jurica, P. J., & Shimamura, A. P. (1999). Monitoring item and source information: Evidence for a negative generation effect in source memory. *Memory & Cognition*, *27*(4), 648–656. https://doi.org/10.3758/BF03211558

*Kaida, K., Niki, K., & Born, J. (2015). Role of sleep for encoding of emotional memory. *Neurobiology of Learning and Memory*, *121*, 72–79. https://doi.org/10.1016/j.nlm.2015.04.002

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*(4), 241–253. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science*, *16*(4), 213–218. https://doi.org/10.1111/j.1467-8721.2007.00506.x

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, *1*(2), 99–113. https://doi.org/10.1177/1754073908100432

*Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, *101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*(4), 575–591. https://doi.org/10.1016/j.jml.2006.05.004

Kensinger, E. A., & Kark, S. M. (2018). Emotion and memory. In J. T. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed., pp. 1–26). New York, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119170174.epcn101

*Kensinger, E. A., O'Brien, J. L., Swanberg, K., Garoff-Eaton, R. J., & Schacter, D. L. (2007). The effects of emotional content on reality-monitoring performance in young and older adults. *Psychology and Aging*, *22*(4), 752–764. https://doi.org/10.1037/0882-7974.22.4.752

*Kensinger, E. A., & Schacter, D. L. (2005). Emotional content and reality-monitoring ability: fMRI evidence for the influences of encoding processes. *Neuropsychologia*, *43*(10), 1429–1443. https://doi.org/10.1016/j.neuropsychologia.2005.01.004

*Kensinger, E. A., & Schacter, D. L. (2006a). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *The Journal of Neuroscience*, *26*(9), 2564–2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

*Kensinger, E. A., & Schacter, D. L. (2006b). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition*, *34*(2), 251–260. https://doi.org/10.3758/BF03193403

Kensinger, E. A., & Schacter, D. L. (2016). Memory and emotion. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 564–578). New York, USA: Guilford Publications.

*Koenig, S., & Mecklinger, A. (2008). Electrophysiological correlates of encoding and retrieving emotional events. *Emotion*, *8*(2), 162–173. https://doi.org/10.1037/1528-3542.8.2.162

Ladouceur, C. D. (2012). Neural systems supporting cognitive-affective interactions in adolescence: The role of puberty and implications for affective disorders. *Frontiers in Integrative Neuroscience*, *6*, 65. https://doi.org/10.3389/fnint.2012.00065

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry, 44*(12), 1248-1263. http://doi.org/10.1016/S0006-3223(98)00275-3

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

*Le Bigot, L., Knutsen, D., & Gil, S. (2018). I remember emotional content better, but I'm struggling to remember who said it! *Cognition*, *180*, 52–58. https://doi.org/10.1016/j.cognition.2018.07.001

*Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and

older adults. *Aging Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Levine, L. J., & Edelstein, R. S. (2009). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition and Emotion*, *23*(5), 833–875. https://doi.org/10.1080/02699930902738863

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition*, *22*(5), 530–554. https://doi.org/10.1521/soco.22.5.530.50767

Lipinska, G., Stuart, B., Thomas, K. G. F., Baldwin, D. S., & Bolinger, E. (2019). Preferential consolidation of emotional memory during sleep: A meta-analysis. *Frontiers in Psychology*, *10*, 1014. https://doi.org/10.3389/fpsyg.2019.01014

MacKay, D. G., Johnson, L. W., Graham, E. R., & Burke, D. M. (2015). Aging, emotion, attention, and binding in the taboo stroop task: Data and theories. *International Journal of Environmental Research and Public Health, 12*(10), 12803-12833. https://doi.org/10.3390/ijerph121012803

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

*MacKenzie, G., Powell, T. F., & Donaldson, D. I. (2015). Positive emotion can protect against source memory impairment. *Cognition and Emotion*, *29*(2), 236–250. https://doi.org/10.1080/02699931.2014.911145

*Maddock, R. J., & Frein, S. T. (2009). Reduced memory for the spatial and temporal context of unpleasant words. *Cognition and Emotion*, *23*(1), 96–117. https://doi.org/10.1080/02699930801948977

Mancuso, L. E., Ilieva, I. P., Hamilton, R. H., & Farah, M. J. (2016). Does transcranial direct current stimulation improve healthy working memory?: A meta-analytic review. *Journal of Cognitive Neuroscience*, *28*(8), 1063–1089. https://doi.org/10.1162/jocn_a_00956

*Mao, W. B., An, S., & Yang, X. F. (2017). The effects of goal relevance and perceptual features on emotional items and associative memory. *Frontiers in Psychology*, *8*, 1223. https://doi.org/10.3389/fpsyg.2017.01223

*Mao, X., You, Y., Li, W., & Guo, C. (2015). Emotion impairs extrinsic source memory-An ERP study. *Biological Psychology*, *110*, 182–189. https://doi.org/10.1016/j.biopsycho.2015.07.005

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on*

Psychological Science, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, *16*(3-5), 437–458. https://doi.org/10.1080/026432999380870

*Mather, M., & Knight, M. (2008). The emotional harbinger effect: Poor context memory for cues that previously predicted something arousing. *Emotion*, *8*(6), 850–860. https://doi.org/10.1037/a0014087

*Mather, M., & Nesmith, K. (2008). Arousal-enhanced location memory for pictures. *Journal of Memory and Language*, *58*(2), 449–464. https://doi.org/10.1016/j.jml.2007.01.004

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competitition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

*Mathews, A., Ridgeway, V., & Holmes, E. A. (2013). Feels like the real thing: Imagery is both more realistic and emotional than verbal thought. *Cognition and Emotion*, *27*(2), 217–229. https://doi.org/10.1080/02699931.2012.698252

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276-282. https://doi.org/10.11613/BM.2012.031

*McKague, M., McAnally, K. I., Skovron, M., Bendall, S., & Jackson, H. J. (2012). Source monitoring and proneness to auditory-verbal hallucinations: A signal detection analysis. *Cognitive Neuropsychiatry*, *17*(6), 544–562. https://doi.org/10.1080/13546805.2012.676311

*Minor, G., & Herzmann, G. (2019). Effects of negative emotion on neural correlates of item and source memory during encoding and retrieval. *Brain Research*, *1718*, 32–45. https://doi.org/10.1016/j.brainres.2019.05.001

Mitchell, K. J., Mather, M., Johnson, M. K., Raye, C. L., & Greene, E. J. (2006). A functional magnetic resonance imaging investigation of short-term source and item memory for negative pictures. *NeuroReport*, *17*(14), 1543–1547. https://doi.org/10.1097/01.wnr.0000234743.50442.e5

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLoS Medicine, 6*(6), e1000097. https://doi.org/10.1371/journal.pmed1000097

*Morrison, A. P., & Haddock, G. (1997). Cognitive factors in source monitoring and auditory hallucinations.

Psychological Medicine, *27*(3), 669–679. https://doi.org/10.1017/S003329179700487X

Murphy, N. A., & Isaacowitz, D. M. (2008). Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging*, *23*(2), 263–286. https://doi.org/10.1037/0882-7974.23.2.263

Murray, B. D., & Kensinger, E. A. (2012). The effects of emotion and encoding strategy on associative memory. *Memory & Cognition*, *40*(7), 1056–1069. https://doi.org/10.3758/s13421-012-0215-3

*Nashiro, K., & Mather, M. (2011a). How arousal affects younger and older adults' memory binding. *Experimental Aging Research*, *37*(1), 108–128. https://doi.org/10.1080/0361073X.2011.536746

*Nashiro, K., & Mather, M. (2011b). The effect of emotional arousal on memory binding in normal aging and Alzheimer's disease. *American Journal of Psychology*, *124*(3), 301–312.

*Newsome, R. N., Dulas, M. R., & Duarte, A. (2012). The effects of aging on emotion-induced modulations of source retrieval ERPs: Evidence for valence biases. *Neuropsychologia*, *50*(14), 3370–3384. https://doi.org/10.1016/j.neuropsychologia.2012.09.024

*Novak, D. L., & Mather, M. (2009). The tenacious nature of memory binding for arousing negative items. *Memory & Cognition*, *37*(7), 945–952. https://doi.org/10.3758/MC.37.7.945

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General*, *129*(2), 242–261. https://doi.org/10.1037/0096-3445.129.2.242

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104

Oliver, M. C., Bays, R. B., & Zabrucky, K. M. (2016). False memories and the DRM paradigm: Effects of imagery, list, and test type. *Journal of General Psychology*, *143*(1), 33–48. https://doi.org/10.1080/00221309.2015.1110558

*Otani, H., Jaffa, M., Libkuman, T. M., Goernert, P. N., & Kato, K. (2012). Does source memory impairment associated with emotionally arousing stimuli occur at encoding or retrieval? *Journal of Cognitive Psychology*, *24*(7), 796–801. https://doi.org/10.1080/20445911.2012.693468

*Otani, H., Libkuman, T. M., Goernert, P. N., Kato, K., Migita, M., Freehafer, S. E., & Landow, M. P. (2012). Emotion, directed forgetting, and source memory. *British Journal of Psychology*, *103*(3), 343–358.

https://doi.org/10.1111/j.2044-8295.2011.02078.x

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, *11*(4), 137–138. https://doi.org/10.3758/BF03331011

Palumbo, R., Mammarella, N., Di Domenico, A., & Fairfield, B. (2018). When and where in aging: The role of music on source monitoring. *Aging Clinical and Experimental Research*, *30*(6), 669–676. https://doi.org/10.1007/s40520-018-0955-4

*Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy on internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology*, *10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2020). Is internal source memory recognition modulated by emotional encoding contexts? *Psychological Research*. Advance online publication. https://doi.org/10.1007/s00426-020-01294-4

Quintana, D. S. (2015). From pre-registration to publication: A non-technical primer for conducting a meta-analysis to synthesize correlational data. *Frontiers in Psychology*, *6*, 1549. https://doi.org/10.3389/fpsyg.2015.01549

Raye, C. L., & Johnson, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society*, *15*(6), 405–408. https://doi.org/10.3758/BF03334572

Rohatgi, A. (2019). WebPlotDigitizer (Version 4.2) [Computer software]. Retrieved from https://automeris.io/WebPlotDigitizer/index.html

Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: A meta-analysis. *Psychonomic Bulletin and Review*, *26*(5), 1529–1547. https://doi.org/10.3758/s13423-019-01649-y

*Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion*, *11*(3). https://doi.org/10.1037/a0024246

*Rimmele, U., Davachi, L., & Phelps, E. A. (2012). Memory for time and place contributes to enhanced confidence in memories for emotinal events. *Emotion*, *12*(4), 834–846. https://doi.org/10.1037/a0028003

*Ritchey, M., Wang, S. F., Yonelinas, A. P., & Ranganath, C. (2019). Dissociable medial temporal pathways

for encoding emotional item and context information. *Neuropsychologia*, *124*, 66–78. https://doi.org/10.1016/j.neuropsychologia.2018.12.015

*Robin, F., Cébron, T., Letellier, M., & Nizard, J. (2018). The effects of emotional content on source monitoring in fibromyalgia patients. *Journal of General Psychology*, *145*(4), 392–414. https://doi.org/10.1080/00221309.2018.1494536

RStudio Team (2018). RStudio: Integrated development for R. RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/

Russel, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sala, G., Aksayli, N. D., Tatlidil, K. S., Gondo, Y., & Gobet, F. (2019). Working memory training does not enhance older adults' cognitive skills: A comprehensive meta-analysis. *Intelligence*, *77*, 101386. https://doi.org/10.1016/j.intell.2019.101386

Schaefer, A., Pottage, C. L., & Rickart, A. J. (2011). Electrophysiological correlates of remembering emotional pictures. *NeuroImage*, *54*(1), 714–724. https://doi.org/10.1016/j.neuroimage.2010.07.030

*Schmidt, K., Patnaik, P., & Kensinger, E. A. (2011). Emotion's influence on memory for spatial and temporal context. *Cognition and Emotion*, *25*(2), 229–243. https://doi.org/10.1080/02699931.2010.483123

Schweizer, S., Satpute, A. B., Atzil, S., Field, A. P., Hitchcock, C., Black, M., ...Dalgleish, T. (2019). The impact of affective information on working memory: A pair of meta-analytic reviews of behavioral and neuroimaging evidence. *Psychological Bulletin*, *145*(6), 566–609. https://doi.org/10.1037/bul0000193

Sharot, T., & Phelps, E. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(3), 294–306. https://doi.org/10.3758/CABN.4.3.294

*Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Shields, G. S., Sazma, M. A., McCullough, A. M., & Yonelinas, A. P. (2017). The effects of acute stress on episodic memory: A meta-analysis and integrative review. *Psychological Bulletin*, *143*(6), 636–675. https://doi.org/10.1037/bul0000100

Song, F., Hooper, L., & Loke, Y. K. (2013). Publication bias: What is it? How do we measure it? How do we avoid it? *Open Access Journal of Clinical Trials*, *5*(1), 71–81. https://doi.org/10.2147/OAJCT.S34419

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging*, *10*(4), 527–539. https://doi.org/10.1037/0882-7974.10.4.527

Steinmetz, K. R. M., & Kensinger, E. A. (2009). The effects of valence and arousal on the neural activity leading to subsequent memory. *Psychophysiology*, *46*(6), 1190–1199. https://doi.org/10.1111/j.1469-8986.2009.00868.x.The

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., ...Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *342*, d4002. https://doi.org/10.1136/bmj.d4002

Stevens, J. S., & Hamann, S. (2012). Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. *Neuropsychologia*, *50*(7), 1578–1593. https://doi.org/10.1016/j.neuropsychologia.2012.03.011

Talmi, D. (2013). Enhanced emotional memory. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Talmi, D., Luk, B. T. C., McGarry, L. M., & Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *Journal of Memory and Language*, *56*(4), 555–574. https://doi.org/10.1016/j.jml.2007.01.002

Talmi, D., & McGarry, L. M. (2012). Accounting for immediate emotional memory enhancement. *Journal of Memory and Language*, *66*(1), 93–108. https://doi.org/10.1016/j.jml.2011.07.009

Taneja, P., Olausson, H., Trulsson, M., Svensson, P., & Baad-Hansen, L. (2019). Defining pleasant touch stimuli: A systematic review and meta-analysis. *Psychological Research*, Advance online publication. https://doi.org/10.1007/s00426-019-01253-8

Teixeira-Santos, A. C., Moreira, C. S., Magalhães, R., Magalhães, C., Pereira, D. R., Leite, J., ...Sampaio, A. (2019). Reviewing working memory training gains in healthy older adults: A meta-analytic review of transfer for cognitive outcomes. *Neuroscience & Biobehavioral Reviews*, *103*, 163–177. https://doi.org/10.1016/j.neubiorev.2019.05.009

Thompson, T., Oram, C., Correll, C. U., Tsermentseli, S., & Stubbs, B. (2017). Analgesic effects of alcohol: A systematic review and meta-analysis of controlled experimental studies in healthy participants.

Journal of Pain, *18*(5), 499–510. https://doi.org/10.1016/j.jpain.2016.11.009

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, *8*, 1454. https://doi.org/10.3389/fpsyg.2017.01454

Urban, E. J., Charles, S. T., Levine, L. J., & Almeida, D. M. (2018). Depression history and memory bias for specific daily emotions. *PLoS ONE*, *13*(9), e0203574. https://doi.org/10.1371/journal.pone.0203574

van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2019). Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. *PLoS ONE*, *14*(4), e0215052. https://doi.org/10.1371/journal.pone.0215052

Ventura-Bort, C., Dolcos, F., Wendt, J., Wirkner, J., Hamm, A. O., & Weymar, M. (2017). Item and source memory for emotional associates is mediated by different retrieval processes. *Neuropsychologia*. Advance online publication. https://doi.org/10.1016/j.neuropsychologia.2017.12.015

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vilberg, K. L., & Rugg, M. D. (2009). An investigation of the effects of relative probability of old and new test items on the neural correlates of successful and unsuccessful source memory. *NeuroImage*, *45*(2), 562–571. https://doi.org/10.1016/j.neuroimage.2008.12.020

*Wang, B. (2012). Females' superiority in item memory, but not source memory for neutral and emotional Chinese words. *Personality and Individual Differences*, *52*(8), 925–929. https://doi.org/10.1016/j.paid.2012.02.005

*Wang, B., & Fu, X. (2011). Time course of effects of emotion on item memory and source memory for Chinese words. *Neurobiology of Learning and Memory*, *95*(4), 415–424. https://doi.org/10.1016/j.nlm.2011.02.001

Waring, J. D., & Kensinger, E. A. (2009). Effects of emotional valence and arousal upon memory trade-offs with aging. *Psychology and Aging*, *24*(2), 412–422. https://doi.org/10.1037/a0015526

Weymar, M., Löw, A., & Hamm, A. O. (2011). Emotional memories are resilient to time: Evidence from the parietal ERP old/new effect. *Human Brain Mapping*, *32*(4), 632–640. https://doi.org/10.1002/hbm.21051

Weymar, M., Löw, A., Schwabe, L., & Hamm, A. O. (2010). Brain dynamics associated with recollective experiences of emotional events. *NeuroReport*, *21*(12), 827–831.

https://doi.org/10.1097/WNR.0b013e32833d180a

Wirkner, J., Ventura-Bort, C., Schulz, P., Hamm, A. O., & Weymar, M. (2018). Event-related potentials of emotional and neutral memories: The role of encoding position and delayed testing. *Psychophysiology*, *55*(7), 1–12. https://doi.org/10.1111/psyp.13069

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

*Yick, Y. Y., Buratto, L. G., & Schaefer, A. (2015). The effects of negative emotion on encoding-related neural activity predicting item and source recognition. *Neuropsychologia*, *73*, 48–59. https://doi.org/10.1016/j.neuropsychologia.2015.04.030

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: An emotional binding account. *Trends in Cognitive Sciences*, *19*(5), 259–267. https://doi.org/10.1016/j.tics.2015.02.009

*Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-Serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology*, *65*(4), 236–244. https://doi.org/10.1027/a000001

CHAPTER III

DIFFERENTIAL EFFECTS OF VALENCE AND ENCODING STRATEGY ON INTERNAL SOURCE MEMORY AND JUDGMENTS OF SOURCE: EXPLORING THE PRODUCTION AND THE SELF-REFERENCE EFFECT

Differential effects of valence and encoding strategy on internal source memory and judgments of source: Exploring the production and the self-reference effect[1,2]

Abstract

Item memory studies show that emotional stimuli are associated with improved memory performance compared to neutral ones. However, emotion-related effects on source memory are less consistent. The current study probed how emotional valence and specific encoding conditions influence internal source memory performance and judgments of source (JOSs). In two independent experiments, participants were required to read silently/aloud (Experiment 1) or to perform self-reference/common judgments (Experiment 2) on a list of negative/neutral/positive words. They also performed immediate JOSs ratings for each word. The study phase was followed by a test phase in which participants performed old-new judgments. In Experiment 1, the production effect was replicated for item memory, but the effects of valence on item and source memory were not significant. In Experiment 2, self-referential processing effects on item and source memory differed as a function of valence. In both experiments, JOSs ratings were sensitive to valence and encoding conditions, although they were not predictive of objective memory performance. These findings demonstrate that the effects of valence on internal source memory and JOSs are modulated by encoding strategy. Thus, the way information is encoded can shed some light on how emotion might enhance, impair or exert no influence on source memory.

Keywords: internal source memory; valence; emotion; production effect; self-reference effect; judgments of source; metamemory.

Evidence for the intricate interactions between emotion and cognition has led to a remarkable shift in the cognitive sciences (Ochsner & Phelps, 2007; Okon-Singer, Hendler, Pessoa, & Shackman, 2015). One example is the relationship between emotion and episodic memory. Specifically, the emotion-enhanced memory (EEM) effect refers to improved memory for emotionally charged (high arousal and/or negative/positive) compared to non-emotional (low arousal and/or neutral) information. This effect has been well documented for different stimulus types, such as single words (e.g., D'Argembeau & Van der Linden, 2004; Davidson, McFarland, & Glisky, 2006; Kensinger & Corkin, 2003; Maddock & Frein, 2009),

[1] Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy in internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology, 10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

[2] The Supplementary Materials associated with this chapter are presented in Appendix B.

word pairs (e.g., Maddox, Naveh-Benjamin, Old, & Kilb, 2012), or pictures (e.g., Nashiro & Mather, 2011; Schmidt, Patnaik, & Kensinger, 2011; Schümann, Bayer, Talmi, & Sommer, 2018; Yick, Buratto, & Schaefer, 2015). The EEM has also been extensively reported in the case of item memory (i.e., memory for the central features and relevant content of an event such as the words in a word pair; e.g., Kensinger & Kark, 2018; Kensinger & Schacter, 2008; Levine & Edelstein, 2009; Murphy & Isaacowitz, 2008). Nonetheless, a similar enhancement effect has been less consistent on other aspects of episodic memory such as source memory, i.e., the memory for the origins and conditions in which a certain event occurred (Johnson, Hastroudi, & Lindsay, 1993; Johnson & Raye, 1981; see Appendix for a selective review of these studies).

According to the source-monitoring framework (Johnson et al., 1993), when we try to specify the origins of an event, our decision may rely on qualitative characteristics that are bound to an item during encoding, including semantic detail, cognitive operations (e.g., elaboration; organization; imagery), perceptual (e.g., color; sound), contextual (e.g., time; space) and/or affective (e.g., emotional reactions) features, as well as on previous general knowledge or schemas (e.g., stereotypes; beliefs). The enhanced memory effect for items with emotional features during encoding has been discussed in light of different theoretical approaches, including the Easterbrook's (1959) cue-utilization hypothesis, the priority-binding theory (Mackay et al., 2004), the object-based binding theory (Mather, 2007), and the arousal-biased competition theory (Mather & Sutherland, 2011; see Bowen, Kark, & Kensinger, 2018 for a review of other approaches). For instance, whereas impaired source memory for emotional stimuli might be explained by narrowed attention to central details and reduced capacity to bind central information with peripheral details (Easterbrook, 1959), a source memory enhancement might stem from a stronger activation of item-context binding mechanisms elicited by emotional events (Mackay et al., 2004). In turn, the object-based binding theory (Mather, 2007) conciliates contradictory reports of impairment *vs.* enhancement *vs.* null effects, by postulating that arousal facilitates the binding of elements that are an integral part of the emotional object (i.e., intrinsic features such as the color of the object), whereas it impairs or exerts no influence on the binding of an emotional object with other contextual information (i.e., extrinsic features such as the color of the border surrounding the object). Similarly, the arousal-biased competition theory (Mather & Sutherland, 2011) also predicts different outcomes for contextual details depending on the bottom-up (stimulus-driven information, such as its perceptual saliency) and/or top-down (goal-driven information, such as relevance for the current goals, expectations, and prior knowledge) priority of the information. In the latter case, arousal enhances memory for high priority

information (i.e., information characterized by salient perceptual features and/or with relevance) while impairing memory for low priority information.

Even considering the former theoretical frameworks, several other factors may impact upon the relationship between emotion and source memory, including stimulus type (e.g., words or pictures; Phelps, LaBar, & Spencer, 1997), the emotional properties of the stimuli (Boywitt, 2015), and the characteristics of the experimental task (e.g., intentional/incidental encoding, retention interval, encoding duration, type of memory test, sensory modality; D'Argembeau & Van der Linden, 2004; Ferré, Comesaña, & Guasch, 2019). Additionally, other factors have been relatively ignored in the study of source memory and its interactions with emotion, such as the source monitoring task and the encoding strategy. Regarding the first, Johnson and colleagues (1993) proposed three distinct source monitoring processes: external source monitoring, which refers to the discrimination between different external sources (e.g., something said speaker X or speaker Y?); internal source monitoring, which relies on the distinction between internally derived information (e.g., something I said silently or aloud?); reality monitoring, which comprises the discrimination between internal and external sources (e.g., something I heard/saw or imagined?). Externally derived memories usually contain more perceptual, temporal, spatial, and affective information, whereas internally derived memories comprise more information regarding cognitive processes that occur during memory acquisition (Johnson & Raye, 1981; Raye & Johnson, 1980). Moreover, due to the overlap of informative cues in both quantity and quality, the discrimination between two or more external/internal sources is more challenging than the discrimination between an external and an internal source (Raye & Johnson, 1980). Accordingly, the way emotion affects source memory may depend on the contextual features that are manipulated, as already pointed out by previous studies showing that different external source memory tasks can lead to differences in source memory performance (e.g., D'Argembeau & Van der Linden, 2004; Koenig & Mecklinger, 2008; Kuhlmann & Touron, 2017; MacKenzie, Powell, & Donaldson, 2015). For example, Boywitt (2015) demonstrated that the relationship between arousal and source memory performance is linear in the case of frame color, but it follows an inverted U-shape function for spatial location. Nonetheless, in comparison with external source memory, the study of other contextual features such as the discrimination between cognitive operations (internal source memory) when using emotional stimuli has received less attention (see Appendix).

In the case of encoding strategies, different approaches have been shown to benefit source memory performance such as semantic clustering (Wegesin, Jacobs, Zubin, Ventura, & Stern, 2000) and unitization (i.e., the ability to bind items or item-context in a single and meaningful combination**;** Tu, Alty,

& Diana, 2017; Tu & Diana, 2016; see also El Haj, Kessels, & Allain, 2016 for a review). Additionally, Kuhlmann and Touron (2012, 2017) observed that when participants spontaneously used interactive imagery and sentence generation during encoding, source memory performance was improved. Notwithstanding, other encoding factors already known to boost item memory, such as production mode and self-referential processing, may also influence source memory (Hamami, Serbun, & Gutchess, 2011; Leshikar & Duarte, 2012, 2014; Ozubko, Major, & MacLeod, 2014; Serbun, Shih, & Gutchess, 2011; Yin, Ma, Xu, & Yang, 2018). Specifically, the production effect is a simple mnemonic strategy that shows a memory benefit for vocal production conditions, such as mouthing, reading aloud, reading aloud loudly, and singing, when compared to silent reading conditions (Dodson & Schacter, 2001; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Ozubko & MacLeod, 2010; Quinlan & Taylor, 2013). The self-reference effect, in turn, represents a memory benefit for information that is encoded in relation to the self in comparison with information processed in other deep (e.g., semantic; other self-referent) or shallow (e.g., phonemic; perceptual) conditions (Kuiper & Rogers, 1979; Leshikar & Duarte, 2012; Leshikar, Dulas, & Duarte, 2015; Rogers, Kuiper, & Kirker, 1977; Symons & Johnson, 1997; Yang, Truong, Fuss, & Bislimovic, 2012).

Acknowledging the modulatory role of the encoding strategy and type of source memory task in the interaction between emotion and source memory, the first aim of the current study was to explore how stimulus valence influences internal source monitoring, when the encoding strategies (production mode and the self-referential processing) are also manipulated. Regarding the production mode, previous studies revealed that a similar beneficial effect is observed on item and source memories when stimuli are produced aloud (Ozubko et al., 2014). Nonetheless, to the best of our knowledge, no previous study has concomitantly explored the role of emotion and production mode in source memory. In the case of the self-reference effect, previous research demonstrated that both item and source memory are enhanced when a self-referential approach is implemented during encoding (Hamami et al., 2011; Kalenzaga et al., 2015; Leshikar & Duarte, 2012, 2014; Serbun et al., 2011; Yin et al., 2018; Zhang, Pan, Li, & Guo, 2018). Also, when stimulus valence is manipulated, both neutral and positive stimuli that are self-referentially encoded are better remembered compared to negative stimuli (e.g., D'Argembeau, Comblain, & Van der Linden, 2005; Leshikar et al., 2015; Yang et al., 2012; Zhang et al., 2018). This pattern appears to hold partially for positive stimuli in the case of internal source monitoring, but it has been less consistently explored in the case of neutral stimuli (e.g., Durbin, Mitchell, & Johnson, 2017). Thus, in the current study, we expected to replicate the self-reference benefit for both item and source memory, and to better qualify this effect when valence is manipulated during encoding.

Together with objective measures of memory performance, a second aim of this study was to examine how stimulus valence and encoding strategy might modulate subjective metamemory processes, specifically judgements of source (JOSs). Metamemory encompasses specific knowledge and beliefs about memory functioning (Flavell, 1979), such as what we will remember or forget in the future. The study of metamemory often relies on prospective tasks in which participants are required to judge if they will remember/forget previously learned information, or on retrospective tasks in which participants are asked to evaluate their memory performance after remembering some information (Nelson & Narens, 1990). One of the most common measures to assess prospective metamemory relies on judgments of learning (JOLs), in which participants are required to make a prediction about the memorability of specific stimuli, during or after their acquisition, through a rating or numeric scale. So far, few studies have explored the role of emotion on metamemory judgments. The existing evidence suggests that participants deem emotional stimuli to be more memorable than neutral stimuli during the encoding stage. This effect has been observed with odors (Jönsson, Olsson, & Olsson, 2005), faces (Nomi, Rhodes, & Cleary, 2013), pictures (Hourihan & Bursey, 2017), and words (Hourihan, Fraundorf, & Benjamin, 2017; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). When considering the influence of emotion on judgments of source (JOSs), the existing data is even more scarce. One of the first studies to introduce the concept of JOSs was led by Dutton and Carroll (2001), in which they explored how three different emotionally arousing conditions modulated both JOLs and JOSs predictions. These authors found that JOLs and JOSs predictions somehow matched the recall performance, particularly in emotional conditions characterized by high arousal. In this context, the results of the current study can represent a novel contribution to the study of emotion and prospective metamemory, especially in the case of internal source memory monitoring.

Based on the evidence reviewed above, our first experiment investigated possible interactions between the production effect and stimulus valence in the case of internal source monitoring and JOSs. First, we predicted a replication of the EEM for item memory (Kensinger & Kark, 2018; Kensinger & Schacter, 2008; Levine & Edelstein, 2009; Murphy & Isaacowitz, 2008), i.e. better recognition performance for both negative and positive stimuli in comparison with neutral stimuli regardless of the production mode, based on the assumption that both production modes recruit similar attentional and elaborative processes. Second, in the case of internal source memory, we hypothesized an impairment or, more likely, a null effect of valence (see Appendix for a selective review of studies). This prediction is supported by the object-based binding hypothesis (Mather, 2007) as the production mode can be deemed as an extrinsic feature of the stimuli. It is also supported by the arousal-biased competition theory (Mather

& Sutherland, 2011): considering that participants were overtly instructed to memorize the stimuli in the current experiment, both emotional and non-emotional stimuli were goal-relevant and received similar top-down priority. Third, we expected JOSs to be sensitive to both mode of production and stimulus valence as previous studies demonstrated that stimuli read aloud receive higher JOLs ratings than stimuli read silently (Castel, Rhodes, & Friedman, 2013), and that emotional words are perceived as more memorable (Hourihan et al., 2017; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). In our second experiment, the effects of both self-referential encoding and valence on internal source memory and JOSs were examined. The replication of the self-reference mnemonic benefit for both item and source memory was expected, especially in the case of neutral and positive stimuli (D'Argembeau et al., 2005; Leshikar et al., 2015; Yang et al., 2012; Zhang et al., 2018). Although these predictions contrast with reports of null or impairment effects of emotion on internal source memory (see Appendix), they highlight how the encoding strategy and top-down processes, such as motivational relevance (Mather & Sutherland, 2011), might modulate the relationship between emotion and source memory. Regarding the metamemory judgments, we hypothesized JOSs to be sensitive to valence and encoding strategy as in Experiment 1. Specifically, we predicted that positive and neutral stimuli would be judged as more memorable than negative stimuli given that they might be perceived as more relevant self-descriptors (D'Argembeau et al., 2005; Kuiper & Rogers, 1979).

Experiment 1

In a previous study by Ozubko and colleagues (2014; Experiment 3), the production mode benefited internal source memory for items read aloud during encoding. Reading aloud seems to be a distinctive feature that can function as a diagnostic cue during the test phase: if we remember that a specific word was read aloud, we will probably classify the item as "old". This will be less likely to occur in the case of silent words as they have the same production status as new words during the test (MacLeod et al., 2010; Ozubko & MacLeod, 2010). In the current experiment, the source memory benefit for items read aloud was tested by manipulating both the production mode (silent *vs.* aloud) and stimulus valence (negative *vs.* neutral *vs.* positive). During the study phase, participants were required to read a list of words with different valence properties, half aloud and half silently. For each word, participants were also asked to make JOSs using a six-point rating scale. Immediately after learning, participants performed a test in which they were asked to discriminate between old and new items and to recall the production mode.

Material and methods

Participants.

Thirty-two college students participated in this experiment. However, one was excluded due to self-reported depression diagnosis and use of antidepressants; three other participants were excluded as they scored above 13 in the Portuguese version of the Beck Depression Inventory-II, which is indicative of depressive symptoms (Coelho, Martins, & Barros, 2002). The final sample was composed of 28 participants (23 females), aged between 18 and 32 years (M = 20.25, SD = 3.62), and with an average of 13.54 years of formal education (SD = 2.28). There were no self-reports of psychiatric/neurological disorders nor psychoactive drug usage. Participants had normal or corrected-to-normal vision, and they reported no auditory or other sensory and motor problems. They provided informed consent prior to their enrolment in the study and received course credit for their collaboration. The research protocol was approved by the local Ethics Committee (University of Minho, Braga, Portugal).

In Experiment 3 of Ozubko and colleagues (2014), where a superiority source memory effect was found for items studied aloud, a total of 24 participants were enrolled and the effect sizes (partial eta squared - η_p^2) ranged between .24 and .77. Based on this information, *a priori* sample size estimation was conducted using G*Power-3 statistical software (Faul, Erdfelder, Lang, & Buchner, 2007), considering the within-factors repeated-measures analysis of variance (RMANOVA; 3 valence: negative/neutral/positive x 2 source: aloud/silent design) as the main statistical test. For an alpha significance level of .05, a power of .80, and an effect size of .24 (as calculated in SPSS), a sample of at least 11 participants would be required to allow the detection of differences between modes of production in source memory. Thus, the number of recruited participants was adequate.

Materials.

A total of 180 words (60 negative/neutral/positive; see Supplementary Table S1) were initially selected from the Portuguese version of the Affective Norms for English Words (ANEW; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012). Words differed in valence, ranging from 1 to 9, with 9 corresponding to the most positive ratings and 1 to the least positive ratings (positive > neutral; positive > negative; neutral > negative; all $p < .001$). In the case of arousal, differences were observed between neutral and both negative and positive stimuli (neutral < negative; neutral < positive; negative = positive**;** all $p < .001$). No significant differences were found in frequency, number of letters and number of syllables as a function of word valence ($p > .05$; see Supplementary Table S1). The selected words were randomly distributed across six lists of 30 stimuli each (with 10 stimuli of each valence category) to be used as

learning and test items. Lists did not differ in words' valence, arousal, frequency, number of letters, and number of syllables ($p > .05$).

Procedure.

The experimental task was completed in a single session, and it consisted of three study-test cycles. During the study phase, each trial began with a fixation cross (500 ms) followed by a blank screen (250 ms), and then a target word (3000 ms), which was presented in the center of the screen in a light grey background (Arial, black color, font size 28). Together with each word, the instruction "read silently" or "read aloud" was presented on top of the screen to inform how the word should be read. Participants were then instructed to judge how likely they were to remember that the previous word was read silently or aloud by using a six-point rating scale following Hourihan and colleagues (2017), in which "1" represented "Sure I will not remember" and "6" represented "Sure I will remember". The rating scale remained on the screen until the participant made the JOSs. Lastly, a blank screen appeared during 500 ms (see Figure 1). Thirty stimuli were randomly presented during encoding to be later tested. Four additional words, two at the beginning and two at the end of the list, were used as fillers to mitigate possible primacy and recency effects, but they were not considered in the analysis. No specifications were provided about which fingers the participants should use to provide a response.

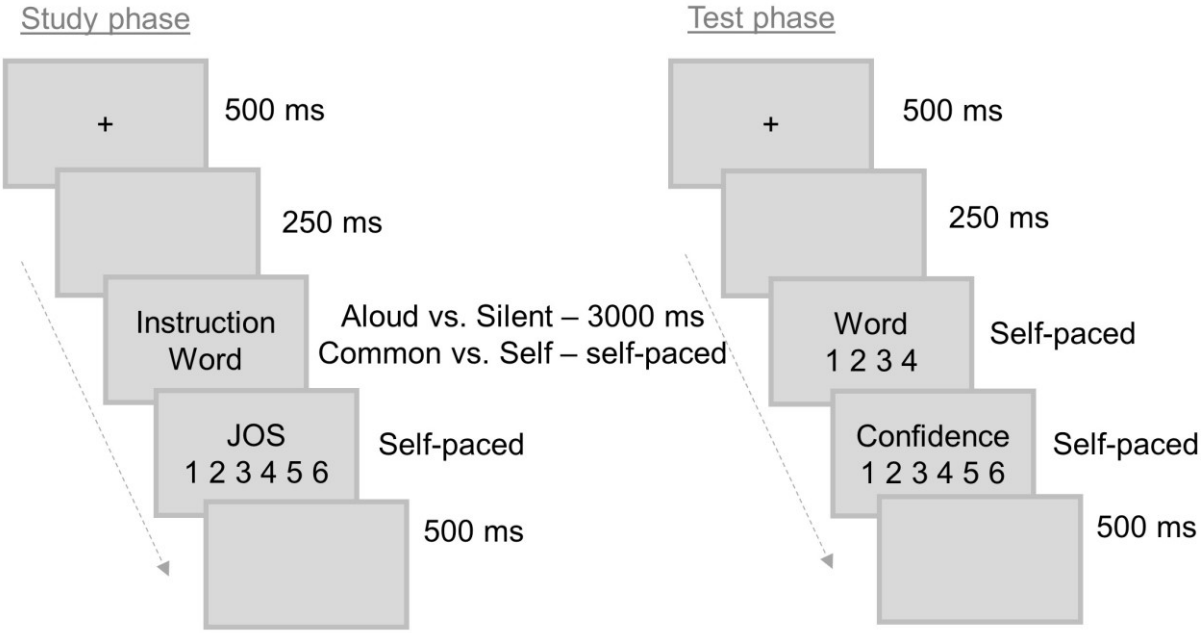


Figure 1. Schematic illustration of the tasks used in Experiments 1 and 2. Note: JOS = Judgment of Source.

The study phase was immediately followed by a test phase in which the previously 30 read words were randomly mixed with 30 new words. A trial started with a fixation cross (500 ms) and a blank screen (250 ms); then a word was presented in the center of the screen, and participants had to perform old-new judgments regarding the source by choosing one of four response options: 1) "read silently"; 2) "read aloud"; 3) "read, but do not know if silently/aloud"; 4) "new". The "do not know" option was adopted following previous studies (Dulas & Duarte, 2014; Leshikar & Duarte, 2012, 2014; Newsome et al., 2012; Sharot & Yonelinas, 2008) with the purpose of reducing guessing and response bias. There was no time limit to provide a response. The response keys were counterbalanced across participants. After the old-new judgment, participants were asked to evaluate the degree of confidence in their response using, once again, a six-point rating scale (1 = "No confidence"; 6 = "Complete confidence"). The confidence judgment was also self-paced. The last event of the trial was a blank screen lasting 500 ms (see Figure 1).

The lists were counterbalanced across participants regarding study/test and read aloud/silent to ensure that each stimulus could be read aloud or silently and could be an old or a new item. Moreover, prior to the experimental task, all participants completed a training period with a brief study-test cycle, and they were instructed to pay attention to the words and their source (aloud or silent), because they would be tested later (intentional learning). Participants were also informed that each study-test cycle was independent, i.e., after finishing a cycle they could forget the stimuli and focus their attention on the new cycle. Stimulus presentation was controlled using SuperLab software (Version 5, Cedrus Corporation, San Pedro, CA, https://www.cedrus.com/superlab/).

Data analysis.

Initially, the proportion of responses during recognition was calculated for six main categories: 1) correct source (the source was correctly identified; for example, the participant recognized an item as "read aloud" when it was indeed paired with a "read aloud" instruction during the study phase); 2) incorrect source (the participant misattributed the source; for example, for an item that was paired with a "read aloud" instruction during encoding, the participant selected "read silently" during test); 3) do not know source (the participant correctly recognized an item as old, but did not remember if it was read silently or aloud); 4) miss (for an item presented during the study phase, the participant misidentified it as "new" during test); 5) correct rejection (the participant correctly identified a new item as "new"); 6) false alarm (the participant considered a new item as "old" by choosing one of the sources or the "do not know" option). The item memory recognition accuracy was then obtained using Pr = ([p(hits) – p(false

alarms)]) (Snodgrass & Corwin, 1988), wherein the p(hits) resulted from combining both correct and incorrect source responses and the do not know responses. The response bias also followed Snodgrass and Corwin (1988): Br = ([p(false alarms)/(1 – Pr)]). A conservative response bias is considered for Br values below 0.50, whereas a liberal response bias is considered for values above 0.50. In a first attempt to compute the Br, the denominator was equal to zero for some participants. To circumvent this problem, we applied a correction to both hit and false alarm rates, following the same authors [hit rates = (number of hits + 0.5) / (number of old items + 1); false alarm rates = (number of false alarms + 0.5) / (number of new items + 1)]. In the case of source recognition, the discrimination measure was obtained by subtracting incorrect source from correct source responses - [p(correct source) – p(incorrect source)] – excluding the "do not know" responses and following prior studies (e.g., Dulas & Duarte, 2014; Leshikar et al., 2015; Newsome et al., 2012). Both item and source discrimination measures were submitted to a 3 (valence: negative/neutral/positive) x 2 (source: aloud/silent) RMANOVA.

It is worth noting that the old-new recognition patterns and the source memory discrimination are somehow superimposed and confounded in this study. Moreover, the computed measures are amenable to response bias. Multinomial models represent a useful approach to separate between old-new detection and source discrimination, by additionally considering the effect of potential biases (Batchelder & Riefer, 1990). Thereby, multinomial models were computed as a supplementary analysis to back up the results of the RMANOVA. A full description of the procedure and results is presented in the Supplementary Materials section.

The effects of source and emotional content on the JOSs during encoding were evaluated through a 3 (valence) x 2 (source) RMANOVA. Moreover, Goodman-Kruskal gamma correlations were used as indices of metamnenomic accuracy for each experimental condition in accordance with previous literature (e.g., Castel et al., 2013). The procedure used to compute gamma is described in the Supplementary Materials section. Additional findings regarding the proportion of incorrect source responses, "do not know" responses, misses, correct rejections, (corrected) false alarms, recognition confidence for source judgments and correct rejections, as well as the response times for JOSs and recognition responses are presented in the Supplementary Materials section.

A Bayesian two-way RMANOVA was conducted (JASP Team, 2018, Version 0.9.0.1) to complement the abovementioned analysis (Wagenmakers, Love et al., 2018; Wagenmakers, Marsman et al., 2018; see Supplementary Table S7 and Table S8). Specifically, Bayes factors (BF_{10}) were considered in favor of the alternative hypothesis if BF_{10} was larger than three, whereas a BF_{10} lower than 0.3 was interpreted as evidence in favor of the null hypothesis (Quintana & Williams, 2018; Wagenmakers, Love et al., 2018).

Finally, the Greenhouse-Geisser method was used as correction for sphericity violations, and Bonferroni-corrected post hoc comparisons were applied to qualify interaction effects.

Results

Descriptive statistics regarding the behavioral performance in Experiment 1 are shown in Table 1.

Recognition accuracy.

Item memory. The RMANOVA yielded a main effect of valence, $F(2, 54) = 4.08$, $p = .022$, $\eta_p^2 = .13$, a main effect of source, $F(1, 27) = 15.93$, $p < .001$, $\eta_p^2 = .37$, but no interaction effect between the two factors, $F(2, 54) = 0.34$, $p = .716$, $\eta_p^2 = .01$. Bayes factors (see Supplementary Table S7) favored the two main effects model in relation to the null model ($BF_{10} = 671.057 \pm 2.21\%$). The comparison between this model and the model that adds the interaction term (671.247/82.160 = 8.17) revealed evidence against the interaction as data were 8.17 more likely in the two main effects model than in the model with the interaction. Specifically, the pairwise comparisons showed a marginally significant difference in the recognition of negative and positive words ($p = .060$, $d = 0.47$, 95% CI [-0.002, 0.14]): negative words ($M = .79$, $SE = .03$) were more accurately recognized than positive words ($M = .72$, $SE = .02$). The words read aloud ($M = .79$, $SE = .02$) were also more accurately recognized than words read silently ($M = .72$, $SE = .02$; $p < .001$, $d = 0.75$, 95% CI [0.03, 0.10]; see Table 1 and Figure 2B).

In the case of the corrected item Br measure, the RMANOVA showed a main effect of valence $F(2, 54) = 3.64$, $p = .033$, $\eta_p^2 = .12$, a main effect of source, $F(1, 27) = 11.79$, $p = .002$, $\eta_p^2 = .30$, yet no interaction effect, $F(2, 54) = 1.03$, $p = .363$, $\eta_p^2 = .04$. Again, the Bayesian analysis supported the two main effects model when compared to the null model ($BF_{10} = 42.206 \pm 1.50\%$). Additionally, when compared to the model with the interaction term (42.206/7.965), the data were 5.30 times more likely in the two main effects model (see Supplementary Table S7). Specifically, there was a statistically significant difference between neutral and positive words ($p = .035$, $d = 0.51$, 95% CI [0.005, 0.16]), indicating that participants used a less conservative response criterion for positive words ($M = .43$, $SE = .04$) in comparison with neutral words ($M = .35$, $SE = .04$). Also, the response criterion was less conservative for words read aloud ($M = .44$, $SE = .04$) than for words read silently during encoding ($M = .37$, $SE = .04$; $p = .002$, $d = 0.65$, 95% CI [0.029, 0.12]). However, the response criterion for all conditions was on average below or near 0.50, thus indicating a general conservative to neutral response bias (see Table 1; Snodgrass and Corwin, 1988).

Source memory. The results regarding the source memory measure did not follow the same pattern as the item memory, as no statistically significant effects were observed [valence: $F(2, 54) = 0.42$, $p = .657$, $\eta_p^2 = .02$; source: $F(1, 27) = 1.37$, $p = .252$, $\eta_p^2 = .05$; interaction: $F(2, 54) = 0.49$, $p = .618$, $\eta_p^2 =. 02$)]. Bayesian analysis supported the null effects (valence: $BF_{10} = 0.072 \pm 0.86\%$; source: $BF_{10} = 1.027 \pm 4.28\%$; valence + source + interaction: $BF_{10} = 0.009 \pm 1.75\%$). Thus, the source memory recognition measures were similar irrespective of stimulus valence and of the encoding strategy (see Table 1 and Figure 2C).

Judgments of source (JOSs).

When assessing how valence and way of production affected the JOSs ratings, we observed a main effect of valence, $F(2, 54) = 15.93$, $p < .001$, $\eta_p^2 = .37$, a main effect of source, $F(1, 27) = 59.61$, $p < .001$, $\eta_p^2 = .69$, but no interaction effect, $F(2, 54) = 0.44$, $p = .649$, $\eta_p^2 = .02$. The Bayes factors supported the two main effects model with no interaction (valence + source: $BF_{10} = 1.620 \times 10^{23} \pm 1.25\%$; 8.29 times more likely than the interaction model). Specifically, the words read aloud ($M = 4.23$, $SE = 0.13$) were regarded as more memorable than the words read silently ($M = 3.44$, $SE = 0.10$; $p < .001$, $d = 1.46$, 95% CI [0.58, 1.00]), and both negative ($M = 3.98$, $SE = 0.11$; $p < .001$, $d = 1.05$, 95% CI [0.19, 0.50]) and positive stimuli ($M = 3.88$, $SE = 0.12$; $p = .001$, $d = 0.80$, 95% CI [0.10, 0.40]) were rated as more memorable when compared to neutral stimuli ($M = 3.64$, $SE = 0.10$; see Table 1 and Figure 2A).

In general, the metacognitive judgments partially agree with the results of item memory, as words read aloud were better recognized than words read silently. However, the predictions in the case of valence were less accurate, since no significant differences were found between emotional conditions. Considering that the JOSs ratings were performed in relation to the source, no match was observed between the predictions and the source memory performance. The analysis of the metamnemonic accuracy revealed that the gamma correlations for neutral/positive words that were read aloud and for negative words read silently were statistically different from zero, $t(27) \geq 2.29$, $p \leq .03$, whereas for the remaining conditions the gamma correlations did not differ from zero, $t(27) \leq 1.78$, $p \geq .09$. On the one hand, some experimental conditions were different from zero, supporting a possible relation between immediate metamemory judgments and recognition. Additionally, even though the mean gamma coefficients were positive and small (see Table 1), they resemble what has been reported in the literature

Table 1

Behavioral Performance (Source x Valence) in Experiments 1 and 2

	Experiment 1 – *M* (*SD*)						Experiment 2 – *M* (*SD*)					
	Aloud			Silent			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
Source correct	0.61 (0.16)	0.55 (0.20)	0.58 (0.19)	0.59 (0.22)	0.57 (0.24)	0.55 (0.24)	0.64 (0.19)	0.76 (0.20)	0.79 (0.17)	0.47 (0.23)	0.58 (0.26)	0.44 (0.24)
Source incorrect	0.14 (0.12)	0.11 (0.14)	0.11 (0.14)	0.05 (0.08)	0.05 (0.08)	0.05 (0.08)	0.08 (0.10)	0.05 (0.09)	0.06 (0.10)	0.08 (0.08)	0.04 (0.06)	0.14 (0.12)
Source DNK	0.18 (0.17)	0.22 (0.19)	0.21 (0.16)	0.24 (0.17)	0.21 (0.22)	0.22 (0.17)	0.14 (0.14)	0.08 (0.10)	0.08 (0.11)	0.19 (0.20)	0.17 (0.19)	0.21 (0.18)
Source miss	0.07 (0.10)	0.11 (0.08)	0.09 (0.08)	0.12 (0.12)	0.17 (0.16)	0.17 (0.14)	0.15 (0.13)	0.11 (0.10)	0.07 (0.11)	0.25 (0.15)	0.21 (0.15)	0.21 (0.17)
Item Pr measure	0.82 (0.13)	0.78 (0.13)	0.76 (0.13)	0.77 (0.15)	0.72 (0.16)	0.68 (0.15)	0.68 (0.14)	0.76 (0.14)	0.74 (0.17)	0.59 (0.15)	0.67 (0.17)	0.60 (0.20)
Item Br measure	0.47 (0.22)	0.36 (0.20)	0.48 (0.26)	0.38 (0.23)	0.33 (0.25)	0.39 (0.23)	0.43 (0.25)	0.43 (0.23)	0.57 (0.26)	0.33 (0.21)	0.33 (0.21)	0.39 (0.26)
Source measure	0.46 (0.20)	0.44 (0.27)	0.46 (0.27)	0.54 (0.24)	0.52 (0.26)	0.49 (0.27)	0.56 (0.25)	0.71 (0.26)	0.72 (0.22)	0.40 (0.27)	0.54 (0.28)	0.30 (0.31)
JOSs ratings	4.40 (0.68)	4.02 (0.71)	4.26 (0.76)	3.56 (0.63)	3.25 (0.52)	3.50 (0.60)	4.05 (0.85)	4.12 (0.85)	4.34 (0.82)	3.91 (0.72)	3.94 (0.69)	3.96 (0.79)
Gamma (JOSs)	0.18 (0.53)	0.29 (0.51)	0.31 (0.49)	0.19 (0.45)	0.19 (0.58)	0.14 (0.48)	0.15 (0.47)	0.16 (0.50)	0.14 (0.52)	0.11 (0.53)	0.06 (0.45)	0.07 (0.48)

	Negative	Neutral	Positive	Negative	Neutral	Positive
CR	0.92 (0.09)	0.93 (0.10)	0.89 (0.11)	0.88 (0.12)	0.91 (0.09)	0.85 (0.15)
FA	0.08 (0.09)	0.07 (0.10)	0.11 (0.11)	0.12 (0.11)	0.09 (0.09)	0.15 (0.15)

Note. CR = Correct Rejections; DNK = Do Not Know; FA = False Alarms; JOSs = Judgments of Source.

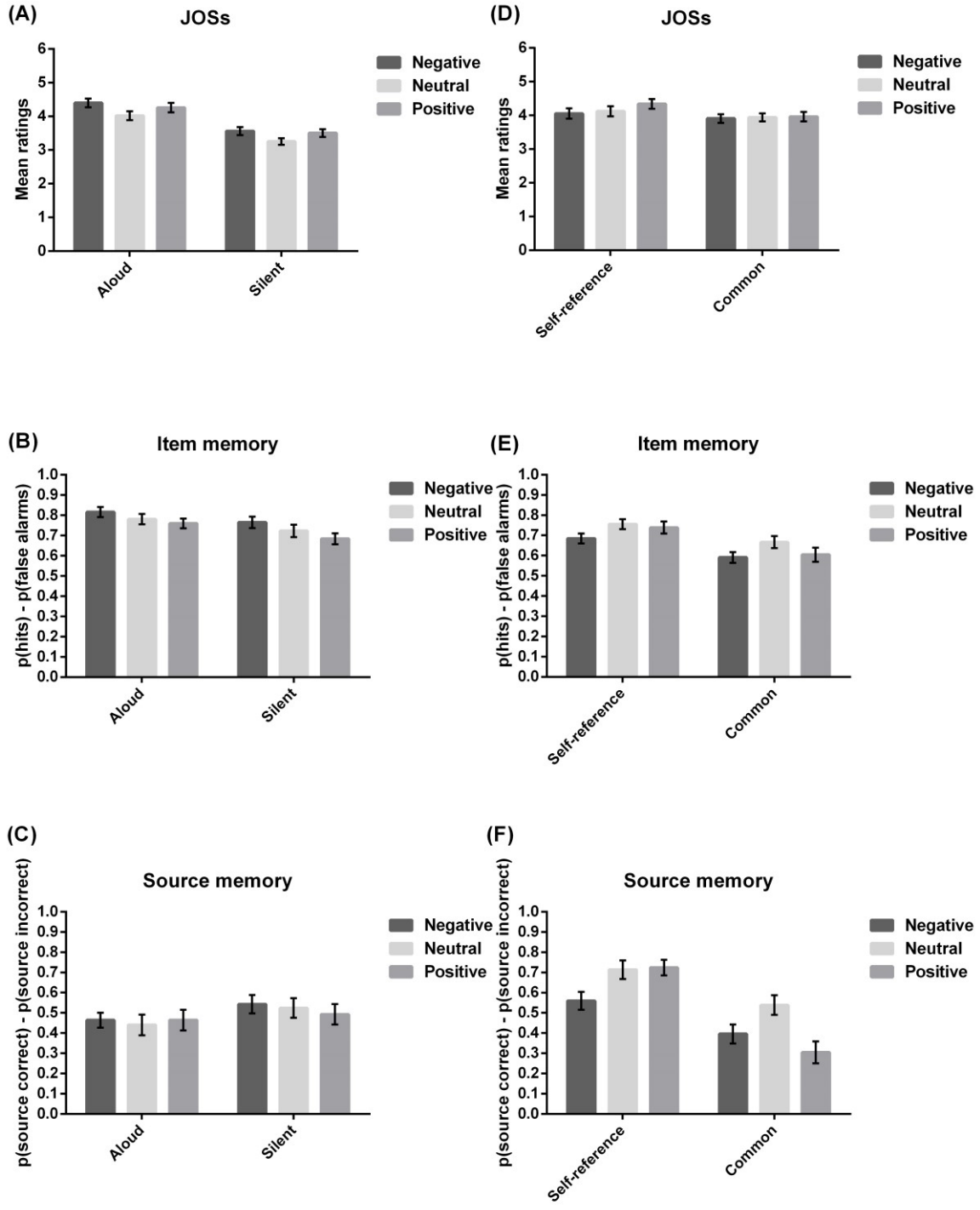


Figure 2. Behavioral results for JOSs, item memory and source memory in Experiments 1 and 2. Specifically, the behavioral results for judgments of source (JOSs; panel A and D), item memory (panel B and E) and source memory (panel C and F) are plotted on the y-axis as a function of source (aloud/silence; self-reference/common) and valence (negative/neutral/positive). Note: the error bars indicate the standard errors of the means.

of metamnemonic accuracy in the case of source memory (e.g., Carroll, Mazzoni, Andrews, & Pocock, 1999). On the other hand, no statistically significant differences between conditions were observed for gamma correlations (valence: $F(2, 54) = 0.14$, $p = .866$, $\eta_p^2 = .01$; source: $F(1, 27) = 1.06$, $p = .313$, $\eta_p^2 = .04$; interaction: $F(2, 54) = 0.61$, $p = .549$, $\eta_p^2 = .02$). Bayes factors were also in agreement (valence: $BF_{10} = 0.069 \pm 0.86\%$; source: $BF_{10} = 0.296 \pm 1.18\%$; valence + source + interaction: $BF_{10} = 0.003 \pm 2.86\%$). Thus, the modulatory role of encoding strategy and valence in the associations between the prospective source predictions and the actual source memory performance was inconclusive.

Discussion: Experiment 1

A major finding of Experiment 1 was the absence of a production effect on internal source memory. Nonetheless, as expected, the effect was replicated in the case of item memory (Dodson & Schacter, 2001; MacLeod et al., 2010; Ozubko & MacLeod, 2010; Quinlan & Taylor, 2013). Moreover, there was no interaction between valence and production mode, indicating that the production effect on item memory occurred regardless of whether stimuli had or not an emotional quality. The EEM on item memory was also absent, against our predictions and the main tendency in the literature (see Kensinger & Kark, 2018, Kensinger & Schacter, 2008, Levine & Edelstein, 2009, and Murphy & Isaacowitz, 2008 for overviews), especially if we take into consideration that both positive/negative words were rated as significantly more arousing than neutral words. Even so, this result is not unprecedented (e.g., Ferré et al., 2019; Hourihan & Bursey, 2017). Considering the response bias, participants revealed a conservative to neutral response criterion (see Table 1). Specifically, in cases of uncertainty, participants showed an increased tendency to regard an item as "new" when it was studied in the "read silently" condition than when it was studied in the "read aloud" condition. A similar tendency was observed for neutral items when compared to positive items. Although valence did not significantly affect objective memory indices, the same was not observed for subjective metamemory judgements. In fact, participants regarded the source of emotional words that were read aloud as more memorable, providing support to previous studies with JOLs (Castel et al., 2013; Hourihan et al., 2017; Jönsson et al., 2005; Nomi et al., 2013; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). However, contrary to the findings of Dutton and Carroll (2001), the JOSs predictions were not associated with internal source memory performance, which favors the notion that metamemory judgments are not always predictive of future memory (e.g., Carroll et al., 1999). Rather, they show that participants were sensitive to variations in the encoding conditions, and that they used these cues to perform metamemory judgments (Koriat, 1997; Mazzoni & Nelson, 1995; Nomi et al., 2013).

Experiment 2

The experimental design was identical to Experiment 1, although a different encoding strategy was tested: the self-referential processing. Given that in Experiment 1 negative and positive words were more arousing than neutral words, in Experiment 2 arousal ratings were also equated across valence categories. The control of arousal is advantageous considering that the effects of valence and arousal on memory function appear to be supported by different processing mechanisms (valence effects have been associated with more effortful semantic and autobiographical elaboration processes, whereas arousal effects have been related to more automatic processes; Cook, Hicks, & Marsh, 2007; Kensinger & Corwin, 2004). Moreover, these differential effects also seem to be supported by distinct neurofunctional mechanisms (prefrontal cortex-hippocampal interactions in the case of valence, and amygdala-hippocampal interactions in the case of arousal; Kensinger & Corwin, 2003, 2004). Hence, the control of arousal allows isolating valence-related effects. Specifically, in this experiment, emotional (negative/positive) and neutral adjectives with medium arousal ratings were used as stimuli. During the encoding phase, participants performed two types of judgments. In half the trials, they were required to judge if the word was related to them. In the other half, they were asked if the word was commonly used by people in everyday life. The latter task represents a non-self-referential condition that requires a deep, semantic analysis of the word, and it was based on previous research (e.g., Hamami et al., 2011; Leshikar et al., 2015; Newsome et al., 2012; Serbun et al., 2011). Participants were also instructed to perform JOSs for each word. During the test phase, participants performed old-new judgments in which they were additionally asked to identify the source of the item (self-reference *vs.* common).

Materials and methods

Participants.

A total of 32 young adults (30 females), aged between 18 and 45 years (M = 22.59, SD = 5.89), and with an average of 14.45 years of formal education (SD = 2.35) participated in this experiment. The same inclusion criteria described in Experiment 1 were applied to Experiment 2. The sample size estimation followed the same strategy as described in Experiment 1 (i.e., within-factors RMANOVA as the main statistical test, η_p^2 as measure of effect size, similar experimental design, and the same parameters of alpha significance level and power), but a recent study by Durbin and colleagues (2017, Experiment 1) was used as reference: here, an effect size of .16 was reported for source memory performance (24 participants were tested). G*Power (Faul et al., 2007) indicated that a minimum of 16 participants would

be necessary to allow the detection of such effects. Hence, an adequate sample size was also tested in the current experiment.

Materials.

A total of 144 adjectives (48 negative/neutral/positive) were selected from the Portuguese version of the ANEW (Soares et al., 2012). The valence ratings differed between the three types of words (positive > neutral; positive > negative; neutral > negative; all $p <$. 001), but no further differences were observed regarding arousal, frequency, number of letters and number of syllables ($p >$ .05; see Supplementary Table S1). The selected stimuli were then equally divided among four lists of 36 words each (12 stimuli from each valence category) to be used in the study and test phases. The lists were similar regarding valence, arousal, frequency, number of letters, and number of syllables ($p >$ .05).

Procedure.

The overall procedure, structure, and characteristics of the experimental task were similar to Experiment 1 (see Figure 1), but with the following differences: a) only two study-test cycles were used; b) during the encoding phase, the participants performed two different subjective judgments - a self-referential judgment (in which they assessed whether each stimulus related somehow to their personal characteristics: self-referential condition) or a common judgment (in which they evaluated whether the word was commonly used by people in their everyday lives: non-self-referential condition); c) the instructions that appeared together with the words were “Does this word describe me?” or “Is this word common?”; d) during the encoding phase, the participants were prompted to respond “yes” (press key “Z”) or “no” (press key “M”) in accordance to the “self-reference” or “common” instructions, and there was no time limit to produce a response (self-paced); e) during the test phase, the participants were instructed to select one of four response options: “self-description”; “common”; “evaluated, but do not know if self-description/common”; “new”.

Data analysis.

The same data analysis procedures planned for Experiment 1 were adopted in Experiment 2.

Results

The main descriptive statistics for the behavioral performance in Experiment 2 are shown in Table 1.

Recognition accuracy.

Item memory. The analysis of the corrected item memory Pr measure showed a main effect of valence, F (2, 62) = 5.46, p = .012, η_p^2 = .15, ε = .78, a main effect of source, F(1, 31) = 41.63, p < .001, η_p^2 = .57, but no valence x source interaction effect, F(2, 62) = 1.15, p = .314, η_p^2 = .04, ε = .76. The two main factors model received the strongest support (BF_{10} = 8.708 x 10^7 ± 1.18%; see Supplementary Table S8), and it was preferred to the interaction model (8.708 x 10^7/1.582 x 10^7 = 5.50). The participants were better at recognizing neutral words (M = .71, SE = .03) relative to negative words (M = .64, SE = .02; p < .001, d = 0.82, 95% CI [0.03, 0.11]). They also demonstrated a better recognition performance for words encoded in the self-referential condition (M = .73, SE = .02) than for words studied in the common condition (M = .62, SE = .03; p < .001, d = 1.14, 95% CI [0.07, 0.14]; see Table 1 and Figure 2E).

Regarding the corrected item Br measure, the RMANOVA revealed a main effect of valence, F (2, 62) = 9.26, p < .001, η_p^2 = .23, a main effect of source, F(1, 31) = 40.97, p <. 001, η_p^2 = .57, yet no interaction effect, F(2, 62) = 2.33, p = .106, η_p^2 = .07. Again, Bayes Factors were in favor of the two main effects model in contrast to the null model (BF_{10} = 3.092 x 10^8 ± 2.06%; see Supplementary Table S8). This model was also favored when compared to the model with the interaction term (3.092 x 10^8/1.201 x 10^8 = 2.57). Specifically, the participants demonstrated a more conservative response criterion for both negative (M = .38, SE = .04) and neutral words (M = .38, SE = .02) when compared to positive words (M = .48, SE = .04; p = .003, d = 0.63/0.64, 95% CI [0.03, 0.18] considering both comparisons). Moreover, participants used a less conservative response criterion for words encoded in the self-referential condition (M = .48, SE = .04) compared to words studied in the common condition (M = .35, SE = .04; p < .001, d = 1.13, 95% CI [0.09, 0.17]; see Table 1).

Source memory. The RMANOVA yielded a main effect of valence, F(2, 62) = 15.17, p < .001, η_p^2 = .33, ε = .79, a main effect of source, F(1, 31) = 39.58, p < .001, η_p^2 = .56, and a valence x source interaction effect, F(2, 62) = 10.69, p < .001, η_p^2 = .26. Positive evidence in favor of the interaction model over the two main effects model was also obtained with Bayes factors (1.007 x 10^{13}/9.218 x 10^{14} = 0.01). Specifically, the interaction effect yielded a differential pattern of results when considering the self-referential and the common condition. Whereas the source of negative words encoded in the self-referential condition was more poorly recognized than the source of both neutral (95% CI [0.07, 0.23]) and positive words (95% CI [0.05, 0.28]), the source of both negative (95% CI [0.04, 0.25]) and positive

words (95% CI [0.13, 0.34]) was less accurately recognized than the source of neutral words in the case of the common condition ($p < .01$; see Table 1 and Figure 2F).

Judgments of source (JOSs).

Concerning the JOSs ratings, the RMANOVA indicated a main effect of valence, $F(2, 62) = 4.97$, $p = .020$, $\eta_p^2 = .14$, $\varepsilon = .72$, a main effect of source, $F(1, 31) = 9.70$, $p = .004$, $\eta_p^2 = .24$, and a valence x source interaction effect, $F(2, 62) = 3.39$, $p = .040$, $\eta_p^2 = .10$. The Bayes factors did not support the interaction model as the two main effects model gathered the most robust evidence against the null model ($BF_{10} = 1888.860 \pm 1.90\%$; 1888.860/935.928 = 2.02). The results indicate that participants rated self-referentially words as more memorable ($M = 4.17$, $SE = 0.14$) than words studied in the common condition ($M = 3.94$, $SE = 0.12$; $p = .004$, $d = 0.55$, 95% CI [0.08, 0.39]; see Table 1 and Figure 2D). The valence effect was inconclusive: Bayes factors only showed anecdotal evidence in favor of the valence model ($BF_{10} = 1.029 \pm 0.73\%$); the pairwise comparisons revealed marginally significant differences between the JOS ratings of positive words and both neutral ($p = .074$, $d = 0.42$, 95%CI [-0.01, 0.25]) and negative words ($p = .064$, $d = 0.43$, 95% CI [-0.01, 0.35]).

The metamnemonic judgments regarding source memory during encoding only partially complied with the pattern of results obtained in the recognition test. Although JOSs ratings indicated that self-referentially encoded words were regarded as more memorable than words studied in the common condition, which was in line with the self-reference benefit observed in the recognition test, the JOS ratings concerning valence were unclear. Additionally, the gamma correlations revealed no reliable association between source judgments and source recognition as the coefficients did not differ from zero, $t(31) \leq 1.86$, $p \geq .07$. Furthermore, no significant differences were found between conditions (valence: $F(2, 54) = 0.10$, $p = .905$, $\eta_p^2 = .003$; source: $F(1, 27) = 0.80$, $p = .378$, $\eta_p^2 = .03$; interaction: $F(2, 54) = 0.09$, $p = .917$, $\eta_p^2 = .003$), and the coefficients remained positive yet small (see Table 1). The null effects were supported by the Bayesian analysis (valence: $BF_{10} = 0.058 \pm 1.06\%$; source: $BF_{10} = 0.287 \pm 0.72\%$; interaction: 0.002/0.016 = 0.13). Thus, the results pointed to the lack of an association between source memory predictions and memory performance.

Discussion: Experiment 2

By changing the encoding strategy from Experiment 1 to Experiment 2, we observed a differential influence of valence on internal source memory, item memory, and JOSs. As expected, a self-referential

benefit on item memory was observed (Kuiper & Rogers, 1979; Rogers et al., 1977; Symons & Johnson, 1997), and this occurred irrespective of stimulus valence. The item memory performance was also enhanced for neutral relative to negative words (yet no significant differences were observed in the case of positive words), and this was observed in both encoding conditions. Thus, we partially replicated prior evidence showing that neutral stimuli encoded in a self-referential manner are better remembered than negative stimuli (Yang et al., 2012). However, the hypothesized difference between negative and positive stimuli (based on prior studies - Yang et al., 2012; Leshikar et al., 2015) was not observed. Notwithstanding, the EEM was not replicated in the common condition. The response bias also changed with the encoding strategy: in cases of uncertainty, participants showed an increased tendency to respond "new" for both negative and neutral stimuli relative to positive stimuli; participants also used a more conservative response criterion for stimuli encoded in the common condition than for stimuli encoded in the self-referential condition. On average, participants appeared to show a conservative to neutral response criterion (see Table 1). The only exception was observed in the case of self-referentially encoded positive words, whose mean Br was above 0.50. In the case of internal source memory performance, an interaction effect between valence and encoding strategy emerged, revealing that source recognition was improved for both neutral and positive words compared to negative stimuli in the self-reference task, which partially corroborates previous findings (e.g., Durbin et al., 2017; Zhang et al., 2018). A different pattern was observed in the common task, in which emotional words led to reduced source memory accuracy in contrast to neutral words. Specifically, valence impaired internal source memory, in agreement with prior studies (see Appendix). The JOSs ratings were also sensitive to self-referential processing as self-referentially encoded words were regarded as more memorable than words studied in the common condition. Regarding valence, the effect was less clear, which stands in contrast with Experiment 1 and previous studies (e.g., Hourihan et al., 2017; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). Nevertheless, no significant associations were found between the JOSs ratings and internal source memory performance.

General discussion

The main goal of this study was twofold: to explore how stimulus valence influences internal source monitoring, specifically when the encoding conditions are also characterized by memory enhancing features (production and self-reference effect); to specify the role played by stimulus valence in immediate prospective judgments concerning internal source memory (JOSs ratings). Overall, the results revealed that internal source monitoring, item memory, and JOSs were differently modulated by the encoding

strategy, which was supported by participant-based results (RMANOVA), Bayes factors, and multinomial models (see Supplementary Materials). Specifically, in Experiment 1, the effects of valence on item and source memory were not significant, even though the production effect was replicated for item memory. The JOSs ratings were sensitive to both valence and production mode. In turn, Experiment 2 revealed that the self-referential processing enhanced item and source memory differently as a function of valence. The JOSs ratings were also sensitive to the self-reference effect.

The current study supported the notion that self-referential conditions benefit both item and internal source memory (Durbin et al., 2017; Hamami et al., 2011; Kalenzaga et al., 2015; Leshikar and Duarte, 2012, 2014; Serbun et al., 2011; Yin et al., 2018; Zhang et al., 2018). Nonetheless, in the case of the production effect, the beneficial effects were only demonstrated for item memory but not for source memory, which stands in contrast with previous evidence (Ozubko et al., 2014). Specific methodological factors may have attenuated the production effect such as intentional encoding instructions, the use of the "do not know" option in the test phase (which may have led to a more conservative response in situations in which the participants were not able to recollect the source but they knew the words were familiar), or the source recognition measure that included the proportion of incorrect source attributions. Indeed, we observed that the proportion of incorrect source attributions for words that were read aloud was significantly higher than for words read silently. As participants believed that words read aloud were more memorable than words read silently (which was reflected in JOSs ratings), they may have been more prone to attribute a specific word read aloud to the other source when they only knew that a specific word was studied before. Nonetheless, a lack of effect was also expected, given that prior studies that probed the effects of valence on internal source memory also reported a null (e.g., Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006a; Sharot & Yonelinas, 2008) or an impairment effect (e.g., Ferré et al., 2019, Experiment 2 and 3; Mao, et al., 2015; Newsome et al., 2012; Otani, Jaffa, et al., 2012; Otani, Libkuman, et al., 2012; see Appendix). Critically, both results were observed in the current study but with distinct encoding conditions, i.e., no effect in Experiment 1 and an impairment effect in the non-self-referential condition in Experiment 2. Furthermore, an enhancement effect emerged for both neutral and positive words studied in a self-referential manner.

Whereas an impairment or lack of effect might be explained by the object-based binding theory (Mather, 2007), as production mode and cognitive operations (self-referential or common judgments) can be deemed as extrinsic features of the item, this theoretical framework cannot account for the self-referential benefit of source memory in the case of neutral and positive words, especially considering that stimuli in Experiment 2 were matched for arousal. In this context, the arousal-biased competition theory

(Mather & Sutherland, 2011) offers a more suitable framework for the source memory effects found in the current study, considering that both emotional and non-emotional information can be prioritized according to the current goals and motivations of the individual. Even though both positive and negative stimuli were more arousing than neutral stimuli in Experiment 1, the participants were instructed to effortfully encode both item and production mode. Hence, even if the items that were read aloud were more perceptually salient, all items were goal-relevant and likely to receive similar processing resources.

Considering that self-referential processing promotes a more organized, elaborative and efficient processing in comparison with other perceptual and semantic tasks, such as common judgments (which can be deemed as unusual tasks; D'Argembeau et al., 2005; Symons & Johnson, 1997), participants may have prioritized information that was relevant to describe themselves in Experiment 2. In fact, words that are perceived as self-descriptive tend to be better remembered than non-self-descriptive words (Kuiper & Rogers, 1979). Participants also tend to favor positive information and to disregard negative self-referential information (D'Argembeau et al., 2005; Watson, Dritschel, Jentzsch, & Obonsawin, 2008; Zhang et al., 2018), a tendency that has been observed irrespective of age, gender, or cultural background (Mezulis, Abramson, Hyde, & Hankin, 2004). Moreover, the information that fits the current self-scheme may receive a deeper and elaborative processing, whereas non-fitting information may be processed in a shallow manner, which may result in a less successful memory performance (D'Argembeau et al., 2005; Kuiper & Rogers, 1979). Thus, enhanced source memory might be expected for both neutral and positive self-referential words as they are likely to match the current self-schema (the proportion of “yes” responses was also higher for positive and neutral words; see Supplementary Materials). In contrast, the negative self-referential words may have been neglected as they are less likely to match the current self-schema, leading to a decline in source memory performance.

In the case of the common condition, a similar result to Experiment 1 could be expected, as identical intentional encoding conditions were required during the learning phase. However, an impairment effect was observed for emotional words encoded in the common condition. This finding might be accounted for by the higher proportion of incorrect source attributions found for positive words in comparison with neutral words (see Supplementary Materials). Although this represents a limitation of the current study, it might be the case that the emotional adjectives selected here were more prototypical self-descriptors than the neutral adjectives. Thereby, participants could be more likely to confound the source of emotional adjectives studied in the common condition. Likewise, the neutral adjectives could offer a more congruent match with the common task, resulting in a better source memory performance. According to the arousal-biased competition theory (Mather & Sutherland, 2011), the former explanations

might be related to top-down processes such as the relevance of the stimuli to the current task. Additionally, they also fit with the source-monitoring account (Johnson et al., 1993), as decisions regarding the origin of an event are influenced not only by qualitative characteristics that take place during encoding (e.g., cognitive operations), but also by previous general knowledge and schemas (e.g., stereotypes; beliefs). Thus, when there are few qualitative features to distinguish between sources, which is the case of internal source memory decisions, participants might rely on their beliefs. In the test phase, if participants consider a specific word as a self-descriptor (which might be more likely to happen in the case of positive adjectives), they might be biased to select the self-reference option, especially when they cannot recall other qualitative features associated with the word, even when a "do not know option" is available. In fact, positive adjectives that were self-referentially encoded were also associated with a more liberal response criterion (see Table 1), suggesting that participants were biased to deem positive new items as "old".

The subjective JOSs ratings were also informative of how sensitive the participants were to both valence and encoding strategy during the encoding phase. In Experiment 1, the source of both positive and negative words was regarded as more memorable than neutral words. This pattern has also been reported in studies using JOLs (e.g., Hourihan & Bursey, 2017; Hourihan et al., 2017; Jönsson et al., 2005; Nomi et al., 2013; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). Here, it was extended to JOSs. As the processing of emotional stimuli shows some advantages in comparison with neutral stimuli in terms of attention, organization, and distinctiveness (D'Argembeau & Van der Linden, 2004; Maddox et al., 2012; Talmi, 2013), these factors may also play a role in metamemory judgments. Furthermore, the source of words that were read aloud was rendered as more memorable than words read silently, following previous studies examining JOLs (Castel et al., 2013). Also, in Experiment 2, JOSs ratings revealed that self-referentially encoded words were judged as more memorable when compared to words studied in the common condition. Notwithstanding, the effect of valence was inconclusive, and our initial predictions were not confirmed as only marginal results supported the preference of the participants for positive information. Overall, both experiments demonstrate that participants were sensitive to the potential advantage of reading items aloud or processing them self-referentially. This finding might be accounted for by the cue-utilization hypothesis (Koriat, 1997), which posits that prospective metamemory judgments can be modulated by specific characteristics that provide a sense of how easy/difficult it is to process and learn an item.

According to the cue-utilization hypothesis, JOLs are inferential in nature and their predictive value depends on whether the cues used to make metamemory judgments converge with variables affecting

memory performance. Thus, in some instances immediate metamemory judgments might predict future memory performance (e.g., Carroll, Davis, & Conway, 2001; Dutton & Carroll, 2001; Mazzoni & Nelson, 1995), whereas the same may not occur in other situations. For instance, in a previous study by Carroll and colleagues (1999), exploring JOLs and JOSs applied to reality monitoring decisions (seen *vs.* imagined), participants' predictions regarding source memory performance did not differ from chance. In the current study, this result is extended to internal source memory, specifically in Experiment 2 and in some conditions of Experiment 1. Although the predictions for some conditions of Experiment 1 were different from chance, no reliable differences in the gamma coefficients were found between encoding strategy and valence conditions. These findings stand in contrast with the findings of Dutton and Carroll (2001) but agree with the JOLs findings (Hourihan & Bursey, 2017; Hourihan et al., 2017). As suggested by Kelly, Carroll and Mazzoni (2002), JOSs can be considered an unfamiliar task and, plausibly, participants might have based their JOSs on item memorability rather than source memory memorability. In fact, it has been previously suggested that JOLs and JOSs are associated and that similar cues might be used to make these predictions (Carroll et al., 2001; Carroll et al., 1999; Kelly et al., 2002). As immediate prospective metamemory judgments are performed item by item, participants may rely their decisions on short-term memory and on other factors (e.g., the nature of the elaborative processes) that do not have enough diagnostic value to predict which information will be later remembered or forgotten.

The indirect analysis of item memory performance did not reveal an EEM effect, contrary to our initial prediction and several prior studies (e.g., Davidson et al., 2006; Kensinger & Schacter, 2006a; Maddock & Frein, 2009; Maddox et al., 2012; Schmidt et al., 2011; Schümann et al., 2018; Yang et al., 2012). Specifically, in Experiment 1, no difference was found when emotional words were compared to neutral words. In Experiment 2, neutral words were associated with higher recognition rates than negative words irrespective of the encoding strategy. Enhanced item memory for neutral compared to negative words encoded in the self-referential condition was expected (Yang et al., 2012). Nonetheless, a similar benefit was also expected for positive words that were processed self-referentially in relation to negative words (Leshikar et al., 2015; Yang et al., 2012), even though this prediction was not confirmed by the current study. A failure to observe such effect is rather in agreement with the studies of D'Argembeau and colleagues (2005, Experiment 2) and of Zhang and colleagues (2018, Experiment 2) that compared positive and negative traits. Even considering the consistency of the EEM effect in the literature (see Murphy & Isaacowitz, 2008 for a meta-analysis), other studies have also failed to report significant emotion effects on memory tasks (e.g., Ferré et al., 2019; Hourihan & Bursey, 2017; Newsome et al., 2012), particularly when using recognition tests (e.g., D'Argembeau & Van der Linden, 2004; Davidson

et al., 2006; Doerksen & Shimamura, 2001) and short intervals between study and test (e.g., Mitchell, Mather, Johnson, Raye, & Greene, 2006; Sharot & Yonelinas, 2008; Yonelinas & Ritchey, 2015; Wang, 2018). The null effect might also be explained by the higher proneness to false alarms observed in the case of emotionally salient words, especially when they are intentionally studied (Cook et al., 2007; D'Argembeau & Van der Linden, 2004; Davidson et al., 2006, Experiment 2B; Sharot & Yonelinas, 2008). Indeed, this explanation could partially account for the item memory difference between neutral and negative words in Experiment 2, irrespective of the encoding task, as the false alarm rates were higher for emotional compared to neutral words (see Supplementary Materials). Additionally, the negative adjectives studied in the self-referential condition were also associated with a greater proportion of misses compared to positive adjectives studied in the same condition (see Supplementary Materials). This last finding might have also contributed to the statistically significant difference between negative and neutral stimuli, and for the intermediate position of positive stimuli, which did not differ from both neutral and negative stimuli.

Although the former interpretation seems to apply to Experiment 2, the same cannot be assumed for Experiment 1, especially because the false alarm rates did not differ across negative, neutral and positive words (see Supplementary Materials). In this regard, the short interval between study and test may be an important factor to consider. Prior studies revealed that the relationship between the brain activity during encoding and the subsequent memory performance for neutral stimuli is stronger at a short interval than at a long interval, whereas the same relationship appears to be more stable over time in the case of emotional stimuli (Mickley Steinmetz, Schmidt, Zucker, & Kensinger, 2012). Moreover, emotional information seems to be more resistant to forgetting (Mitchell et al, 2006; Ritchey, Dolcos, & Cabeza, 2008; Schaefer, Pottage, & Rickart, 2011; Sharot & Phelps, 2004; Sharot & Yonelinas, 2008; Yonelinas & Ritchey, 2015; Weymar, Löw, & Hamm, 2011; Weymar, Löw, Melzig, & Hamm, 2009). As in Experiment 1 the study phase was immediately followed by the test phase, it could be the case that the EEM effect would emerge if longer study-test intervals were implemented (Mitchell et al., 2006; Schaefer et al., 2011; Sharot & Phelps, 2004; Sharot & Yonelinas, 2008; Yick et al., 2015), consistent with consolidation processes (Hamann, 2001; Talmi, 2013). However, previous studies also indicated that even in immediate conditions, negative arousing stimuli can be better recognized than neutral stimuli (e.g., Wirkner, Ventura-Bort, Schulz, Hamm, & Weymar, 2018). Whereas in Experiment 2 stimuli were controlled for arousal, in Experiment 1 both negative and positive stimuli had higher arousal ratings than neutral ones (according to Soares et al., 2012). As such, a difference could be at least expected between negative and neutral stimuli in the context of Experiment 1. Nevertheless, we only observed a trend for a

beneficial effect of negative information compared to positive information, which is in line with prior evidence showing that memory superiority effects are more consistent for negative information and less clear for positive information (e.g., Ochsner, 2000; Otani, Jaffa, et al., 2012; Otani, Libkuman, et al., 2012; Rossi-Arnaud, Spataro, Costanzi, Saraulli, & Cestari, 2018; Wang, 2018; Wang & Fu, 2011). In this context, differences in valence/arousal between the original ANEW ratings and the participants' ratings in the current study should be also considered. We cannot rule out the possibility that the findings of this experiment were somehow affected by a mismatch between the original ratings and the subjective evaluations of the words by the participants of the current study (see Davidson et al., 2006, Experiment 1 and 2A, and Koenig & Mecklinger, 2008 for examples).

Another plausible explanation for the lack of an EEM effect may relate to the experimental conditions as participants were aware that their memory would be tested later, and they were overtly instructed to memorize item and source. Thus, despite a processing advantage of emotional words over neutral words in terms of attention, distinctiveness, and organization (D'Argembeau & Van der Linden, 2004; Maddox et al., 2012; Talmi, 2013), the impact of these factors might have been attenuated by intentional learning. Hence, it is possible that valence effects were overshadowed by similar attentional resources and effortful encoding for all types of stimuli (e.g., Ferré et al., 2019, Experiment 1). When the relevance of the stimuli to the current goals is identical across emotional and neutral conditions, the memory performance may not differ as a function of valence and/or arousal (Ochsner, 2000), in line with the arousal-biased competition theory (Mather & Sutherland, 2011). Furthermore, experimental conditions relying on a recognition test instead of free recall (e.g., D'Argembeau & Van der Linden, 2004; Doerksen & Shimamura, 2001), the type of recognition test (e.g., direct source memory test instead of old/new judgments followed by a source memory test conditioned to old items), different study-test cycles with few trials (e.g., Adelman & Estes, 2013), and the short period between study and test (e.g., Mitchell et al., 2006; Sharot & Phelps, 2004; Sharot & Yonelinas, 2008; Wang, 2018), might have also contributed to the current pattern of findings.

Limitations

The manipulation of valence was critical in the current study, but the full control of stimulus arousal was not possible in the case of Experiment 1. Therefore, it is not possible to clearly dissociate valence from arousal effects in this case. Moreover, word selection relied on norms considering the combined ratings of male and female participants, yet the sample was mainly composed of female participants. Considering prior evidence for sex differences in the encoding and recall of emotional events (e.g., Galli,

Wolpe, & Otten, 2011; Glaser, Mendrek, Germain, Lakis, & Lavoie, 2012), this may limit the generalization of the current findings. Also, the role of specific stimulus properties (e.g., familiarity; concreteness; imageability; age of acquisition; self-descriptive/commonness characteristics**;** Adelman & Estes, 2013; Fan et al., 2016) was not accounted for in this study. Another limitation concerns the lack of control of factors such as attention, organization, distinctiveness, or personal motivations, which may impact upon the emotion-memory dynamics (Talmi, 2013), and are critical for a more thorough discussion of theoretical frameworks such as the object-based binding theory (Mather, 2007) and the arousal-biased competition theory (Mather & Sutherland, 2011). Likewise, it was not possible to clarify whether the results observed in the test phase depended on mechanisms that operated during encoding and/or recognition (see Hamann, 2001, and Levine & Edelstein, 2009 for reviews), even if the experimental manipulations were restricted to the encoding phase and even if the influence of consolidation processes is not expected in conditions of immediate recognition.

Conclusion

The current study confirmed the role of task-related factors in the interplay between emotion and memory. Specifically, emotional stimuli encoded in different conditions (read aloud *vs.* silently; self-reference *vs.* common) led to distinct findings in terms of internal source memory, item memory, and JOSs ratings. In line with the source-monitoring framework, we showed that emotional events do not always enhance episodic memory recognition, and that their impact is not the same for source and item memory (Johnson et al., 1993; Jurica & Shimamaura, 1999). Although the encoding strategies tested here - the production and the self-reference effect - are known to benefit item memory, a similar benefit was only found for internal source memory in the case of positive and neutral self-referenced words. Further research is needed to clarify the relationship between emotion and source memory, considering that stimulus type (e.g., Durbin et al., 2017), the type of source discrimination (e.g., Boywitt, 2015), and the encoding strategy (e.g., Kuhlmann & Touron, 2012, 2017) may affect this relationship. The same applies to JOSs ratings, which were sensitive to both valence and encoding strategy manipulations, although their predictive value was not confirmed. Together, the current experiments demonstrate how different encoding strategies modulate the effects of valence on distinct features of episodic memory, and on prospective metamemory judgments. Future studies should account for the way information is encoded when probing how emotion influences different facets of episodic memory.

References

Adelman, J. S., & Estes, Z. (2013). Emotion and memory: A recognition advantage for positive and negative words independent of arousal. *Cognition, 129*(3), 530-535. https://doi.org/10.1016/j.cognition.2013.08.014

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*(4), 548-564. https://doi.org/10.1037/0033-295X.97.4.548

Bowen, H. J., Kark, S. M., & Kensinger, E. A. (2018). NEVER forget: Negative emotional valence enhances recapitulation. *Psychonomic Bulletin and Review*, *25*(3), 870–891. https://doi.org/10.3758/s13423-017-1313-9

Boywitt, C. D. (2015). Non-monotonic relationships between emotional arousal and memory for color and location. *Cognition and Emotion*, *29*(8), 1335–1349. https://doi.org/10.1080/02699931.2014.977850

Carroll, M., Davis, R., & Conway, M. (2001). The effects of self-reference on recognition and source attribution. *Australian Journal of Psychology, 53*(3), 140-145. https://doi.org/10.1080/00049530108255136

Carroll, M., Mazzoni, G., Andrews, S., & Pocock, P. (1999). Monitoring the future: Object and source memory for real and imagined events. *Applied Cognitive Psychology, 13*(4), 373-390. https://doi.org/10.1002/(SICI)1099-0720(199908)13:4<373::AID-ACP605>3.0.CO;2-F

Castel, A. D., Rhodes, M. G., & Friedman, M. C. (2013). Predicting memory benefits in the production effect: The use and misuse of self-generated distinctive cues when making judgments of learning. *Memory & Cognition*, *41*(1), 28–35. https://doi.org/10.3758/s13421-012-0249-6

Coelho, R., Martins, A., & Barros, H. (2002). Clinical profiles relating gender and depressive symptoms among adolescents ascertained by the Beck Depressive Inventory II. *European Psychiatry, 17*(4), 222-226. https://doi.org/10.1016/S0924-9338(02)00663-6

Cook, G. I., Hicks, J. L., & Marsh, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition*, *35*(2), 222–230. https://doi.org/10.3758/BF03193443

D'Argembeau, A., Comblain, C., & Van der Linden, M. (2005). Affective valence and the self-reference effect: Influence of retrieval conditions. *British Journal of Psychology*, *96*(4), 457–466. https://doi.org/10.1348/000712605X53218

D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion, 4*(2), 173-188. https://doi.org/10.1037/1528-3542.4.2.173

D'Argembeau, A., & Van der Linden, M. (2005). Influence of emotion on memory for temporal information. *Emotion, 5*(4), 503-507. https://doi.org/10.1037/1528-3542.5.4.503

Davidson, P. S. R., McFarland, C. P., & Glisky, E. L. (2006). Effects of emotion on item and source memory in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 306–322. https://doi.org/10.3758/CABN.6.4.306

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, *8*(1), 155–161. https://doi.org/10.3758/BF03196152

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1*(1), 5–11. https://doi.org/10.1037/1528-3542.1.1.5

Dulas, M. R., & Duarte, A. (2014). Aging affects the interactio between attentional control and source memory: An fMRI study. *Journal of Cognitive Neuroscience*, *26*(12), 2653–2669. https://doi.org/10.1162/jocn_a_00663

Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory, 25*(9), 1191-1200. https://doi.org/10.1080/09658211.2017.1282517

Dutton, A., & Carroll, M. (2001). Eyewitness testimony: Effects of source. *Australian Journal of Psychology*, *53*(2), 83–91. https://doi.org/10.1080/00049530108255128

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*(3), 183-201. https://doi.org/10.1037/h0047707

El Haj, M., Kessels, R. P. C., & Allain, P. (2016). Source memory rehabilitation: A review toward recommendations for setting up a strategy training aimed at the "What, Where, and When" of episodic retrieval. *Applied Neuropsychology: Adult*, *23*(1), 53–60. https://doi.org/10.1080/23279095.2014.992071

Fan, W., Zhong, Y., Li, J., Yang, Z., Zhan, Y., Cal, R., & Fu, X. (2016). Negative emotion weakens the degree of self-reference effect: evidence from ERPs. *Frontiers in Psychology, 7*, 1408. https://doi.org/10.3389/fpsyg.2016.01408

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. https://doi.org/10.3758/BF03193146

Ferré, P., Comesaña, M., & Guasch, M. (2019). Emotional content and source memory for language: Impairment in an incidental encoding task. *Frontiers in Psychology, 10*, 65. https://doi.org/10.3389/fpsyg.2019.00065

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-development inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Galli, G., Wolpe, N., & Otten, L. J. (2011). Sex differences in the use of anticipatory brain activity to encode emotional events. *The Journal of Neuroscience, 31*(34), 12364-12370. https://doi.org/10.1523/JNEUROSCI.1619-11.2011

Glaser, E., Mendrek, A., Germain, M., Lakis, N., & Lavoi, M. E. (2012). Sex differences in memory of emotional images: A behavioral and electrophysiological investigation. *International Journal of Psychophysiology, 85*(1), 17-26. https://doi.org/10.1016/j.ijpsycho.2012.01.007

Hamami, A., Serbun, S. J., & Gutchess, A. H. (2011). Self-Referencing enhances memory specificity with age. *Psychology and Aging*, *26*(3), 636–646. https://doi.org/10.1038/jid.2014.371

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences, 5*(9), 394–400. https://doi.org/10.1016/S1364-6613(00)01707-1

Hourihan, K. L., & Bursey, E. (2017). A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory*, *25*(1), 35–43. https://doi.org/10.1080/09658211.2015.1122809

Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory & Cognition*, *45*(1), 121–136. https://doi.org/10.3758/s13421-016-0646-3

JASP Team (2018). JASP (Version 0.9.0.1) [Computer software].

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28. https://doi.org/10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67–85. https://doi.org/10.1037/0033-295X.88.1.67

Jönsson, F. U., Olsson, H., & Olsson, M. J. (2005). Odor emotionality affects the confidence in odor naming. *Chemical Senses, 30*(1), 29–35. https://doi.org/10.1093/chemse/bjh254

Jurica, P. J., & Shimamura, A. P. (1999). Monitoring item and source information: Evidence for a negative generation effect in source memory. *Memory & Cognition, 27*(4), 648-656. https://doi.org/10.3758/BF03211558

Kalenzaga, S., Sperduti, M., Anssens, A., Martinelli, P., Devauchelle, A., Gallarda, T., ...Piolino, P. (2015). Episodic memory and self-reference via semantic autobiographical memory: Insights from an fMRI study in younger and older adults. *Frontiers in Behavioral Neuroscience, 8*, 449. https://doi.org/10.3389/fnbeh.2014.00449

Kelly, A., Carroll, M., & Mazzoni, G. (2002). Metamemory and reality monitoring. *Applied Cognitive Psychology, 16*(4), 407-428. https://doi.org/10.1002/acp.803

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences, 101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., & Kark, S. M. (2018). Emotion and memory. In E. J. Wagenmakers & J. T. Wixted (Eds.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol.1, pp.1-26). New York: Wiley.

Kensinger, E. A., & Schacter, D. L. (2006a). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *Journal of Neuroscience, 26*(9), 2564-2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

Kensinger, E. A., & Schacter, D. L. (2006b). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition, 34*(2), 251–260. https://doi.org/10.3758/BF03193403

Kensinger, E. A., & Schacter, D. L. (2008). Memory and emotion. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.). *Handbook of emotions* (4th ed., pp. 564-578). New York: Guilford.

Koenig, S., & Mecklinger, A. (2008). Electrophysiological correlates of encoding and retrieving emotional events. *Emotion*, *8*(2), 162–173. https://doi.org/10.1037/1528-3542.8.2.162

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1352–1364. https://doi.org/10.1037/a0027863

Kuhlmann, B. G., & Touron, D. R. (2017). Relate it! Objective and subjective evaluation of mediator-based strategies for improving source memory in younger and older adults. *Cortex*, *91*, 25–39. https://doi.org/10.1016/j.cortex.2016.11.015

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self-other differences. *Journal of Personality and Social Psychology*, *37*(4), 499–514. https://doi.org/10.1037/0022-3514.37.4.499

Le Bigot, L., Knutsen, D., & Gil, S. (2018). I remember emotional content better, but I'm struggling to remember who said it. *Cognition, 180*, 52-58. https://doi.org/10.1016/j.cognition.2018.07.001

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150.

Leshikar, E. D., & Duarte, A. (2012). Medial prefrontal cortex supports memory accuracy for self-referenced items. *Social Neuroscience, 7*(2), 126–145. https://doi.org/10.1080/17470919.2011.585242

Leshikar, E. D., & Duarte, A. (2014). Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 236–252. https://doi.org/10.3758/s13415-013-0198-y

Levine, L. J., & Edelstein, R. S. (2009). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition and Emotion, 23*(5), 833-875. https://doi.org/10.1080/02699930902738863

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32*(3), 474–488. https://doi.org/10.3758/BF03195840

MacKenzie, G., Powell, T. F., & Donaldson, D. I. (2015). Positive emotion can protect against source memory impairment. *Cognition and Emotion*, *29*(2), 236–250. https://doi.org/10.1080/02699931.2014.911145

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 671–685. https://doi.org/10.1037/a0018785

Maddock, R. J., & Frein, S. T. (2009). Reduced memory for the spatial and temporal context of unpleasant words. *Cognition and Emotion*, *23*(1), 96–117. https://doi.org/10.1080/02699930801948977

Maddox, G. B., Naveh-Benjamin, M., Old, S., & Kilb, A. (2012). The role of attention in the associative binding of emotionally arousing words. *Psychonomic Bulletin & Review*, *19*(6), 1128–1134. https://doi.org/10.3758/s13423-012-0315-x

Mao, X., You, Y., Li, W., & Guo, C. (2015). Emotion impairs extrinsic source memory-An ERP study. *Biological Psychology*, *110*, 182–189. https://doi.org/10.1016/j.biopsycho.2015.07.005

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., & Nesmith, K. (2008). Arousal-enhanced location memory for pictures. *Journal of Memory and Language*, *58*(2), 449–464. https://doi.org/10.1016/j.jml.2007.01.004

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

Mazzoni, G., & Nelson, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1263–1274. https://doi.org/10.1037/0278-7393.21.5.1263

McKague, M., McAnally, K. I., Skovron, M., Bendall, S., & Jackson, H. J. (2012). Source monitoring and proneness to auditory-verbal hallucinations: A signal detection analysis. *Cognitive Neuropsychiatry*, *17*(6), 544–562. https://doi.org/10.1080/13546805.2012.676311

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*(5), 711–747. https://doi.org/10.1037/0033-2909.130.5.711

Mickley Steinmetz, K. R., Schmidt, K., Zucker, H. R., & Kensinger, E. A. (2012). The effect of emotional arousal and retention delay on subsequent-memory effects. *Cognitive Neuroscience, 3*(3-4), 150-159. https://doi.org/10.1080/17588928.2012.677421

Mitchell, K. J., Mather, M., Johnson, M. K., Raye, C. L., & Greene, E. R. (2006). A functional magnetic resonance imaging investigation of short-term source and item memory for negative pictures. *NeuroReport, 17*(14), 1543-1547. https://doi.org/10.1097/01.wnr.0000234743.50442.e5

Murphy, N. A., & Isaacowitz, D. M. (2008). Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging, 23*(2), 263-286. https://doi.org/10.1037/0882-7974.23.2.263

Nashiro, K., & Mather, M. (2011). How arousal affects younger and older adults' memory binding. *Experimental Aging Research*, *37*(1), 108–128. https://doi.org/10.1080/0361073X.2011.536746

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, *26*, 125–173. https://doi.org/10.1016/S0079-7421(08)60053-5

Newsome, R. N., Dulas, M. R., & Duarte, A. (2012). The effects of aging on emotion-induced modulations of source retrieval ERPs: Evidence for valence biases. *Neuropsychologia*, *50*(14), 3370–3384. https://doi.org/10.1016/j.neuropsychologia.2012.09.024

Nomi, J. S., Rhodes, M. G., & Cleary, A. M. (2013). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition and Emotion*, *27*(1), 141–149. https://doi.org/10.1080/02699931.2012.679917

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology General, 129*(2), 242-261. https://doi.org/10.1037//0096-3445.129.2.242

Ochsner, K. N., & Phelps, E. (2007). Emerging perspectives on emotion–cognition interactions. *Trends in Cognitive Sciences, 11*(8), 317–318. https://doi.org/10.1016/j.tics.2007.06.008

Okon-Singer, H., Hendler, T., Pessoa, L., & Shackman, A. J. (2015). The neurobiology of emotion-cognition interactions: Fundamental questions and strategies for future research. *Frontiers in Human Neuroscience, 9*, 58. https://doi.org/10.3389/fnhum.2015.00058

Otani, H., Jaffa, M., Libkuman, T. M., Goernert, P. N., & Kato, K. (2012). Does source memory impairment associated with emotionally arousing stimuli occur at encoding or retrieval? *Journal of Cognitive Psychology*, *24*(7), 796–801. https://doi.org/10.1080/20445911.2012.693468

Otani, H., Libkuman, T. M., Goernert, P. N., Kato, K., Migita, M., Freehafer, S. E., & Landow, M. P. (2012). Emotion, directed forgetting, and source memory. *British Journal of Psychology*, *103*(3), 343–358. https://doi.org/10.1111/j.2044-8295.2011.02078.x

Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1543–1547. https://doi.org/10.1037/a0020604

Ozubko, J. D., Major, J., & MacLeod, C. M. (2014). Remembered study mode: Support for the distinctiveness account of the production effect. *Memory*, *22*(5), 509–524. https://doi.org/10.1080/09658211.2013.800554

Phelps, E. A., LaBar, K. S., & Spencer, D. D. (1997). Memory for emotional words following unilateral temporal lobectomy. *Brain and Cognition, 35*(1), 85-109. https://doi.org/10.1006/brcg.1997.0929

Quinlan, C. K., & Taylor, T. L. (2013). Enhancing the production effect in memory. *Memory*, *21*(8), 904–915. https://doi.org/10.1080/09658211.2013.766754

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null hypothesis significance tests in psychiatry: A non-technical guide using JASP. *BMC Psychiatry, 18*, 178. https://doi.org/10.1186/s12888-018-1761-4

Raye, C. L., & Johnson, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society*, *15*(6), 405–408. https://doi.org/10.3758/BF03334572

Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion, 11*(3), 553-562. https://doi.org/10.1037/a0024246

Rimmele, U., Davachi, L., & Phelps, E. A. (2012). Memory for time and place contributes to enhanced confidence in memories for emotional events. *Emotion, 12*(4), 834-846. https://doi.org/10.1037/a0028003

Ritchey, M., Dolcos, F., & Cabeza, R. (2008). Role of amygdala connectivity in the persistence of emotional memories over time: An event-related fMRI investigation. *Cerebral Cortex, 18*(11), 2494-2504. https://doi.org/10.1093/cercor/bhm262

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Rossi-Arnaud, C., Spataro, P., Costanzi, M., Saraulli, D., & Cestari, V. (2018). Divided attention enhances recognition of emotional stimuli: Evidence from the attentional boost effect. *Memory, 26*(3), 42-52. https://doi.org/10.1080/09658211.2017.1319489

Schaefer, A., Pottage, C. L., & Rickart, A. J. (2011). Electrophysiological correlates of remembering emotional pictures. *NeuroImage, 54*(1), 714-724. https://doi.org/10.1016/j.neuroimage.2010.07.030

Schmidt, K., Patnaik, P., & Kensinger, E. A. (2011). Emotion's influence on memory for spatial and temporal context. *Cognition and Emotion, 25*(2), 229-243. https://doi.org/10.1080/02699931.2010.483123

Schümann, D., Bayer, J., Talmi, D., & Sommer, T. (2018). Dissociation of immediate and delayed effects of emotional arousal on episodic memory. *Neurobiology of Learning and Memory, 148*, 11-19. https://doi.org/10.1016/j.nlm.2017.12.007

Serbun, S. J., Shih, J. Y., & Gutchess, A. H. (2011). Memory for details with self-referencing. *Memory*, *19*(8), 1004–1014. https://doi.org/10.1080/09658211.2011.626429

Sharot, T., & Phelps, E. A. (2004). How arousal modulated memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience, 4*(3), 294-306. https://doi.org/10.3758/CABN.4.3.294

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Talmi, D. (2013). Enhanced emotional memory. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Tauber, S. K., & Dunlosky, J. (2012). Can older adults accurately judge their learning of emotional information? *Psychology and Aging*, *27*(4), 924–933. https://doi.org/10.1037/a0028447

Tu, H. W., Alty, E. E., & Diana, R. A. (2017). Event-related potentials during encoding: Comparing unitization to relational processing. *Brain Research*, *1667*, 46–54. https://doi.org/10.1016/j.brainres.2017.05.003

Tu, H. W., & Diana, R. A. (2016). Two are not better than one: Combining unitization and relational encoding strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(1), 114–126. https://doi.org/10.1037/xlm0000170

Wagenmakers, E-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ...Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*(1), 58-78. https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ...Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications, *Psychonomic Bulletin & Review, 25*(1), 35-57. https://doi.org/10.3758/s13423-017-1343-3

Wang, B. (2018). Retention interval modulates the effect of negative arousing pictures on recognition memory. *Memory, 26*(8), 1105-1116. https://doi.org/10.1080/09658211.2018.1427763

Wang, B., & Fu, X. (2011). Time course of effects of emotion on item memory and source memory for Chinese words. *Neurobiology of Learning and Memory*, *95*(4), 415–424. https://doi.org/10.1016/j.nlm.2011.02.001

Watson, L. A., Dritschel, B., Jentzsch, I., & Obonsawin, M. C. (2008). Changes in the relationship between self-reference and emotional valence as a function of dysphoria. *British Journal of Psychology*, *99*(1), 143–152. https://doi.org/10.1348/000712607X248689

Wegesin, D. J., Jacobs, D. M., Zubin, N. R., Ventura, P. R., & Stern, Y. (2000). Source memory and encoding strategy in normal aging. *Journal of Clinical and Experimental Neuropsychology*, *22*(4), 455–464. https://doi.org/10.1076/1380-3395(200008)22:4;1-0;FT455

Weymar, M., Löw, A., and Hamm, A. O. (2011). Emotional memories are resilient to time: evidence from the parietal ERP old/new effect. *Human Brain Mapping, 32*(4), 632-640. https://doi.org/10.1002/hbm.21051

Weymar, M., Löw, A., Melzig, C. A., & Hamm, A. O. (2009). Enhanced long-term recollection for emotional pictures: Evidence from high-density ERPs. *Psychophysiology, 46*(6), 1200-1207. https://doi.org/10.1111/j.1469-8986.2009.00869.x

Wirkner, J., Ventura-Bort, C., Schulz, P., Hamm, A. O., & Weymar, M. (2018). Event-related potentials of emotional and neutral memories: The role of encoding position and delayed testing. *Psychophysiology, 55*(5), e13069. https://doi.org/10.1111/psyp.13069

Yang, L., Truong, L., Fuss, S., & Bislimovic, S. (2012). The effects of ageing and divided attention on the self-reference effect in emotional memory: Spontaneous or effortful mnemonic benefits? *Memory, 20*(6), 596–607. https://doi.org/10.1080/09658211.2012.690040

Yick, Y. Y., Buratto, L. G., & Schaefer, A. (2015). The effects of negative emotion on encoding-related neural activity predicting item and source recognition. *Neuropsychologia*, *73*, 48–59. https://doi.org/10.1016/j.neuropsychologia.2015.04.030

Yin, Y., Ma, Y., Xu, X., & Yang, H. (2018). The effect of self-referencing on memory for different kinds of source information. *Memory, 27*(4), 519-527. https://doi.org/10.1080/09658211.2018.1532009

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: An emotional binding account. *Trends in Cognitive Sciences, 19*(5), 259–267. https://doi.org/10.1016/j.tics.2015.02.009

Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology, 65*(4), 236-244. http://dx.doi.org/10.1027/1618-3169/a000409

Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*, *62*(3), 240–253. https://doi.org/10.1016/j.jml.2009.11.004

Appendix

Summary of Findings Reported in Studies Probing the Effects of Stimulus Emotional Properties on Source Memory Performance, Considering External, Internal, and Reality Monitoring Tasks

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
External SM								
Doerksen & Shimamura, 2001	E1	24 (16f; 19.6)	128 words	32 Pos: V ≥ 5.25 32 Neg: V ≤ 2.5 64 Neu: n/a	Read word silently (intentional)	Stimuli color	Recognition test	↑ Pos, Neg
	E2	24 (17f; 19.9)	= E1	= E1	= E1	Color of the border	= E1	↑ Pos, Neg
Kensinger & Corkin, 2003	E2	18 (18m; 22.7)	280 words (ANEW)	140 Neg: set A (V = 1.78; A = 6.95); set B (V = 1.91; A = 7.16) 140 Neu: set A (V = 5.83; A = 3.12); set B (V = 5.65; A = 3.35)	Rate each word as abstract or concrete (intentional)	Stimuli color	Recognition test	↑ Neg
	E4	20 (20m; 21.5)	180 words	60 taboo: V = 3.58; A = 8.02 60 Neg: V = 1.78; A = 3.35 60 Neu: V = 5.22; A = 4.51	= E2	= E2	= E2	↑ Neg (↑ taboo)
D'Argembeau & Van der Linden, 2004	E1	32 (12f; 22.9)	96 words	24 Pos: V ≥ 4.10 24 Neg: V ≤ 1.90 48 Neu: 2.70 ≤ V ≥ 3.30	Read word (16 participants intentional; 16 incidental)	Stimuli color (two colors)	Recognition test	✕
	E2	32 (22f; 21.8)	64 words	16 Pos: n/a 16 Neg: n/a 32 Neu: n/a	Read word (intentional)	= E1	= E1	✕

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
D'Argembeau & Van der Linden, 2004	E3	48 (34f; 21.5)	= E1	= E1	= E1	= E1 (four colors)	= E1	↑ Pos, Neg (incidental) ✕ (intentional)
	E4	48 (34f; 22.5)	= E1	= E1	= E1	Spatial location	= E1	↑ Pos, Neg (both incidental and intentional)
D'Argembeau & Van der Linden, 2005		48 (24f; 23.7)	126 pictures (IAPS)	42 Neg: V = 2.73; A = 5.69 42 Neu: V = 4.98; A = 3.15 42 Pos: V = 7.40; A = 5.39	Look at pictures to recognize them later (intentional for the item, not for the source)	Temporal context (list 1 *vs.* list 2 *vs.* list 3)	Recognition test	↑ Neg (.38 < *d* < .58)
Davidson et al., 2006	E1	32 (n/a; 21.6) only YA	48 words (ANEW)	24 Emotional: n/a 24 Neu: n/a	Indicate weather each word was emotional or neutral (incidental)	Voice (male *vs.* female)	Recognition test (two-alternative forced choice)	↑ Emotional (if normative ratings) ✕ (if participants' ratings)
	E2A	32 (n/a; 18.5)	= E1	= E1	= E1	Stimuli color (two colors)	= E1	✕ (participants' ratings)
	E2B	33 (n/a; 18.3)	= E1	= E1	= E1 (intentional)	= E2A	= E2A	✕
Mitchell et al., 2006		19 (10f; 20.6)	360 pictures (266 IAPS + other sources)	180 Neg: V = 2.63; A = 6.15 180 Neu: V = 6.44; A = 3.67	Memorize picture and location (intentional)	Spatial location (six locations)	Recognition test	↓ Neg (short-term memory)
Koenig & Mecklinger, 2008		20 (10 f; 21)	480 pictures (IAPS)	160 Pos: V = 7.13; A = 5.04 160 Neg: V = 2.97; A = 5.20 160 Neu: V = 5.09; A = 3.43	Respond if indoors or outdoors (intentional)	Temporal context (list 1 *vs.* list 2) and spatial location (booth 1 *vs.* booth 2)	Recognition test	✕

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Maddock & Frein, 2009	E1A	48 (34f; 19.77)	36 words	12 Pos: V = 2.74; A = 6.98 12 Neg: V = 3.04; A = 7.13 12 Neu: V = 0.26; A = 2.57	Rate pleasantness (incidental)	Spatial location (6 x 7 cell spatial grid)	Free recall	↓ Neg
	E1B	48 (40f; 19.08)	36 words	12 Pos: V = 2.36; A = 6.67 12 Neg: V = 2.62; A = 6.78 12 Neu: V = 0.22; A = 2.62	= E1A	= E1A	= E1A	↓ Neg
	E3	48 (29f; 19.9)	= E1A	= E1A	= E1A	Temporal context (which word appeared first in word pairs)	= E1A	↓ Neg
	E4	24 (17f; 20.67)	72 words (E1A + E1B)	= E1A + E1B	= E1A (intentional)	= E1A (2 x 2 cell spatial grid)	Recognition test	↓ Neg
Rimmele et al., 2011	E1	25 (13f; 22.58)	120 scenes (IAPS)	60 Neg: V = 2.88; A = 5.62 60 Neu: V = 5.58; A = 3.87	Judge if the frame color appeared in the scene (incidental)	Color of the border	Recognition test	↓ Neg
	E2	26 (16f; 24.35)	= E1	= E1	= E1	Spatial location (four locations)	= E1	↑ Neg
	E3	32 (18f; 23.90)	= E1	= E1	= E1	Temporal context (first *vs.* second *vs.* third block)	= E1	↑ Neg

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Schmidt et al., 2011	E1	24 (14f; 19.7)	540 pictures (IAPS)	Pos: V = 6.98 Neg: V = 3.06 HA: A = 5.92 LA: A = 4.19 Neu: V = 5.08; A = 3.10	Living judgment (something alive) or common judgment (something that might be encounter within a typical month) or approach judgment (move closer in "real life") (incidental)	Temporal context (first *vs.* second *vs.* third list) Spatial location (left *vs.* center *vs.* right)	Recognition test	↑ HA (compared to LA) (η_p^2 = .17) ✕ (valence) (for R judgments)
	E2	24 (13f; 19.7)	n/a	n/a	Memorize objects and background scenes (intentional)	Order (first *vs.* second *vs.* third in the order of appearance) Spatial location (3 x 3 grid)	Recognition test	↑ HA (compared to LA) (η_p^2 = .24) ✕ (valence)
Wang & Fu, 2011		136 (86f; 22.91)	120 Chinese words	20 Pos: set 1 (V = 6.31; A = 6.02); set 2 (V = 6.42; A = 6.01) 20 Neg: set 1 (V = 2.44; A = 5.77); set 2 (V = 2.42; A = 5.79) 20 Neu: set 1 (V = 5.06; A = 4.97); set 2 (V = 5.05; A = 4.97)	Memorize words (intentional)	Stimuli color	Recognition test	✕ (for different delay conditions: immediate; 19 min; 63 min; 4.9 h; 8.75h; 1 day; 6 days; 14 days) (.004 < η_p^2 < .051)
Rimmele et al., 2012	E1	25 (13f; 22.52)	120 scenes (IAPS)	60 Neg: V = 2.88; A = 5.62 60 Neu: V = 5.58; A = 3.87	Judge if the color of dots appeared elsewhere in the scene	Color of dots presented in the scene	Recognition test	↓ Neg (*d* = .84)

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Boywitt, 2015	E1	101 (84%f; n/a)	120 pictures (IAPS)	LA: V = 5.28; A = 2.88 MA: V = 5.10; A = 4.82 HA: V = 4.55; A = 6.76	Memorize pictures (intentional)	Color of the border	Recognition test	✕ (K judgments) (η_p^2 = .028) ↓ Increasing arousal (R judgments) (η_p^2 = .26)
	E2	83 (77%; n/a)	= E1	= E1	= E1	Spatial location (left *vs.* right)	= E1	∩ (R and K) (η_p^2 = .139)
MacKenzie et al., 2015	E1	32 (23f; 19)	60 pictures (IAPS)	20 Pos: V = 7.92; A = 4.53 20 Neg: V = 2.74; A = 4.53 20 Neu: V = 5.27; A = 4.53	Memorize picture and border color (intentional)	Color of the border	Recognition test	↓ Pos, Neg
	E2	32 (20f; 21)	= E1	= E1	Memorize picture and color (intentional)	Stimuli color	= E1	↓ Neg
Yick et al., 2015		42 (28f; 22)	720 pictures (IAPS + Google image)	Emotional: V = 1.98; A = 3.21 Neu: V = 3.21; A = 1.95	Rate emotionality (intentional)	Temporal context (block 1 *vs.* block 2)	Recognition test	↑ Emotional pictures
Reality monitoring								
Kensinger & Schacter, 2006b	E1	32 (18f; 21.3)	128 words (ANEW)	64 Neg: V = 3.00; A = 6.2 64 Neu: V = 5.80; A = 3.9	Verify if the height of the first letter was lower than the height of the fourth letter (half of the participants incidental and other half intentional)	Heard/imagined *vs.* seen	Recognition test	↑ Neg (both incidental and intentional) (η_p^2 = .45)

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Kensinger & Schacter, 2006b	E2	20 (12f; 21.5)	450 words + 450 pictures of objects	HA: words (A = 2.96); pictures (A = 3.07) LA: words (A = 0.59); pictures (A = 0.67)	Indicate if each object was bigger or smaller than a shoebox (incidental)	Imagined *vs.* seen	Recognition test	↑ HA (η_p^2 = .33)
	E3A	16 (8f; 21.9)	= E1	= E1	= E1 (incidental)	= E1	= E1	↑ Neg (η_p^2 = .62)
	E3B	16 (n/a; 21.3)	= E2	= E2	= E2 (incidental)	= E2	= E2	↑ HA (η_p^2 = .80)
Cook et al., 2007	E1	34 (n/a)	270 words (ANEW)	90 Pos: V = 7.70; A = 5.71 90 Neg: V = 2.31; A = 5.89 90 Neu: V = 5.18; A = 5.70	Study the words (incidental)	Heard/imagined *vs.* seen	Recognition test	↓ Neg (only in comparison with Neu, not Pos)
	E2	120 (n/a)	= 90	= E1 (but only neg and neu words)	= E1	= E1	= E1	↓ Neg (for both mixed and pure lists)
	E3	106 (n/a)	= 90	= E1 (but only pos and neu)	= E1	= E1	= E1	↓ Pos (for both mixed and pure lists)
	E4	68 (n/a)	160 categorial words representing love, joy, fear, sadness, and anger (and also neutral condition)	128 emotional 32 neutral	Solving an anagram or rate the frequency of encounter in the last two weeks (incidental)	= E1	= E1	↓ Pos, Neg

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
McKague et al., 2012		133 (98f; 19.3); only 20 scoring lowest in the auditory-verbal hallucination questionnaire	192 words (ANEW)	96 Pos: V = 7.48; A = 5.44 96 Neg: V = 2.63; A = 5.60	Generate aloud or heard sentences including a target word	Self-generated *vs.* other-generated	Recognition test	↓ Neg (in comparison with Pos) (η_p^2 = .19)
Le Bigot et al., 2018	E1	42 (38f; 20.86)	104 words (ANEW)	52 Neg: V = 2.28; A = 5.20 52 Neu: V = 4.93; A = 3.36	Produce a sentence out loud containing a specific word or hear another participant producing a sentence (incidental)	Self-generated *vs.* partner-generated	Free recall	↓ Neg
	E2	46 (46f; 21.28)	= E1	52 Pos: V = 7.89; A = 6.36 Neu = E1	= E1	= E1	= E1	↓ Pos
Internal SM								
Kensinger & Schacter, 2006a		21 (11m; 18-35)	360 words 360 pictures	120 Pos: n/a 120 Neg; n/a 120 Neu: n/a	Animate judgment (something animate) or common judgment (something encountered in a typical month) (incidental)	Animate judgment *vs.* common judgment	Recognition test	✕
Cook et al., 2007	E4	68 (n/a)	160 categorial words representing love, joy, fear, sadness, and anger (and also neutral condition)	128 emotional 32 neutral	Solving an anagram or rate the frequency of encounter in the last two weeks (incidental)	Solving an anagram *vs.* rate the frequency of encounter in the last two weeks	Recognition test	↓ Pos, Neg

(continued)

Study	*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Sharot & Yonelinas, 2008	24 (n/a; 18-22)	360 photos (IAPS)	180 Neg: V = 7.69; A = 6.79 180 Neu: V = 3.75; A = 3.03	Rate visual complexity of the photo (incidental)	Colour task (colours of the photo to judge the complexity) *vs.* detail task (details of the photo to judge the complexity	Recognition test	✕ (for different delay conditions: 5 min; 24h)
Newsome et al., 2012	16 YA (8f; 25) 14 OA (5f; 67.7)	432 pictures (IAPS) + 450 pictures (Google)	Pos: YA (V = 1.69; A = 2.50); OA (V = 1.59; A = 2.78) Neg: YA (V = 4.25; A = 3.65); OA (V = 4.22; A = 4.04) Neu: YA (V = 2.74; A = 2.12); OA (V = 2.77; A = 2.50)	Common judgment (something that might be seen within a typical month) or indoors judgment (the image depicts something indoor) (incidental)	Common judgment *vs.* indoors judgment	Recognition test	↓ Neg
Otani, Jaffa, et al., 2012	40 (29f; n/a)	60 pictures (IAPS)	20 Pos: set A (V = 5.52; A = 4.07); set B (V = 6.31; A = 4.39) 20 Neg: set A (V = 2.10; A = 3.88); set B (V = 1.75; A = 4.65) 20 Neu: set A (V = 4.41; A = 2.34); set B (V = 4.27; A = 2.39)	Memorize the pictures and if apple- or orange-kind (intentional)	Apple-kind *vs.* orange-kind	Free recall	↓ Neg (more misattribution errors than Neu pictures, yet no difference in relation to Pos pictures) (η_p^2 = .08)

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Otani, Libkuman, et al., 2012		40 (26f; n/a)	60 pictures (IAPS)	20 Pos: set 1 (V = 7.11; A = 6.34); set 2 (V = 7.40; A = 6.26) 20 Neg: set 1 (V = 2.45; A = 6.02); set 2 (V = 2.17; A = 6.24) 20 Neu: set 1 (V = 4.99; A = 2.84); set 2 (V = 4.92; A = 2.75)	Remember pictures that were followed by a RRRR instruction slide (intentional)	To remember *vs.* to forget items	Free recall	↓ Neg (more misattribution errors than Neu pictures, yet no difference in relation to Pos pictures) (η_p^2 = .11)
Leshikar et al., 2015	E1	24 YA (14f; 21.25) 24 OA (14f; 66.83)	288 adjectives (ANEW)	144 Neg: n/a 144 Pos: n/a	Self-referential judgment (self-descriptive) or common judgment (commonly used word) (intentional)	Self-referential judgement *vs.* common judgment	Recognition test	↑ Pos (both age groups in the self-referential condition)
Mao et al., 2015		17 (10f; 23.4)	1080 pictures (IAPS + CAPS)	360 Pos: V = 2.12; A = 3.04 360 Neg: V = 3.90; A = 3.52 360 Neu: V = 2.84; A = 2.62	Respond if the item belongs to the scene or people category (n/a)	Scenes *vs.* people	Recognition test	↓ Neg, Pos (only for R judgments, not K judgments) (η_p^2 = .25)
Durbin et al., 2017	E1	24 (12f; 20.6)	432 words (ANEW)	144 Pos: V = 7.41; A = 5.68 144 Neg: V = 2.79; A = 5.70 144 Neu: V = 5.10; A = 3.84	Self-referential judgment (degree of personal relevance) or story judgment (likelihood of appearance in a current news story) (n/a)	Self-referential judgement *vs.* story judgment	Recognition test	↑ Pos (only for words encoded self-referentially compared to the story judgment) (η_p^2 = .16)

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Durbin et al., 2017	E2	24 (12f; 20.5)	504 pictures (IAPS)	168 Pos: V = 7.07; A = 5.28 168 Neg: V = 2.90; A = 5.34 168 Neu: V = 5.05; A = 3.36	Self-referential judgment (= E1) or quality judgment (perceptual quality of the picture)	Self-referential judgement *vs.* quality judgment	Recognition test	↑ Neg, Neu (for pictures encoded non-self-referentially than self-referentially) ($.17 < \eta_p^2 < .30$)
Zhang et al., 2018	E1	39 (27f; 20.2)	80 trait words (pool of Chinese personality-trait adjectives)	40 Pos: n/a 40 Neg: n/a	Self-referential judgment (indicate whether the word could be used to describe themselves) or other-referential judgment (indicate whether the word could be used to describe others) (incidental)	Self-referential judgment *vs.* other-referential judgment	Recognition test	↓ Neg (in comparison with Pos in the self-referential condition) ↑ Pos (for self-referent than other-referent words) (η^2 = .34)
	E2	70 (54f; 20.6)	160 trait words (pool of Chinese personality-trait adjectives)	80 Pos: n/a 80 Neg: n/a	= E1	= E1	= E1	↓ Neg (in comparison with Pos in the self-referential condition) ↓ Neg (for self-referential than other-referential condition)

(continued)

Study		*N* (m/f; M_{age})	Type of stimuli	Valence and/or arousal properties: *M*	Encoding task (intentional/incidental)	SM manipulation	Type of SM test	SM effect (effect size range)
Ferré et al., 2019	E1	30 (23f; 22.3)	120 words (emoFinder search engine; mostly ANEW)	60 Neg: study set (V = 2.24; A = 6.53); test set (V = 2.21; A = 6.56) 60 Neu: study set (V = 5.13; A = 4.13); test set (V = 5.11; A = 4.08)	Memorize the words and the language of appearance (intentional)	Spanish *vs.* Catalan	Recognition test	✕
	E2	30 (28f; 21.4)	= E1	= E1	Read aloud task (incidental)	= E1	Recognition test	↓ Neg ($.19 < \eta_p^2 < .33$)
	E3	30 (20f; 24.3)	= E1	60 Pos: study set (V = 7.46; A = 6.02); test set (V = 7.48; A = 6.17) Neu = E1	= E1	= E1	Recognition test	↓ Pos (η_p^2 - .26)

Note. A = Arousal; ANEW = Affective Norms for English Words; CAPS = Chinese Affective Picture Set; E = Experiment; HA = High Arousal; IAPS = International Affective Picture System; K = Know; LA = Low Arousal; MA = Medium Arousal; Neg = Negative; Neu = Neutral; n/a = Not Available; OA = Old Adults; Pos = Positive; R = Remember; SM = Source Memory; V = Valence; YA = Young Adults; ↑ = Enhancement Effect on Source Memory; ↓ = Impairment Effect on Source Memory; ✕ = No Difference Between Emotional and Non-emotional Stimuli on Source Memory; ∩ = Inverted "U" Relationship.

CHAPTER IV

IS INTERNAL SOURCE MEMORY RECOGNITION MODULATED BY EMOTIONAL ENCODING CONTEXTS?

Is internal source memory recognition modulated by emotional encoding contexts?[1,2]

Abstract

The influence of emotion on memory has been mainly examined by manipulating the emotional valence and/or arousal of critical items. Few studies probed how emotional information presented during the encoding of critical neutral items modulates memory recognition, particularly when considering source memory features. In this study, we specified the role of emotional encoding contexts in internal source memory performance (discrimination between encoding tasks) using a mixed (Experiment 1) and a blocked design (Experiment 2). During the study phase, participants were required to evaluate a set of neutral words, using either a self-referential or a semantic (common judgment) encoding strategy. Prior and concomitantly with each word, negative, neutral or positive pictures were presented in the background. The beneficial effect of self-referential encoding was observed for both item and internal source memory in both experiments. Remarkably, item and internal source memory recognition was not modulated by emotion, even though a secondary analysis indicated that the consistent exposure to negative (*vs.* positive) information led to worse source memory performance. These findings suggest that internal source memory of neutral items is not always affected by changing or repetitive emotional encoding contexts.

Keywords: internal source memory; valence; emotion; emotional interference; recognition memory; self-referential processing.

Emotional information appears to have a privileged status in cognitive processing relative to non-emotional information: we are faster at detecting, identifying and directing our attentional resources to emotional stimuli (Calvo & Lang, 2004; Carretié, Hinojosa, Martín-Loeches, Mercado, & Tapia, 2004; Meinhardt & Pekrun, 2003; Pratto & John, 1991; see also Hamann, 2001). Such advantage has been also demonstrated in the context of episodic memory. The emotion-enhanced memory (EEM) effect indicates that emotional stimuli are remembered better than neutral ones (Dolcos & Cabeza, 2002; Kensinger, Gutchess, & Schacter, 2007; Marchewka al., 2016; Talmi, Anderson, Riggs, Caplan, & Moscovitch, 2008; see Murphy & Isaacowitz, 2008 for a meta-analytic review). Whereas most studies addressed the impact of emotional encoding conditions on memory for neutral target items (item

[1] Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2020). Is internal source memory recognition modulated by emotional encoding contexts? *Psychological Research*. Advance online publication. https://doi.org/10.1007/s00426-020-01294-4

[2] The Supplementary Materials associated with this chapter are presented in Appendix C.

memory), other aspects of episodic memory such as memory binding and source memory have received less attention. Specifically, source memory represents the memory for qualitative features of a study episode, which might include perceptual, contextual (e.g., spatiotemporal features), cognitive (e.g., elaboration processes), and affective (e.g., feelings elicited by the situation) details, being also permeable to the effects of previous experiences (e.g., stereotypes; beliefs; Johnson, Hastroudi, & Lindsay, 1993; Johnson & Raye, 1981).

Studies probing the effects of emotional stimuli on source memory have yielded mixed findings (see Appendix table from D. R. Pereira, Sampaio, & Pinheiro, 2019 for a selective overview). Emotional stimuli have been associated with improved source memory compared to neutral ones (e.g., Doerksen & Shimamura, 2001; Kensinger & Corkin, 2003). Notwithstanding, other studies reported impaired source memory for emotional stimuli (e.g., Cook, Hicks, & Marsh, 2007; Ferré, Comesaña, & Guasch, 2019, Experiment 2 and 3; K. J. Mitchell, Mather, Johnson, Raye, & Green, 2006; Newsome, Dulas, & Duarte, 2012), whereas others reported null effects (e.g., Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006; Koenig & Mecklinger, 2008; D. R. Pereira et al., 2019, Experiment 1).

The study of the factors that may account for these discrepant results (e.g., retention interval; type of stimuli; intentionality of encoding; encoding strategy) may shed light on the relationship between memory and emotion (Murphy & Isaacowitz, 2008). In this context, the way emotion is manipulated during encoding should also be considered (e.g., Han. Mao, Kartvelishvili, Li, & Guo, 2018). Specifically, the EEM effect is usually observed by manipulating stimulus valence (the hedonic value of the stimuli, ranging from unpleasant/negative to pleasant/positive) and/or arousal (the degree of activation elicited by the stimuli, ranging from a relaxed/calm to an excited/aroused state; Lang, Bradley, & Cuthbert, 1990; Lang, Greenwald, Bradley, & Hamm, 1993). In this case, emotional salience represents an intrinsic feature of the target stimuli. Nonetheless, emotional cues might be unrelated and extrinsic to the target stimuli, such as when emotional information is concomitantly presented (an emotional context/background - e.g., face; picture; sound) during the encoding of neutral items (e.g., Cui et al., 2016; A. P. R. Smith, Henson, Dolan, & Rugg, 2004; Takashima, van der Ven, Kroes, & Fernández, 2016). For example, the processing of neutral items in emotional vs. neutral contexts has been associated with differences in electrophysiological brain responses, reflected in larger amplitudes of the Late Positive Complex (LPC) in emotional conditions during encoding (Ventura-Bort, Löw, Wendt, Dolcos, et al., 2016), and by enhanced parietal old-new differences during retrieval (Ventura-Bort et al., 2017; Ventura-Bort, Löw, Wendt, Moltó, et al., 2016).

Even though some studies support an EEM effect for neutral stimuli studied in emotional contexts relative to neutral contexts (Takashima et al., 2016; Ventura-Bort, Löw, Wendt, Moltó, et al., 2016), others report a memory advantage only in the case of positive contexts (A. P. R. Smith et al., 2004; A. P. R. Smith, Henson, Rugg, & Dolan, 2005). Null effects were also documented (Jaeger, Johnson, Coronna, & Rugg, 2009; Jaeger & Rugg, 2012). One could hypothesize that the presentation of an emotional context interferes with the encoding of critical items since emotional information receives prioritized processing and requires additional resources that otherwise could be exclusively devoted to the target stimuli (Lv, Wang, Tu, Zheng, & Qiu, 2011; Meinhardt & Pekrun, 2003; M. G. Pereira et al., 2006). Such interference has been shown when emotional target stimuli, particularly negative/arousing items, disrupted the associations between the targets and their backgrounds (e.g., Bisby & Burgess, 2013, Experiment 1; Mather, Gorlick, & Nesmith, 2009, Experiment 3) and between target items as observed in some associative memory studies (e.g., Bisby & Burgess, 2013, Experiment 3; Bisby, Horner, Hørlyck, & Burgess, 2016; Han et al., 2018; see also Bisby & Burgess, 2017 for a review). Nonetheless, when considering neutral targets presented in emotional contexts, emotional encoding conditions were found to elicit beneficial or null effects on memory performance (e.g., Jaeger et al., 2009; Jaeger & Rugg, 2012 Takashima et al., 2016; Ventura-Bort, Löw, Wendt, Moltó, et al., 2016).

The studies reviewed so far indicate that an emotional context may influence how a neutral critical stimulus is encoded and remembered, specifically when both stimulus and context are deemed relevant during encoding. However, concurrent emotional information is not always relevant for the task, and sometimes it is necessary to control for possible interference effects elicited by distracting emotional information (Dolcos, Iordan, & Dolcos, 2011; Iordan, Dolcos, & Dolcos, 2013). In this context, it is important to clarify in which conditions the processing of contextual emotional cues facilitates or hampers task performance, such as when one needs to code and subsequently remember certain types of information. This may also have practical implications, namely to the development of strategies targeting the reappraisal of emotional distractors. These strategies could be valuable, for example, in mood disorders characterized by a processing bias towards negative information (e.g., Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Koster, De Raedt, Goeleven, Franck, & Crombez, 2005; Mogg, Bradley, Williams, & Mathews, 1993; L. A. Watson, Dritschel Jentzsch, & Obonsawin, 2008). Furthermore, in conditions wherein the presence of emotion is beneficial, this knowledge can also be used to develop strategies aiming to boost the encoding and retrieval of neutral information in daily life.

Typically, emotional interference paradigms involve the presentation of an emotional stimulus (e.g., face; odor; picture; sound) prior, after and/or together with a critical neutral item (Erk et al., 2003; Guo, Li, Zhang, Cui, & Wei, 2018; Macri, Pavard, & Versace, 2018; Pierguidi et al., 2016; Ventura-Bort et al., 2017). Contrary to the studies discussed before, here the emotional context is task-irrelevant since participants are not required to associate context and item. When considering memory for neutral target stimuli, results are mixed and seem to vary according to the type of context (e.g., face vs. scenes: Bowen & Kensinger, 2017). Some studies support an EEM for both positive and negative information (Bridge, Chiao, & Paller, 2010; Brierley, Medford, Shaw, & David, 2007; Graf & Yu, 2015), whereas others indicate selective benefits of either positive (Erk et al., 2003; Martínez-Galindo & Cansino, 2015, 2017; Pierguidi et al., 2016) or negative contexts (Bowen & Kensinger, 2017). However, some studies do not find such emotion-related enhancement (Baeken et al., 2012; Bowen & Kensinger, 2017; Fenker, Schott, Richardson-Klavehn, Heinze, & Düzel, 2005; Macri et al., 2018), and some even report a deleterious effect, particularly when high-arousing stimuli are used (Guo et al., 2018; Zhang, Liu, An, Yang, & Wang, 2015). Despite the putative role played by the type of encoding contexts and their arousal properties in the explanation of the mixed findings (e.g., Bowen & Kensinger, 2017; Guo et al., 2018; Zhang et al., 2015), it is also critical to examine the mechanisms underpinning the effects of emotional irrelevant information.

Even when emotional stimuli are used as distractors, they could still easily attract attentional resources due to their evolutionary and motivational value (Bradley et al., 2003; Lang et al., 1990; Lang, Bradley, & Cuthbert, 1997). Consequently, the resources available for concurrent tasks could be reduced, resulting in longer response times and/or increased errors (e.g., Blair et al., 2007; Erthal et al., 2005; Lv et al., 2011; Meinhardt & Pekrun, 2003; D. Mitchell et al., 2007; M. G. Pereira et al., 2006). These deleterious effects appear particularly problematic when both emotional distractors and ongoing task are presented in a short temporal window or in contiguity. Notwithstanding, when this competition is minimized, the processing of emotional information may have a positive effect on the concurrent task (Bocanegra & Zeelenberg, 2009). The presentation of emotional stimuli may also lead to interference effects by inducing transient mood changes (e.g., M. G. Pereira et al., 2006; Sakaki, Gorlick, & Mather, 2011). Furthermore, the presentation of negative pictures before tasks requiring semantic processing was found to result in longer response times but not before perceptual tasks (Sakaki et al., 2011). This suggests that negative irrelevant information may only interfere with tasks that strongly rely on semantic processing (see also Kensinger, 2009). This hypothesis fits well with evidence showing that memory for

negative items is associated with more internal details such as thoughts and feelings than positive items, whereas there is no difference with respect to visual and temporal features (Mickley & Kensinger, 2009).

The evidence presented so far emerged from item memory studies. Few studies have addressed how emotional encoding conditions impact upon other features of source memory. In this context, the current study examined how emotional (negative/positive) and neutral encoding contexts modulate internal source memory, that is the ability to discriminate between internally derived processes (e.g., "did I judge this item as self-descriptive or as commonly used?"; Johnson et al., 1993). Internally derived memories usually include information concerning the cognitive operations (e.g., encoding strategy; imagery; organization of the information; autobiographical elaborations; verbal thoughts) taking place during encoding (Ferguson, Hashtroudi, & Johnson, 1992). Thus, this provides an ideal opportunity to probe how irrelevant emotional information affects the elaborative process during encoding, and consequently impacts upon source memory.

As far as we know, no previous study has examined emotional interference effects on internal source monitoring. Notwithstanding, more recent studies (Macri et al., 2018; Ventura-Bort et al., 2017; Xie & Zhang, 2017) offer some indirect evidence for these effects. Macri and collaborators (2018) presented negative and neutral odors during the learning phase. In the test phase, the source memory task consisted of remembering the locations of each item: source memory performance was improved for items presented in a negative relative to a neutral context. Ventura-Bort and colleagues (2017) used neutral objects paired with emotional and non-emotional scenes during encoding and tested the memory for spatial locations of the objects one week after the encoding session. Their results revealed no emotion effects on source memory. Finally, Xie and Zhang (2017) presented negative, positive and neutral picture contexts and asked participants to memorize a neutral object including its color and orientation. During test, participants needed to reproduce the color and the orientation of each object: source memory for both color and orientation benefited from negative encoding conditions compared to neutral and positive conditions (Xie & Zhang, 2017). Overall, this evidence suggests that negative emotional encoding contexts facilitate the distinction between different externally derived sources (i.e., external source memory; Johnson et al., 1993). Specifically, the benefit associated with negative contexts (Macri et al., 2018; Xie & Zhang, 2017) suggests that the presence of emotional information can enhance the ongoing processing of the main tasks (Bocanegra & Zeelenberg, 2009). These findings are also consistent with the idea that negative information interferes less with the processing of perceptual features (Sakaki et al., 2011), such as location or color of critical items, when compared to semantic processing.

As internal source memory requires discriminating memories of different cognitive operations, the impact of negative information on internal source memory may also differ from external source memory, which is based on more perceptual, temporal, and spatial features (Ferguson et al., 1992). Thereby, the effects of emotion on source memory may vary as a function of the type of detail that is manipulated (Schmidt, Patnaik, & Kensinger, 2011). To the best of our knowledge, this is the first study to explore how emotional information might affect the cognitive operations taking place during encoding and subsequent internal source memory decisions. Two possible outcomes were considered. First, negative information could have a deleterious effect on internal source memory. This hypothesis is grounded in the study of Sakaki and colleagues (2011), which revealed that negative information might interfere with semantic processing, and hence affect encoding and subsequent recognition. Second, a null effect could be expected if we consider the chance-level results reported by Ventura-Bort and colleagues (2017). As a matter of fact, the experimental design adopted in the current study had more in common with the former study than with the studies of Macri and colleagues (2018) and Xie and Zhang (2017). Nonetheless, since we adopted an immediate study-test interval and intentional encoding conditions, source memory performance is likely to be above chance-level. To test these predictions, neutral and emotional (negative and positive) pictures were presented together with target items in a self-reference task or a common judgment task (Experiment 1). Both emotional/neutral pictures and semantic tasks were intermixed during the study phase. Response times during encoding and recognition were also considered given that they provide additional evidence about which conditions modulate performance.

Considering that the time of exposure to item and background in memory studies is usually long, transient affect changes could account for the effects observed in emotional interference studies (e.g., M. G. Pereira et al., 2006; Sakaki et al., 2011; Xie & Zhang, 2017). Thereby, this possibility was explored in a second experiment wherein the neutral vs. emotional (positive and negative) encoding conditions were presented in a blocked design. Such design is commonly used in studies probing the effects of the sustained exposure to emotional stimuli, such as pictures (e.g., Figueira et al., 2017; Güntekin & Tülay, 2014; Pinheiro, Barros, Dias, & Niznikiewicz, 2017; J. C. Smith, Bradley, & Lang, 2005; Subramaniam et al., 2016). Although few studies explored how affect changes modulate source memory performance, reality monitoring studies (testing our ability to distinguish internally- from externally-derived information; Johnson et al., 1993) revealed that participants in a positive affective state were more prone to source memory misattributions when compared to participants in a negative affective state (Gingerich & Dodson, 2013). Positive affect was also associated with more accurate source memory performance when compared to a neutral condition (Subramaniam et al., 2016; Subramaniam, Ranasinghe, Mathalon,

Nagarajan, & Vinogradov, 2017). Specifically, source memory performance in a reality monitoring task (self-generated vs. externally presented information) was found to be improved in a positive (vs. neutral) condition, particularly in the case of self-generated information (Subramaniam et al., 2016). No difference in source memory performance was found between the negative and the neutral conditions. Such results stand in contrast with studies probing external source memory (Macri et al., 2018; Ventura-Bort et al., 2017; Xie & Zhang, 2017). Thereby, Experiment 2 aimed to specify how affect impacts upon internal source memory performance.

Experiment 1

In this experiment, the influence of emotional encoding contexts on internal source memory was examined by presenting emotional (negative and positive) and neutral pictures together with to-be-remembered neutral words. The study phase was followed by an old-new recognition test in which participants were instructed to discriminate between two internally derived sources associated with the neutral words during encoding. To enhance interference effects during encoding, we manipulated the valence of the distractor stimuli while keeping their arousal constant for two main reasons. First, differences in the neurofunctional mechanisms underlying effects of valence vs. arousal on memory have been documented (amygdala-hippocampal interactions in the case of arousal, and prefrontal cortex–hippocampal interactions in the case of valence; Kensinger, 2004; Kensinger & Corkin, 2003, 2004). Hence, the control of stimulus arousal properties is required to establish that any observed effects are due to valence. Second, previous studies indicated that whereas arousal effects are associated with a biological advantage of high-arousing stimuli, valence effects seem to rely on additional deliberate encoding resources through semantic and autobiographic elaborations (Cook et al., 2007; Kensinger, 2004; Kensinger & Corkin, 2004). In the latter case, by manipulating valence, interference effects may emerge from a conflict between the elaborative processes elicited by the emotional information and the elaborative processes required by the ongoing task, particularly in the case of negative information, which appears to elicit more thoughts and feelings than positive information (Mickley & Kensinger, 2009).

Also, to enhance control over the elaborative processes associated with each item during encoding, we used two tasks requiring a semantic and thus a deep analysis of the word following the levels of processing framework (Craik & Lockhart, 1972; Craik & Tulving, 1975): a common judgment task in which participants evaluated if the word was frequently used by people in their everyday life; a self-reference task in which participants stated if the word described or was related to them. The self-referential processing is a well-known approach to boost memory performance as items encoded self-

referentially are better remembered than items encoded in other deep and perceptual processing tasks (e.g., Culcea & Freitas, 2017; Leshikar, Dulas, & Duarte, 2015; Leshikar & Duarte, 2012; Rogers, Kuiper, & Kirker, 1977; Symons & Johnson, 1997). The contrast between the self-referential task and the common judgment task provides an opportunity to explore how emotional encoding contexts, particularly negative ones (Sakaki et al., 2011), interfere with tasks putatively involving distinct levels of processing. Although the common judgment task may be considered a deep semantic processing task, the self-referential processing has been proposed to enhance organization and elaboration (Symons & Johnson, 1997). Accordingly, Fan and colleagues (2016) reported that, relative to a neutral condition, the self-reference effect was weakened when participants were exposed to negative information, which was reflected in electrophysiological changes: reduced amplitudes of the P200, N200 and P300 event-related potential components for words processed self-referentially and preceded by negative pictures when compared to words processed self-referentially and preceded by neutral pictures. Although negative information is expected to interfere with both self-referential and common encoding tasks (Sakaki et al., 2011), it is not clear how it modulates internal source memory. As such, during the test phase of Experiment 1, the internal source memory task involved recognizing whether a certain item was studied in the self-referential condition, in the common condition, or whether it was a new item.

Methods

Participants.

The initial sample included 32 college students, who received course credit for their participation in the experiment. The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996; Coelho, Martins, & Barros, 2002; Martins, Coelho, Ramos, & Barros, 2000), and the State-Trait Anxiety Inventory - Form Y (STAI-Y; Santos & Silva, 1997; Silva & Campos, 1998; Silva & Spielberger, 2007) were used to rule out the presence of mood disorders. This control was necessary considering evidence showing that the processing of mood-congruent information is favored in relation to less congruent information, reflected in a processing bias towards negative information in depression or dysphoria (e.g., Koster et al., 2005; L. A. Watson et al., 2008), or a bias towards negative and threatening information in anxiety disorders (e.g., Bar-Haim et al., 2007; Mogg et al., 1993). Accordingly, data from two participants were discarded, one due to self-reported depression diagnosis, and another due to moderate depressive symptoms (score equal or above 20; Kendall, Hollon, Beck, Hammen, & Ingram, 1987). The final sample was composed of 30 participants (27 females; 28 right-handed), with ages between 18 and 36 years (M = 21.00, SD = 4.15), with an average of 13.63 years of education (SD = 2.40). These participants reported no

past/current medical or psychiatric conditions that could affect task performance. They also reported normal or corrected-to-normal vision, hearing, and motor abilities. Written informed consent was obtained from all participants, and all the procedures were approved by the local Ethics Committee of the University of Minho.

As no previous study explored how the presentation of emotional information during encoding of neutral items influence internal source memory, the study of Macri and colleagues (2018) and a previous experiment with a similar source memory task (D. R. Pereira et al., 2019) were used as references to estimate sample size based on reported values of partial eta squared (η_p^2). The η_p^2 values ranged between 0.08 and 0.56 in the case of source memory performance, using samples between 24 and 32 participants. Hence, using G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007), a sample of at least 14 participants was indicated as necessary to achieve an effect size of 0.08 (calculated as in SPSS and using a 3 [valence: negative/neutral/positive] x 2 [source: self/common] repeated-measures ANOVA as the statistical test of reference), with a power parameter of 0.80, and an alpha significance level of 0.05.

Materials.

Stimuli were 180 neutral adjectives and 90 pictures differing in emotional valence (see Supplementary Table S1). Adjectives were selected from the European Portuguese (EP) version of the Affective Norms for English Words (ANEW; Bradley & Lang, 1999; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), and also from a pilot study aiming to extend the ANEW database. A brief description of the procedure and main results from the pilot study is presented as Supplementary Material. The valence ratings of the adjectives ranged between 4.00 and 6.11 (*M* = 5.03, *SD* = 0.61), with medium arousal ratings (*M* = 5.10, *SD* = 0.81; see Supplementary Table S1). The 180 adjectives were randomly distributed across six lists of 30 items, which were used in both study and test phases. No statistically significant differences were found between lists regarding valence ($p = .84$), arousal ($p = .92$), frequency per million ($p = .42$), number of letters ($p = .81$), and number of syllables ($p = .98$). Additionally, 14 adjectives were selected to be used as training and filler items.

Pictures were selected from the EP adaptation of the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008; Soares et al., 2015). Specifically, using the IAPS valence ratings as reference, 30 negative, 30 neutral, and 30 positive pictures were selected[3]. As expected, statistically

[3] The pictures selected from the EP version of the IAPS were: negative – 1271, 1274, 1280, 2120, 2276, 2694, 2700, 2722, 6010, 6200, 6800, 7135, 8231, 9001, 9008, 9031, 9160, 9171, 9186, 9290, 9330, 9390, 9440, 9468, 9480, 9582, 9584, 9592, 9830, 9832; neutral – 1303, 1321, 1505, 1560,

significant differences were found between negative, neutral, and positive lists regarding valence (positive > neutral; positive > negative; neutral > negative; all $p < .001$), but no significant differences regarding arousal ratings were observed ($p = .22$), so arousal was held constant at medium levels. Furthermore, a similar count of human [$X^2(2, 90) = 1.16$, $p = .56$], animal [$X^2(2, 90) = 0.58$, $p = .75$], common objects [$X^2(2, 90) = 3.75$, $p = .15$], landscape [$X^2(2, 90) = 1.19$, $p = .38$], food [$X^2(2, 90) = 2.17$, $p = .34$], and animate [$X^2(2, 90) = 1.35$, $p = .51$] elements was ensured between valence categories. The 30 images of each valence category were divided into three lists including 10 negative, 10 neutral, and 10 positive images each. Each list served as a study block. Additional 14 images were used as fillers and training items. Different picture-word compounds were arranged to ensure that each picture was presented with different words across participants.

Procedure.

At the beginning of the experimental task, participants were told they were about to perform a memory task, in which words and pictures would be shown. For some words, participants were asked to evaluate if the word described or related to themselves (self-referential condition) by responding "yes" or "no". For other words, they were asked to evaluate if those words were frequently used by people in day-by-day settings (common [non-self-referential] condition; see Leshikar et al., 2015), and were required to respond "yes" or "no". A picture was presented concomitantly with each word. Participants were instructed not only to look at the pictures as they would be asked to rate them later (see Pinheiro et al., 2013; Pinheiro et al., 2015), but also to focus on the words as they would have to recognize them and their source (self-referential or common) among new words (intentional learning conditions). Participants were instructed to respond as fast as possible in all phases of the experiment, and to be as accurate as possible particularly during the test phase. They were also informed that they were going to perform three independent study-test cycles. Each study phase was composed of 30 word-picture trials, 10 with negative, 10 with neutral, and 10 with positive pictures presented in a random order. In half of the trials, participants had to make self-referential judgments, while in the other half they had to perform common judgments. The participants were also required to keep their right and left index fingers on the "v" and "n" keys to respond "yes" or "no", respectively. The "yes"/"no" responses were counterbalanced across participants. In each test phase, 30 old adjectives presented during the study were randomly mixed with

1726, 2309, 2312, 2590, 2605, 3360, 4233, 5920, 5950, 5961, 6910, 7360, 7450, 7497, 7632, 7640, 7920, 8060, 8116, 8117, 8160, 8192, 8220, 8325, 9070, 9422; positive – 1540, 1710, 1720, 1722, 1920, 2070, 2080, 2155, 2170, 2360, 4601, 5300, 5621, 5623, 5626, 7340, 7350, 7430, 7501, 7660, 8031, 8163, 8185, 8186, 8193, 8370, 8420, 8496, 8500, 8501.

30 new adjectives, and participants had to perform old-new recognition judgments. Also, every time the participants judged an item as "old", they were asked if they remembered the emotional content of the image that appeared together with the item. The aim of this question was twofold: first, to check whether participants remembered the emotional content of the pictures of a specific valence category better; second, to evaluate whether participants made additional efforts to associate word and picture, even if they were not instructed to do so and asked to prioritize the words and their source. During the test phase, participants were required to keep their right and left index and middle fingers on the "c", "v", "n", and "m" keys to provide a response. The response mapping of these keys was also counterbalanced across participants. Prior to the experimental trials, participants performed a brief training to get familiarized with the instructions and procedures.

Specifically, each study trial started with a black fixation cross ("+") displayed on the center of the screen in a light gray background, for 500 ms. A negative, neutral or positive picture appeared first alone for 1000 ms, and then together with a superimposed neutral adjective for 3000 ms, on the center of the screen. Hence, each image was displayed on the screen for a total of 4000 ms. We decided to present pictures before and together with the target items, since we were interested in the effects of distractor stimuli on encoding and not on consolidation processes (e.g., Anderson, Wais, & Gabrieli, 2006). Together with both picture and word, the instruction "Self-description" or "Common" was presented on top of the screen to inform the participants about the judgment they should perform in each trial. Two small cues "Y" and "N" were also added to the bottom of the screen to remember the participants when they should respond, and to minimize the cognitive load of remembering the response mapping (see Figure 1). Participants had to provide a response within the 3000 ms after word onset. If time elapsed and no response was registered, the next trial was automatically presented to ensure all participants had the same exposure time to both picture and word. The first two and the last two trials of each study phase were filler items to mitigate possible primacy and recency effects, and they were not considered in the statistical analysis.

During the test phase, each trial started with a black fixation cross (500 ms), followed by an old or new word that remained on the center of the screen until a response was made (self-paced). For each word, participants had to choose one of four options: "self-description"; "common"; "evaluated, but do not know if self-description/common"; "new". The "do not know option" was included to minimize guessing and response bias following previous studies (e.g., Dulas & Duarte, 2014; Leshikar & Duarte, 2012, 2014; Newsome et al., 2012; Sharot & Yonelinas, 2008). In a similar vein, if the participant judged an item as "old" by responding "self-description", "common" or "evaluated, but do not know if self-

description/common", an additional screen was displayed, and participants were required to report if they recalled the emotional content of the image paired with the previously judged word. Another four options were available in this case: "negative"; "neutral"; "positive"; "do not know". The mapping between options and response keys was always shown on the bottom of the screen (see Figure 1).

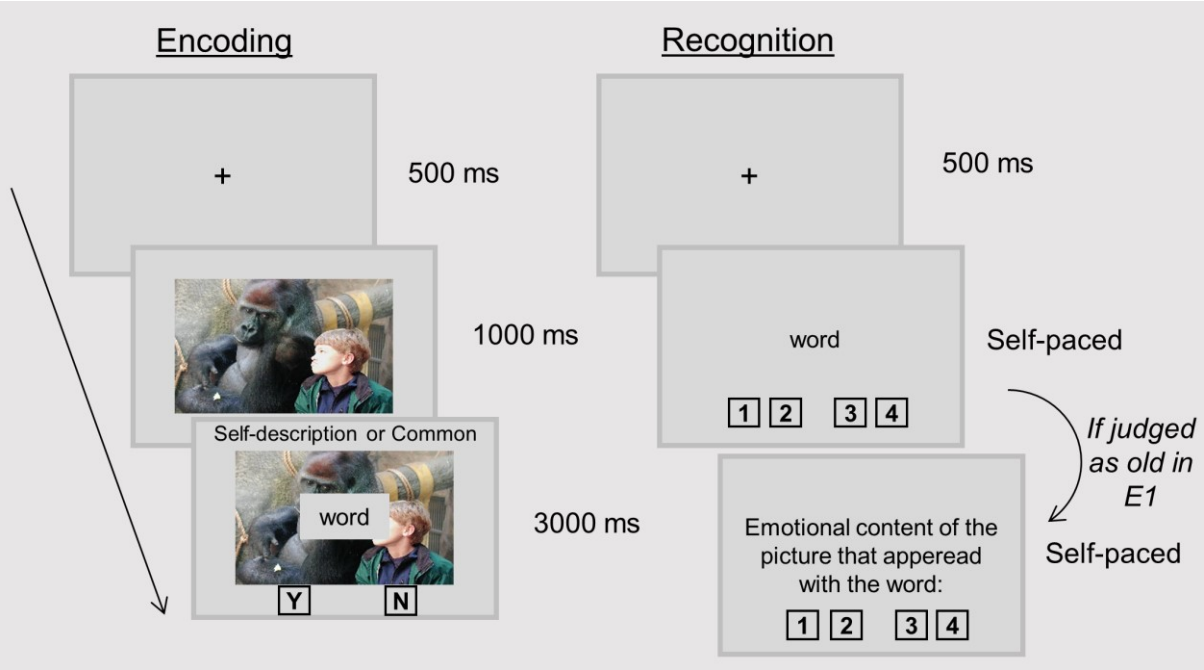


Figure 1. Schematic illustration of an experimental trial from encoding and retrieval phases. Whereas in Experiment 1 each block included mixed trials with negative, neutral and positive pictures, in Experiment 2 each block was composed by pictures of the same valence category. In the encoding phase, the position of "Yes" (Y) and "No" (N) keys was counterbalanced across participants. In the retrieval phase, keys "1" to "4" could be mapped into "self-description", "common", "evaluated, but do not know if self-description/common" or "new" for both experiments, and into "negative", "neutral", "positive", "do not know" in the case of Experiment 1.

At the end of the session, as a manipulation check, participants were also required to rate the valence of the pictures shown during the experimental task, using the 9-point Self-Assessment Manikin (SAM) scale (Bradley & Lang, 1999; 1 = the most unpleasant/negative, 9 = the most pleasant/positive). Superlab software 5.0 (Cedrus Corporation, 2015) was used to program and control stimulus presentation.

Data analysis.

Item recognition accuracy was based on the Snodgrass and Corwin (1988) recognition measure: Pr = ([p(hits) – p(false alarms)]). Of note, the p(hits) were obtained by merging correct/incorrect source responses, and "do not know" responses. The response bias was also computed: Br = ([p(false

alarms)/(1 – Pr)]). Values below 0.50 are indicative of a conservative response criterion, whereas values above 0.50 suggests a liberal response criterion (Snodgrass & Corwin, 1988). Following previous studies (e.g., Dulas & Duarte, 2014; Leshikar et al., 2015; Newsome et al., 2012), source memory recognition accuracy was based on the difference between correct source responses and incorrect source responses, [p(correct) – p (incorrect)]. After computing these measures, a 3 (valence: negative *vs*. neutral *vs*. positive) x 2 (source: self *vs*. common) repeated-measures analysis of variance (ANOVA) was run. Of note, a Friedman's ANOVA was used in the case of response bias as data were non-normally distributed. To support the ANOVA results, a Bayesian analysis with JASP (JASP Team, 2018, Version 0.9.0.1; Wagenmakers, Love et al., 2018; Wagenmakers, Marsman et al., 2018; see Supplementary Table S8-S9 and Figure S1) and a multinomial model approach were implemented, as described elsewhere (D. R. Pereira et al., 2019; see Supplementary Material section and Supplementary Tables S2-S4).

A similar 3 (valence: negative *vs*. neutral *vs*. positive) x 2 (source: self *vs*. common) repeated-measures ANOVA was used to examine response times obtained during the encoding phase, in the specific case of correct source responses. As a manipulation check, a one-way ANOVA was conducted with the picture valence ratings provided by the participants at the end of the session, in order to confirm if their valence ratings match those originally extracted from the EP version of the IAPS database (Soares et al., 2015).

In the Supplementary Material section, together with the results from the Bayesian and multinomial analysis, we present additional findings regarding the one sample *t*-tests run on item and source recognition measures to ensure that old-new and self-common source discriminations were not at chance level. Additional analyses are also presented regarding the proportion of incorrect source responses, "do not know" responses, misses, correct rejections, false alarms, "yes" responses provided during encoding, and a repeated-measures 2 (congruency: "yes" *vs*. "no" response in the encoding) x 3 (valence) x 2 (source) ANOVA including the factor congruency.

Statistical analyses were performed with the IBM Statistical Package for the Social Sciences Statistics, version 24 (IBM SPSS Statistics for Windows, Armonk, NY). The Greenhouse-Geisser approach was applied as correction when violations of sphericity were observed, and Bonferroni-corrected post-hoc tests were used to explore significant differences between conditions.

Results

Table 1 presents a summary of the main descriptive statistics regarding the behavioral performance in Experiment 1.

Recognition accuracy.

Item memory. The 3 x 2 repeated-measures ANOVA yielded a main effect of source, $F(1, 29) = 28.25$, $p < .001$, $\eta_p^2 = .49$, but no effect of valence, $F(2, 58) = 0.66$, $p = .521$, $\eta_p^2 = .02$, nor an interaction effect between source and valence factors, $F(2, 58) = 1.33$, $p = .272$, $\eta_p^2 = .04$. Pairwise comparisons supported a beneficial effect of self-referential encoding on item memory: words encoded in a self-referential manner ($M = 0.72$, $SE = 0.02$) were better recognized than words encoded in the common condition ($M = 0.60$, $SE = 0.02$; $p < .001$, $d = 0.97$, 95% CI [0.07, 0.15]; see Table 1 and Figure 2a). Regarding response bias, the 3 x 2 Friedman's ANOVA revealed a statistically significant effect, $X^2(5) = 27.89$, $p < .001$. Specifically, for words presented together with positive pictures, the Br values were higher in the case of self-referential condition in comparison with the common condition ($p = .002$; significant for $p < .006$ applying the Bonferroni correction for multiple comparisons). Following the Snodgrass and Corwin (1988) interpretation of the Br value, participants were less conservative in the self-referential condition than in the common condition (see Table 1 and Supplementary Table S6). Nonetheless, on average, Br values were below 0.50, suggesting an overall conservative response criterion.

Source memory. In the case of source memory, a main effect of valence, $F(2, 58) = 4.30$, $p = .018$, $\eta_p^2 = .13$, and a main effect of source, $F(1, 29) = 77.47$, $p < .001$, $\eta_p^2 = .73$, were observed, but no interaction effect emerged, $F(2, 58) = 0.68$, $p = .511$, $\eta_p^2 = .02$. Bonferroni-corrected post-hoc tests revealed that the source (self-referential/common) of words studied with neutral pictures as background was better recognized ($M = 0.42$, $SE = 0.04$) when compared to words studied in a negative background ($M = 0.36$, $SE = 0.03$; $p = .009$, $d = 0.59$, 95% CI [0.01, 0.11]). Even though Bayesian analysis supported a main effect of source, i.e., an advantage in source memory recognition for words in the self-referential condition ($M = 0.52$, $SE = 0.04$) relative to the non-self-referential condition ($M = 0.26$, $SE = 0.04$; $p < .001$, $d = 1.61$, 95% CI [0.20, 0.32]; see Table 1 and Figure 2c), it did not support the valence effect nor the difference between words studied in negative *vs.* neutral contexts (see Supplementary Material section).

Given the importance of the valence effect for the current study and considering that participants provided valence ratings for each picture, the measure of source memory accuracy was recomputed for each participant based on these ratings instead of the EP IAPS ratings. Based on prior studies (e.g., Kensinger & Corkin, 2004; Ritchey, LaBar, & Cabeza, 2011), pictures receiving valence ratings between 1 and 3 were considered negative, between 4 and 6 were considered neutral, and between 7 and 9 were considered positive. Then a 3 (valence) x 2 (source) repeated-measures ANOVA was run on source

memory accuracy. When stimulus ratings are organized in a bidimensional valence x arousal affective space, ratings typically fit a boomerang- or U-shape distribution (Lang et al., 1997; Lang, Bradley, & Cuthbert, 1998), revealing that more extreme valence scores are also associated with more extreme arousal scores. Thus, by selecting trials containing pictures rated with more extreme valence scores, it is possible to examine the influence of more arousing emotional pictures on source memory. In this context, a second 3 x 2 repeated-measures ANOVA was also run using the most extreme valence values, i.e., 1 and 2 for negative and 8 and 9 for positive stimuli, while keeping the neutral category within the same range.

Contrary to the initial analysis following the EP IAPS norms (Soares et al., 2015), the analyses using the participants' ratings only yielded a main effect of source (F(1, 28) = 54.77, p < .001, η_p^2 = .66 for the first analysis; F(1, 24) = 25.73, p < .001, η_p^2 = .52 for the second analysis), confirming the memory advantage for words studied self-referentially (M = 0.50, SE = 0.04; M = 0.54, SE = 0.05) in contrast to words studied in the common condition (M = 0.26, SE = 0.04; p < .001, d = 1.37, 95% CI [0.18, 0.32] for the first analysis; M = 0.28, SE = 0.05; p < .001, d = 1.02, 95% CI [0.16, 0.37] for the second analysis; see Supplementary Table S12). Thus, in line with the Bayesian results, no support was found in favor of a valence effect in the current analyses.

Response time.

Study phase. The mean response times inform about possible interference effects and elaborative processes taking place during the encoding phase. The statistical analysis yielded a main effect of source, F(1, 29) = 17.99, p < .001, η_p^2 = .38 [valence: F(2, 58) = 1.25, p = .294, η_p^2 = .04; interaction: F(2, 58) = 0.073, p = .929, η_p^2 = .003], showing that participants responded faster to words in the common condition (M = 1934, SE = 45.01) than to words in the self-referential condition (M = 2041, SE = 44.19; p < .001, d = 0.774, 95% CI [55.34, 158.42]; see Table 1 and Figure 3a).

Table 1

Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Source x Valence) Considering Both Experiments

	Experiment 1 – *M* (*SD*)						Experiment 2 – *M* (*SD*)					
	Self			Common			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
Correct source	0.57 (0.20)	0.60 (0.22)	0.60 (0.19)	0.36 (0.17)	0.43 (0.19)	0.37 (0.21)	0.59 (0.17)	0.62 (0.18)	0.61 (0.18)	0.35 (0.19)	0.37 (0.19)	0.40 (0.18)
Incorrect source	0.08 (0.07)	0.06 (0.06)	0.06 (0.08)	0.12 (0.12)	0.12 (0.09)	0.13 (0.09)	0.06 (0.09)	0.07 (0.11)	0.09 (0.12)	0.13 (0.10)	0.12 (0.14)	0.10 (0.11)
Source DNK	0.14 (0.11)	0.14 (0.13)	0.14 (0.13)	0.20 (0.17)	0.16 (0.13)	0.16 (0.16)	0.20 (0.15)	0.16 (0.15)	0.14 (0.12)	0.27 (0.19)	0.27 (0.16)	0.21 (0.16)
Source miss	0.21 (0.14)	0.20 (0.15)	0.19 (0.14)	0.31 (0.17)	0.29 (0.15)	0.34 (0.17)	0.15 (0.14)	0.14 (0.11)	0.15 (0.11)	0.25 (0.16)	0.24 (0.17)	0.29 (0.15)
Item Pr measure	0.70 (0.15)	0.71 (0.17)	0.73 (0.14)	0.60 (0.15)	0.63 (0.14)	0.58 (0.17)	0.72 (0.16)	0.74 (0.13)	0.72 (0.16)	0.63 (0.17)	0.64 (0.16)	0.59 (0.18)
Item Br measure	0.30 (0.26)	0.35 (0.31)	0.37 (0.31)	0.25 (0.25)	0.25 (0.24)	0.23 (0.23)	0.45 (0.34)	0.45 (0.33)	0.41 (0.33)	0.33 (0.25)	0.36 (0.29)	0.29 (0.24)
SM measure	0.48 (0.24)	0.54 (0.24)	0.54 (0.21)	0.23 (0.20)	0.31 (0.22)	0.24 (0.24)	0.53 (0.20)	0.55 (0.22)	0.53 (0.26)	0.23 (0.21)	0.26 (0.23)	0.30 (0.20)
RT – Study phase	2058 (267)	2038 (279)	2027 (259)	1961 (252)	1927 (279)	1915 (255)	2005 (273)	1914 (248)	1974 (248)	1937 (264)	1822 (233)	1862 (262)
RT – Correct source	2518 (794)	2522 (671)	2494 (697)	2987 (805)	3014 (858)	3088 (1122)	1911 (645)	1971 (557)	1986 (598)	2480 (1148)	2379 (974)	2499 (960)

	New	New negative	New neutral	New positive
CR	0.92 (0.08)	0.87 (0.11)	0.88 (0.09)	0.88 (0.11)
FA	0.08 (0.08)	0.13 (0.11)	0.12 (0.09)	0.12 (0.11)
RT – CR	1546 (520)	1409 (459)	1391 (437)	1442 (445)
RT – FA	3102 (1496)	2119 (794)	2759 (1763)	2807 (1522)

Note. CR = Correct Rejections; DNK = Do Not Know; FA = False Alarms; RT = Response Time; SM = Source Memory.

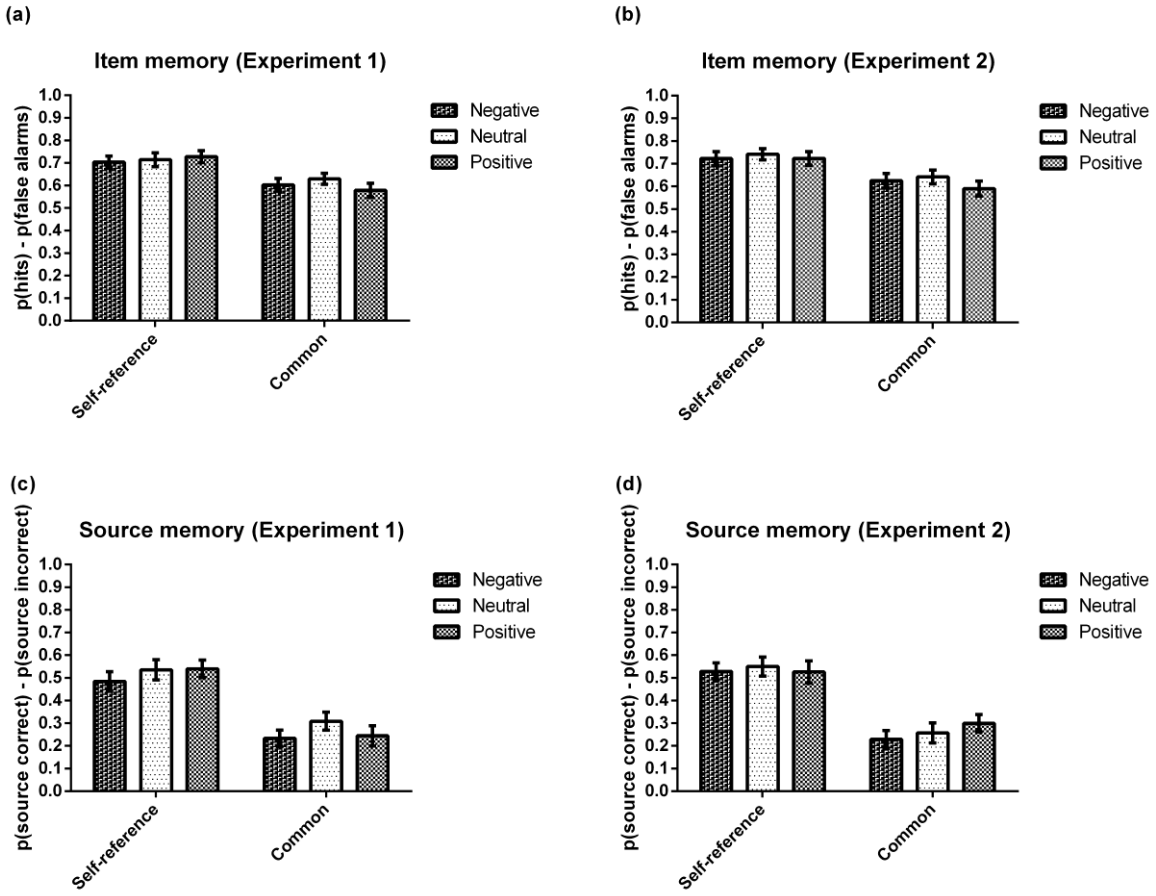


Figure 2. The recognition scores obtained for item (panel a – Experiment 1; panel b – Experiment 2) and source memory (panel c – Experiment 1; panel d – Experiment 2) are plotted on the y-axis as a function of source (self-reference/common) and valence of the background picture (negative/neutral/positive) presented during encoding. The error bars indicate the standard errors of the means.

Correct source responses. Similarly to the study phase results, the repeated-measures ANOVA only revealed a main effect of source $F(1, 26) = 22.43$, $p < .001$, $\eta_p^2 = .46$ [valence: $F(2, 52) = 0.14$, $p = .872$, $\eta_p^2 = .005$; interaction: $F(2, 58) = 0.73$, $p = .487$, $\eta_p^2 = .03$], but the difference was in the opposite direction: whereas responses during the study phase were slower in the self-referential condition, the RTs associated with accurate source responses were faster in the self-referential condition ($M = 2536$, $SE = 117.18$) than in the common condition ($M = 3034$, $SE = 167.84$; $p < .001$, $d = 0.91$, 95% CI [281.49, 713.28]; see Table 1 and Figure 3c).

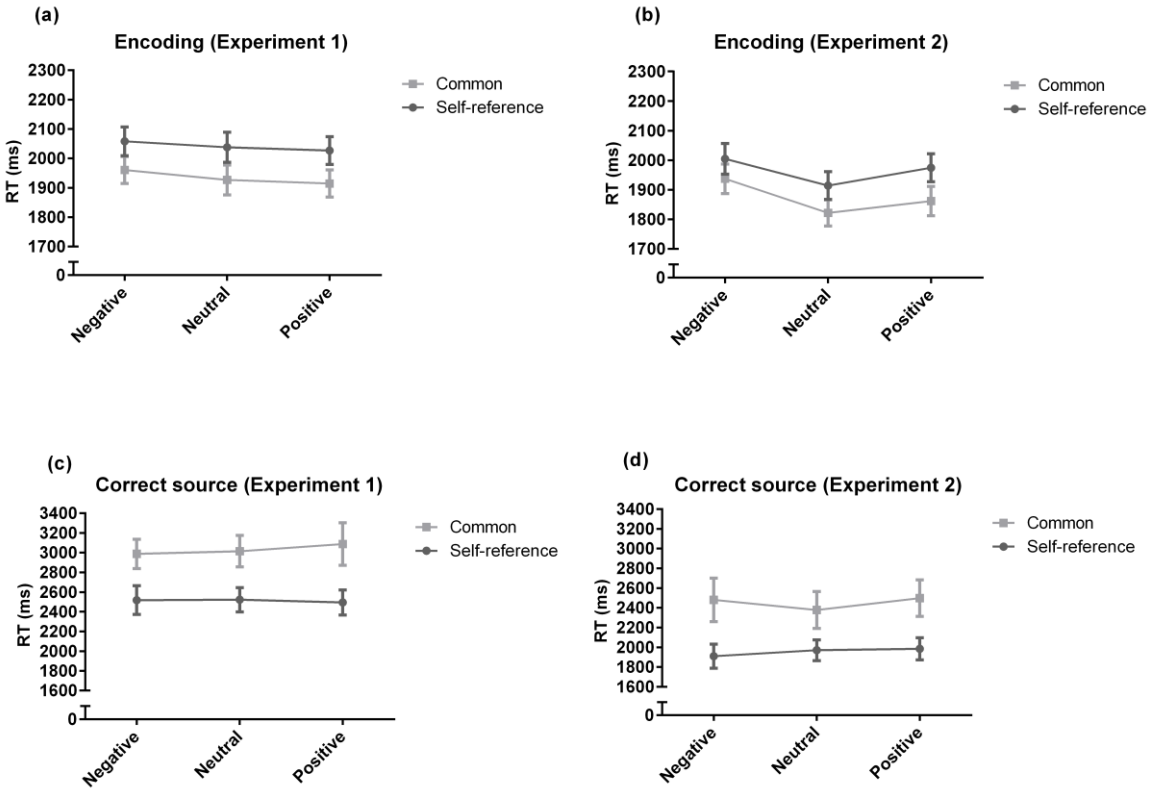


Figure 3. The response times obtained during the study phase (panel a – Experiment 1; panel b – Experiment 2) and during the recognition phase for source hits are plotted on the y-axis as a function of source (self-reference/common) and valence of the background picture (negative/neutral/positive) shown during encoding. The error bars indicate the standard errors of the means.

Manipulation check.

The valence ratings of the pictures presented during the experimental task were comparable to the ratings of the EP version of the IAPS (Soares et al., 2015; see Supplementary Table S1). Specifically, statistically significant differences were observed between the mean valence ratings of the negative, neutral and positive pictures presented in the current experiment, $F(2, 87) = 303.95$, $p < .001$, $\eta_p^2 = .88$. As expected, positive pictures received the highest valence ratings ($M = 7.00$, $SD = 0.61$), followed by neutral pictures ($M = 4.67$, $SD = 0.77$), with negative pictures receiving the lowest valence ratings ($M = 2.91$, $SD = 0.52$; $p < .001$ and $d > 2.66$ for all pairwise comparisons; positive *vs.* neutral: 95% CI [1.36, 2.17]; positive *vs.* negative: 95% CI [3.69, 4.50]; neutral *vs.* negative: 95% CI [1.93, 2.74]).

Discussion: Experiment 1

In the current experiment, the initial participant-based analysis (repeated-measures ANOVA) suggested that the internal source memory performance of neutral words encoded in a negative context

was impaired when compared to words encoded in a neutral context. No statistically significant differences emerged when comparing positive encoding contexts with both negative and neutral encoding contexts. As source memory scores were computed based on incorrect source responses, the former finding was not due to significant differences in the proportion of incorrect source responses across valence conditions (see Supplementary Material section). However, additional participant-based analyses based on individual valence ratings and both Bayesian and multinomial models provided weak evidence in favor of such valence effect (see Supplementary Material section). Previous studies exploring the influence of emotional interference during encoding in external source memory reported a positive effect of negative information (Macri et al., 2018; Xie & Zhang, 2017) or null effects accompanied by a floor effect (Ventura-Bort et al., 2017). Thus, in the case of internal source memory, the current experiment did not find any memory advantage for stimuli encoded in emotional contexts, but rather supported the lack of an effect. In the case of item memory, even though the item memory scores were indirectly derived from correct/incorrect source responses, and "do not know" responses, no EEM effect on recognition performance was observed (e.g., Baeken et al., 2012; Bowen & Kensinger, 2017; Fenker et al., 2005).

Considering that the external source memory features examined in previous studies were perceptual and spatial in nature (see Macri et al., 2018; Ventura-Bort et al., 2017; Xie & Zhang, 2017), and that internal source memory features are mainly characterized by cognitive operations (Johnson & Raye, 1981; Raye & Johnson, 1980), one of our hypothesis was that negative information would result in the worst internal source memory performance by interfering with semantic processing (Sakaki et al., 2011). Additionally, given previous evidence showing that the presence of negative stimuli can attenuate the degree of self-reference, at least at the electrophysiological level (Fan et al., 2016), we were also interested in exploring if the putative interference would differ across encoding tasks. Contrary to our hypothesis, the current experiment showed that the presentation of negative pictures during word encoding did not significantly affect neither immediate memory performance, nor the way participants behaviorally responded to the common condition and, particularly, to the self-referential condition (see Supplementary Material section).

Indeed, during encoding there was no evidence supporting a significant interference effect in the case of negative contexts when compared to both positive and neutral contexts. This claim is supported by our results regarding response time analysis as no significant negative valence modulations were observed during encoding. Slowed response times in main tasks have been reported for negative valence distractors and interpreted as an index of competition for attentional resources (Blair et al., 2007; Erthal

et al., 2005; D. Mitchell et al., 2007; Nielen et al., 2009). Although the absence of a significant response time modulation is not unprecedented (e.g., Erk, Martin, & Walter, 2005), it could also be the case that the stimuli were not presented in enough temporal proximity to cause a deleterious competition for attentional resources (Bocanegra & Zeelenberg, 2009). Thus, considering the experimental conditions used in Experiment 1, another mechanism by which negative information might interfere with encoding is by inducing transient negative mood states (M. G. Pereira et al., 2006; Sakaki et al., 2011; Xie & Zhang, 2017), which have been associated with deleterious effects on memory performance (e.g., Ellis, Thomas, & Rodriguez, 1984; Gärtner & Bajbouj, 2014; Zlomuzica, Preusser, Totzeck, Dere, & Margraf, 2016). In this context, one possible way to enhance the interference effect is to repeatedly present negative information to produce such mood changes. This possibility was tested in Experiment 2.

Experiment 2

The repetitive exposure to emotional pictures of the same valence category has been shown to induce affective changes (Figueira et al., 2017; Pinheiro et al., 2017; J. C. Smith, Bradley et al., 2005; Subramaniam et al., 2016), to result in sustained interference effects on performance in the main task (M. G. Pereira et al., 2006), and even to modulate spontaneous brain oscillatory activity (Güntekin & Tülay, 2014). Thus, in this experiment, the emotional and neutral pictures of Experiment 1 were presented in three valence-related blocks: one negative, one positive, and one neutral. By presenting emotional and neutral trials in a blocked design, we aimed to increase the interference effect of negative information (see M. G. Pereira et al., 2006) and to probe changes in the mood state of participants. As such, after each study phase, participants were asked to report their subjective negative and positive affect. An internal source memory impairment was hypothesized for neutral words encoded in negative contexts through transient mood changes. Specifically, based on previous studies (e.g., Figueira et al., 2017; Pinheiro et al., 2017), we predicted an increase in subjective negative affect and a decrease in positive affect after the presentation of unpleasant pictures, but no significant affect changes after both positive and neutral blocks.

Methods

Participants.

The same inclusion/exclusion and ethical procedures implemented in Experiment 1 were adopted in Experiment 2. Even though the initial sample was composed of 33 young adults, three participants were excluded due to self-report of moderate to severe depressive symptomology as measured by the

BDI-II (Beck et al., 1996; Coelho et al., 2002; Kendall et al., 1987; Martins et al., 2000), one was excluded due to medical diagnosis of depression, and another was excluded due to current pharmacological treatment with anticonvulsant drugs. The final sample was composed of 28 young adults (26 females; 23 right-handed), aged between 18 and 32 years (M = 21.63, SD = 3.96), and with 14.61 average years of formal education (SD = 3.22).

Materials.

The same pool of adjectives and pictures described in Experiment 1 were used in Experiment 2. Besides the BDI-II (Beck et al., 1996; Coelho et al., 2002; Martins et al., 2000), and the STAI-Y (Santos & Silva, 1997; Silva & Campos, 1998; Silva & Spielberger, 2007), we also administered the EP version of the Positive and Negative Affect Schedule (PANAS; Galinha & Pais-Ribeiro, 2005; D. Watson, Clark, & Tellegen, 1988) four times across the experimental session. The PANAS is a self-report questionnaire that measures current positive and negative affect by asking the participants to rate, on a 5-point Likert-type scale (1 = "not at all"; 5 = "extremely"), a group of 20 adjectives (e.g., negative: distressed, nervous, guilty; positive: interested; inspired; attentive). Participants were required to assess their current mood state.

Procedure.

The procedure followed Experiment 1 (see Figure 1). Nonetheless, some changes were introduced. Specifically, before providing the instructions to perform the experimental task, participants were asked to complete the PANAS questionnaire, and they were told that the questionnaire would appear more three times during the experimental task. The first time served as a baseline measure, and then PANAS was administered immediately after the encoding phase of each study-test block.

Contrary to Experiment 1, in which negative, neutral and positive pictures were mixed within a block, each study phase in Experiment 2 included only pictures of the same valence. Of note, the presentation of negative, neutral and positive blocks was counterbalanced and randomized across participants. Additionally, whereas in the test phases of Experiment 1 participants were asked if they could recall the emotional content of the pictures that appeared together with words recognized as "old", in Experiment 2 this question was not included. Thus, each test trial was only composed of a fixation cross (500 ms) followed by an old or new item (self-paced).

Data analysis.

The same data analysis procedure planned for Experiment 1 was followed in Experiment 2. The only exception was the use of a 3 (valence: negative *vs*. neutral *vs*. positive) x 2 (source: self *vs*. common) Friedman's ANOVA, instead of a repeated-measures ANOVA, for mean response time of source hits. This option was adopted as some of the data were non-normally distributed, following the Shapiro-Wilk test and the absolute values of skewness and kurtosis. Wilcoxon tests with corrections were used as follow-up comparisons in the case of statistically significant results. Concerning the PANAS mood scores, the difference between self-reported affect before and after each pictures block was computed, following previous studies (e.g., Figueira et al., 2017; Kohn et al., 2014). Second, a repeated-measures ANOVA was run with the difference scores for the positive affect scale, and the non-parametric equivalent, the Friedman's ANOVA, in the case of the negative affect scale as the data were non-normally distributed. In the latter case, Wilcoxon tests were planned for statistically significant results.

Results

Table 1 displays the main descriptive statistics of the behavioral performance obtained in Experiment 2.

Recognition accuracy.

Item memory. As in Experiment 1, we replicated the self-referential encoding advantage in word recognition (M = 0.73, SE = 0.02) when compared to the common condition (M = 0.62, SE = 0.03; p < .001), regardless of valence, $F(1, 27) = 29.43$, $p < .001$, $\eta_p^2 = .52$, $d = 1.03$, 95% CI [0.07, 0.15]; see Table 1 and Figure 2b). No other significant effects were observed [valence: $F(2, 54) = 0.92$, $p = .404$, $\eta_p^2 = .03$; interaction: $F(2, 54) = 0.73$, $p = .489$, $\eta_p^2 = .03$]. The analysis of the response bias indicated a main effect of source, $F(1, 25) = 11.22$, $p = .003$, $\eta_p^2 = .31$, and no other significant effects were found [valence: $F(2, 50) = 1.12$, $p = .336$, $\eta_p^2 = .04$; interaction: $F(2, 50) = 0.19$, $p = .827$, $\eta_p^2 = .01$]. Specifically, participants were more conservative in the recognition of words encoded in the common condition (M = 0.32, SE = 0.05) than words encoded in the self-referential condition (M = 0.44, SE = 0.06; p = .003, d = 0.66, 95% CI [0.04, 0.18]; see Table 1). Once more, on average, participants showed a Br value below 0.50, which is indicative of a conservative response criterion (Snodgrass & Corwin, 1988).

Source memory. Similarly to item recognition, only a main effect of source was observed, $F(1, 27) = 71.96$, $p < .001$, $\eta_p^2 = .73$. This effect showed once again a general benefit of the self-referential condition (M = 0.54, SE = 0.04) compared to the non-self-referential condition (M = 0.26, SE = 0.03; p

< .001; d = 1.60, 95% CI [0.21, 0.34]; see Table 1 and Figure 2d). Neither a main effect of valence, F(2, 54) = 0.55, p = .580, η_p^2 = .02, nor an interaction effect reached statistical significance, F(2, 54) = 1.15, p = .314, η_p^2 = .04, ε = .77.

As in Experiment 1, two additional repeated-measures ANOVA were run considering the individual IAPS valence ratings provided by the participants. Nonetheless, since block design was used in Experiment 2, we only considered trials in a given block whose individual ratings matched the valence category previously assigned to that block. For example, in the positive block, we only considered trials in which pictures were deemed positive for a given participant. In the first analysis, a main effect of source, F(1, 27) = 69.23, p < .001, η_p^2 = .72, and a main effect of valence were found, F(2, 54) = 3.58, p = .035, η_p^2 = .12, yet no interaction effect, F(2, 54) = 0.66, p = .482, η_p^2 = .02, ε = .75. Once again, the source memory performance associated with words studied in the self-referential condition (M = 0.55, SE = 0.04) was improved compared to words studied in the common condition (M = 0.27, SE = 0.03; p < .001, d = 1.57, 95% CI [0.21, 0.34]). Moreover, the valence effect was reflected in an improved source memory performance for neutral words presented in the positive (M = 0.46, SE = 0.04) relative to the negative block (M = 0.37, SE = 0.04; p = .030, d = 0.53, 95% CI [0.01, 0.19]; see Supplementary Table S12). Bayesian analysis supported the two main effects model as well as the difference between negative and positive conditions (see Supplementary Material section). In the second analysis including trials with more extreme values of negative and positive valence only a main effect of source was observed, F(1, 24) = 43.53, p < .001, η_p^2 = .65 [valence: F(1, 24) = 43.53, p < .001, η_p^2 = .65; interaction: F(2, 48) = 0.55, p = .582, η_p^2 = .02], supporting the memory advantage for words studied in the self-referential condition (M = 0.54, SE = 0.05 for self-referenced items; M = 0.27, SE = 0.04 for items studied in the common condition; p < .001, d = 1.32, 95% CI [0.19, 0.36]).

Response time.

Study phase. Statistically significant main effects of source, F(1, 27) = 11.45, p = .002, η_p^2 = .30, and of valence, F(2, 54) = 3.66, p = .032, η_p^2 = .12, were observed. No interaction effect emerged, F(2, 54) = 0.39, p = .681, η_p^2 = .01. The mean response time for words presented in the neutral block during the encoding phase was faster (M = 1868, SE = 43.03) than for words presented in the negative block (M = 1971, SE = 46.40; p = .011, d = 0.60, 95% CI [20.53, 185.63]). Additionally, participants were slower at responding in self-referential trials (M = 1964, SE = 41.18) compared to common trials (M = 1874, SE = 39.34; p = .002, d = 0.64, 95% CI [35.70, 145.74]; see Table 1 and Figure 3b).

Correct source responses. The 3 x 2 Friedman's ANOVA was statistically significant, $X^2(5) = 21.93$, $p = .001$. The follow-up Wilcoxon tests showed that, despite no significant differences between the response time of correct source responses in the distinct valence categories (see Supplementary Table S5), an effect of source was observed ($p < .006$ applying the Bonferroni correction for multiple comparisons). Specifically, a statistically significant difference in response time was found between common and self-encoding conditions for words presented in emotional backgrounds, both negative ($p = .002$) and positive ($p < .001$). This is in line with the main effect of source already described in Experiment 1 (see Supplementary Table S5, Table 1, and Figure 3d).

Manipulation check.

As in the original ratings reported in the ANEW (see Supplementary Table S1) and in Experiment 1, the mean valence ratings provided by the participants differed significantly between negative, neutral, and positive blocks, $F(2, 87) = 286.53$, $p < .001$, $\eta_p^2 = .87$. Negative pictures received the lowest ratings ($M = 2.99$, $SD = 0.58$), neutral images received intermediate ratings ($M = 4.69$, $SD = 0.84$), and positive images obtained the highest valence ratings ($M = 7.18$, $SD = 0.59$; $p < .001$ and $d > 2.35$ for all pairwise comparisons; positive *vs.* neutral: 95% CI [2.06, 2.92]; positive *vs.* negative: 95% CI [3.75, 4.61]; neutral *vs.* negative: 95% CI [1.26, 2.12]).

Of note, and relevant for the current experiment was the analysis of the PANAS difference scores considering both positive and negative affect scores (see Supplementary Table S7). For positive affect scores, the repeated-measures ANOVA yielded no statistically significant differences between negative, neutral, and positive encoding contexts, $F(2, 54) = 1.54$, $p = .224$, $\eta_p^2 = .05$ (negative: $M_{difference} = -2.11$, $SD = 4.87$; neutral: $M_{difference} = -2.75$, $SD = 4.59$; positive: $M_{difference} = -1.82$, $SD = 5.30$). A similar result was obtained in the case of negative affect scores, $X^2(2) = 3.62$, $p = .164$ (negative: $M_{difference} = 0.18$, $SD = 1.91$; neutral: $M_{difference} = -0.07$, $SD = 1.49$; positive: $M_{difference} = -0.43$, $SD = 0.92$). As a supplementary analysis, we also tested if there were any significant differences between the PANAS scores reported after the negative, neutral, and positive encoding blocks (without considering the baseline). Once again, for both positive and negative affect, no statistically significant differences emerged [$F(2, 54) = 1.54$, $p = .224$, $\eta_p^2 = .05$; $X^2(2) = 2.88$, $p = .237$, respectively]. Overall, in the current experiment, the repetitive exposure to emotional or neutral pictures did not result in significant changes in self-reported positive or negative affect.

Discussion: Experiment 2

In the current experiment, the main analyses failed to support a valence effect on both source and item memory recognition. Thus, the main effect of valence initially found in Experiment 1 did not emerge when considering the participant-based analysis in Experiment 2. Nonetheless, contrary to Experiment 1, the secondary analysis involving the participants' valence ratings yielded a main effect of valence, which was reflected in a worse source memory performance for neutral words encoded in the negative block in contrast to words encoded in the positive block. Even so, caution is warranted in the interpretation of these findings as only a portion of the initial trials was added to the analyses, creating an imbalance in the number of trials included across experimental conditions. Moreover, in the additional analysis of incorrect source responses, "do not know" responses, misses, and false alarms (see Supplementary Material section), no effects of valence were observed. Taken together, the absence of an EEM effect in the case of item memory was in line with Experiment 1 and with previous experiments (e.g., Baeken et al., 2012; Bowen & Kensinger, 2017; Fenker et al., 2005). Notwithstanding, in the case of source memory recognition, only one of the secondary analyses provided partial evidence in favor of the working hypothesis that an enhancement of sustained interference or the elicitation of mood changes during encoding, especially in a negative context, would affect internal source memory performance.

With respect to the self-reported positive and negative affect, the repetitive exposure to emotionally charged pictures, especially to negative ones, did not elicit significant changes in the subjective assessment of mood states. The qualitative observation of both negative and positive affect scores (see Supplementary Table S7) revealed that, despite a small change in negative affect between baseline and experimental blocks, the positive affect appeared to decrease after each experimental block as reported in previous studies (Denkova et al., 2010; Iordan, Dolcos, Denkova, & Dolcos, 2013). The decrease in positive affect was task-related and similar in all experimental blocks. Together with the small change in negative affect, this decrease might explain why the difference between experimental blocks and baseline did not reach statistical significance. Consequently, it may account for the memory performance findings.

The repetitive exposure to negative pictures was expected to enhance the interference effect on memory. Nonetheless, the slower response times observed during encoding in the negative block when compared to the neutral block represented the only evidence in favor of an interference effect on behavioral performance, which was not observed in Experiment 1. Slower responses in unpleasant blocks could be accounted for by mood changes, possibly associated with the activation of the motivational defensive system (M. G. Pereira et al., 2006).

The absence of subjective affect changes contrasts with previous studies (e.g., Figueira et al., 2017; Pinheiro et al., 2017). However, two main factors may explain the null effect. First, the emotional pictures used in other studies were more arousing than non-emotional ones (Bradley, Cuthbert, & Lang, 1996; Figueira et al., 2017; Güntekin & Tülay, 2014; Pinheiro et al., 2017), whereas in the current study pictures were controlled for arousal. Second, the time of exposure to each emotional picture and the number of pictures in each block was higher in prior studies (Bradley et al., 1996; Figueira et al., 2017; J. C. Smith, Bradley et al., 2005). Overall, under some specific circumstances (e.g., shorter blocks of emotional pictures that are controlled for arousal), the repetitive exposure to emotional stimuli might not lead to significant self-reported affect changes.

General discussion

The main goal of this study was to explore how emotional pictures presented during encoding might impact upon the internal source memory recognition of neutral words. Even though previous evidence revealed that the presence of negative (*vs.* neutral) information has a beneficial (Macri et al., 2018; Xie & Zhang, 2017) or a null effect (Subramaniam et al., 2016; Ventura-Bort et al., 2017) on source memory performance, at least when considering external source memory and reality monitoring, the effects of emotional encoding contexts on internal source memory remained to be clarified. In two experiments, one using a mixed design and another using a blocked design, we tested the hypothesis that emotional pictures, particularly negative ones, would interfere with the processing of neutral words (Sakaki et al., 2011) and, consequently, impact upon internal source memory. This hypothesis was based on the notion that accurate internal source memory performance relies on remembering details about cognitive operations that take place during encoding, such as how information was elaborated, organized or even imagined (Johnson et al., 1993). Also, it was based on prior evidence showing that the discrimination of externally *vs.* internally derived memories differs as the former memories are richer in temporal, spatial, and perceptual details, whereas the latter are mainly characterized by cognitive operations (Ferguson et al., 1992). Hence, the influence of emotional stimuli could also differ as a function of the nature and type of source memory discrimination (see Schmidt et al., 2011 for a similar argument), as already documented by the few available studies on external source memory (Macri et al., 2018; Ventura-Bort et al., 2017; Xie & Zhang, 2017) and reality monitoring (Subramaniam et al., 2016). Nonetheless, the results did not provide evidence supporting impaired internal source memory for neutral items initially encoded in negative contexts. The only partial evidence for a detrimental effect of negative valence on source memory was obtained in Experiment 2 considering an additional analysis based on

participants' valence ratings. Notwithstanding, in both experiments we replicated the beneficial effect of self-referential conditions on both item and internal source memory when compared to other deep encoding tasks such as the common judgment task (Culcea & Freitas, 2017; Durbin et al., 2017; Fossati et al., 2004; Hamami, Serbun, & Gutchess, 2011; Leshikar & Duarte, 2012, 2014; Leshikar et al., 2015; D. R. Pereira et al., 2019; Serbun, Shih, & Gutchess, 2011; Rogers et al., 1977; Symons & Johnson, 1997). Additionally, the self-referential (*vs.* common) condition was associated with longer response times during encoding and faster response times in the case of source hits, which might indicate enhanced elaborative processes in the self-referential condition during encoding and easier access to self-referentially encoded items during recognition.

A null finding concerning source memory was also hypothesized given the behavioral findings of Ventura-Bort and colleagues (2017) on memory for the location of neutral objects, as well as the findings of Subramaniam and colleagues (2016) documenting the lack of an effect of negative affect on reality monitoring. It is worth noting that in studies probing emotion effects on internal source memory in which emotion is an intrinsic feature of the target stimuli (see Appendix table from D. R. Pereira et al., 2019 for a selective overview of studies on internal source memory), one of two possible outcomes have been reported: a null or an impairment effect for emotional stimuli when compared to neutral stimuli. The current findings indicate some degree of convergence between intrinsic and extrinsic manipulations of emotion during encoding on internal source memory performance. Even so, these findings also contrast with previous studies using an extrinsic manipulation of emotion, which reported a memory advantage for neutral items studied in negative contexts, considering location (Macri et al., 2018), orientation, and color features of neutral items (Xie & Zhang, 2017). Whereas these studies support the priority-binding theory (MacKay et al., 2004) by showing that negative arousing information facilitates the recollection of details associated with a specific item, the results of the current study are better accounted for by the arousal-biased competition theory (Mather & Sutherland, 2011). Specifically, as in the current study participants were instructed to pay attention to both item and background, they could have assigned similar relevance to information presented in both emotional and neutral encoding contexts. Moreover, the use of intentional encoding conditions might have enhanced the goal-relevance of the words irrespective of whether they were studied in an emotional or neutral context.

An alternative account relates to the arousal properties of the stimuli used as distractors. Whereas in the studies mentioned above negative background stimuli were more arousing than neutral background stimuli (Macri et al., 2018; Ventura-Bort et al., 2017; Xie & Zhang, 2017), in the current study both negative, positive, and neutral backgrounds were controlled for arousal (see Supplementary Table S1).

Furthermore, prior studies indicated that the presentation of arousing stimuli contiguously or concomitantly with non-arousing stimuli might not disrupt or benefit feature binding in the case of non-arousing stimuli (Mather et al., 2009; Mather & Nesmith, 2008). Similarly, as the words used in the current study were neutral, their associations with the details of the encoding task (self-referential *vs.* common) might not have been particularly affected by the simultaneous presentation of arousing background stimuli. Taken together, controlled levels of arousal and effortful encoding conditions may have resulted in a similar prioritization of item and source details for negative, positive, and neutral encoding contexts.

Another critical aspect to consider is the time elapsed between study and test as prior studies indicated that optimal effects of emotion on memory are obtained when longer study-test intervals are used (Yonelinas & Ritchey, 2015), supporting the operation of consolidation processes (Hamann, 2001). As this factor may account for the lack of emotion effects on item memory studies (e.g., K. J. Mitchell et al., 2006; Sharot & Yonelinas, 2008; Wang, 2018), it is plausible that emotion-related effects on source memory are observed if longer study-test intervals are considered. Notwithstanding, most of the studies that explored the effect of emotional encoding contexts on source memory also used short study-test intervals (e.g., Macri et al., 2018; Xie & Zhang, 2017; Subramaniam et al., 2016) and reported significant effects of emotional relative to neutral conditions. The only exception was the study of Ventura-Bort and colleagues (2017), which used a one-week interval, showing that memory for the location of objects was at chance-level. Hence, the presence of emotional information during encoding in this case did not ensure that certain elements of the encoding episode resisted forgetting. Collectively, these studies highlight the need to further investigate the role of study-test time in the relationship between source memory and emotion.

Considering that a deep analysis of each word was required during encoding (Craik & Lockhart, 1972; Craik & Tulving, 1975), it is also possible that the cognitive demands imposed by the self-reference/common judgment task moderated the influence of emotional contexts. Effortful tasks were found to modulate the response to emotional information in previous studies (e.g., Kellermann et al., 2012; Van Dillen & Koole, 2007). Furthermore, the encoding tasks allowed participants to have a more consistent encoding strategy across emotional and neutral distractors. Although this is a plausible explanation for the lack of valence effects, it cannot be tested here and should be examined in future studies.

Even though the arousal-biased competition theory (Mather & Sutherland, 2011) accommodates the null findings reported here, an additional analysis based on the individual valence ratings provided

by the participants in Experiment 2 partially supported our initial hypothesis concerning the effects of negative contexts. Specifically, this analysis revealed that words presented in a negative context were associated with a worse performance compared to those presented in a positive context. We initially hypothesized a difference between negative and neutral contexts, particularly considering behavioral indices of interference during encoding, i.e., longer response times in the negative compared to neutral block (Blair et al., 2007; Erthal et al., 2005; D. Mitchell et al., 2007; Nielen et al., 2009). Nevertheless, a difference was found between negative and positive contexts, which partially agrees with the study of Subramaniam and colleagues (2016) showing a benefit in reality monitoring performance when a positive mood was induced. Indeed, Experiment 2 examined the possibility that interference effects were driven by temporary affect changes prompted by the exposure to emotional pictures. Specifically, negative affect has been associated with the occurrence of irrelevant thoughts and with the narrowing of attention to specific features (e.g., an object such as a gun presented in a visually complex background), which in turn can interfere with cognitive processes required by the tasks at hand (Fredrickson & Branigan, 2005; Kensinger, 2009; Sakaki et al., 2011). In its turn, positive affect promotes broader attention and heuristic processing (also enhancing cognitive flexibility and working memory), which may facilitate task-relevant performance (e.g., Ashby, Isen, & Turken, 1999; Fredrickson & Branigan, 2005). Since memory for negative information seems to include more internal details (e.g., thoughts; feelings) from the encoding episode when compared to positive information (see Mickley & Kensinger, 2009), a stronger conflict might have occurred between the internal details spontaneously prompted by the negative pictures and the cognitive operations required by the neutral words. On the contrary, positive mood seems to promote a wider attentional scope and better cognitive flexibility during conflict resolution, which may have accounted for the advantage observed in internal source memory performance. However, caution is needed when interpreting this result, since it is only supported by an additional analysis with fewer trials. Moreover, no changes in negative or positive affect were detected in self-report measures, and only the slower response times verified during the negative encoding block might be indicative of temporary changes in affect (M. G. Pereira et al., 2006). Therefore, the lack of a more conclusive valence effect on source memory might also be related to the fact that possible emotional interference effects experienced during the encoding phase were not detrimental, giving the participants the opportunity to manage their cognitive resources and to ensure a stable memory performance.

Limitations and future research

First, as no manipulation check was conducted in terms of pictures' arousal, it is possible that participants' ratings differed from the normative ones (e.g., García-Pacios, Río, Villalobos, Ruiz-Vargas, & Maestú, 2015). As such, it is not possible to rule out potential effects of individual differences in arousal ratings on the current results. Future studies should also consider the possibility of collecting ratings for relevant variables prior to the experimental session in order to select a pool of stimuli that comply with both experimental requirements and individual differences. Such idiosyncratic selection would also accommodate sex differences in the processing of emotional stimuli (e.g., García-García et al., 2016; Stevens & Hamann, 2012), which are manifested in affective ratings (e.g., Soares et al., 2015). Indeed, since the experiments described here were mainly composed of female participants, this precludes the generalization of results.

Second, the type of recognition test here might have facilitated both item and source memory recognition as immediate study-test conditions were used. Furthermore, the retrieval phase of prior studies appeared to be more challenging as participants were required to do free-recall in the case of item memory (e.g., Erk et al., 2003), or to remember precise details such as the exact color and location in the case of source memory (e.g., Macri et al., 2018; Xie & Zhang, 2017).

As a final remark, the experimental design of emotional interference studies has been quite variable. While some studies present the interfering information before the critical item (Erk et al., 2003; Xie & Zhang, 2017), others present both types of stimuli concomitantly (Guo et al., 2018; Macri et al., 2018; Pierguidi et al., 2016; Ventura-Bort et al., 2017). Future studies should test whether these approaches lead to different effects and explore the role of factors such as exposure time, number and type of interfering stimuli.

Conclusion

To the best of our knowledge, this is the first study to demonstrate that the encoding of critical neutral words in emotional (*vs.* neutral) contexts did not significantly affect behavioral performance in an internal source memory task. Notwithstanding, participants recognized more words that were encoded in the self-referential (*vs.* common) condition. This effect confirms a memory advantage of self-referential encoding, not only in the case of item memory but also of source memory attributions (Durbin et al., 2017; Leshikar & Duarte, 2014; D. R. Pereira et al., 2019; Serbun et al., 2011). The lack of a valence effect in the current study (in which stimuli were controlled for arousal) raises the question of whether memory changes associated with emotional interference during encoding might become more evident

when negative stimuli are highly arousing. Nonetheless, the null effect does not imply that the presentation of negative or positive distractors fails to modulate encoding and recognition processes, as in previous studies null behavioral findings were accompanied with differences in electrophysiological brain responses (Ventura-Bort et al., 2017). Furthermore, when individual valence ratings were considered in source memory analysis of Experiment 2, a valence effect emerged. In this context, future studies should consider not only orthogonal manipulations of valence and arousal to better disentangle the effects of each affective dimension, but they should also include more fine-grained trial-by-trial measures of emotional-related modulations (e.g., electrophysiological brain responses and self-reported measures of affect) that better capture individual responses to emotional information. Furthermore, more studies are needed to probe if the null findings reported here are replicated when testing different internal source memory features in similar experimental conditions. Likewise, other external source memory and reality-monitoring features could be further tested. The current findings open a new avenue for future research shedding light on how emotional information, even when irrelevant to the current task goals, influences different episodic memory features.

References

Anderson, A. K., Wais, P. E., & Gabrieli, J. D. E. (2006). Emotion enhances remembrance of neutral events past. *Proceedings of the National Academy of Sciences of the United States of America, 103*(5), 1599–1604. httpss://doi.org/10.1073/pnas.0506308103

Ashby, F. G., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review, 106*(3), 529-550. httpss://doi.org/10.1037/0033-295X.106.3.529

Baeken, C., De Raedt, R., Van Schuerbeek, P., De Mey, J., Bossuyt, A., & Luypaert, R. (2012). The influence of emotional priming on the neural substrates of memory: A prospective fMRI study using portrait art stimuli. *NeuroImage, 61*(4), 876–883. httpss://doi.org/10.1016/j.neuroimage.2012.03.043

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1-24. httpss://doi.org/10.1037/0033-2909.133.1.1

Beck, A. T., Steer, R., & Brown, G. (1996). *Beck Depression Inventory-II manual*. New York: The Psychological Corporation.

Bisby, J. A., & Burgess, N., (2013). Negative affect impairs associative memory but not item memory. *Learning & Memory, 21*(1), 760-766. httpss://doi.org/10.1101/lm.032409.113

Bisby, J. A., & Burgess, N. (2017). Differential effects of negative emotion on memory for items and associations, and their relationship to intrusive imagery. *Current Opinion in Behavioral Sciences, 17,* 124-132. httpss://doi.org/10.1016/j.cobeha.2017.07.012

Bisby, J. A., Horner, A., J., Hørlyck, L. D., & Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social Cognitive and Affective Neuroscience, 11*(6), 981-990. httpss://doi.org/10.1093/scan/nsw028

Blair, K. S., Smith, B. W., Mitchell, D. G. V, Morton, J., Vythilingam, M., Pessoa, L., ...Blair, R. J. R. (2007). Modulation of emotion by cognition and cognition by emotion. *NeuroImage, 35*(1), 430–440. httpss://doi.org/10.1016/j.neuroimage.2006.11.048

Bocanegra, B. R., & Zeelenberg, R. (2009). Dissociating emotion-induced blindness and hypervision. *Emotion, 9*(6), 865–873. httpss://doi.org/10.1037/a0017749

Bowen, H. J., & Kensinger, E. A. (2017). Recapitulation of emotional source context during memory retrieval. *Cortex, 91*, 142–156. httpss://doi.org/10.1016/j.cortex.2016.11.004

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1996). Picture media and emotion: Effects of a sustained affective context. *Psychophysiology, 33*(6), 662-670. httpss://doi.org/10.1111/j.1469-8986.1996.tb02362.x

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience, 117*(2), 369–380. httpss://doi.org/10.1037/0735-7044.117.2.369

Bridge, D. J., Chiao, J. Y., & Paller, K. A. (2010). Emotional context at learning systematically biases memory for facial information. *Memory & Cognition, 38*(2), 125–133. httpss://doi.org/10.3758/MC.38.2.125

Brierley, B., Medford, N., Shaw, P., & David, A. S. (2007). Emotional memory for words: Separating content and context. *Cognition and Emotion, 21*(3), 495–521. httpss://doi.org/10.1080/02699930600684963

Calvo, M. G., & Lang, P. J. (2004). Gaze patterns when looking at emotional pictures: Motivationally biased attention. *Motivation and Emotion, 28*(3), 221–243. httpss://doi.org/10.1023/B:MOEM.0000040153.26156.ed

Carretié, L., Hinojosa, J. A., Martín-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping, 22*(4), 290–299. httpss://doi.org/10.1002/hbm.20037

Cedrus Coporation. (2015). Superlab (Version 5.0) [Computer software].

Coelho, R., Martins, A., & Barros, H. (2002). Clinical profiles relating gender and depressive symptoms among adolescents ascertained by the Beck Depressive Inventory II. *European Psychiatry, 17*(4), 222-226. httpss://doi.org/10.1016/S0924-9338(02)00663-6

Cook, G. I., Hicks, J. L., & Marsh, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition, 35*(2), 222–230. httpss://doi.org/10.3758/BF03193443

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684. httpss://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294. httpss://doi.org/10.1037/0096-3445.104.3.268

Cui, L., Shi, G., He, F., Zhang, Q., Oei, T. P. S., & Guo, C. (2016). Electrophysiological correlates of emotional source memory in high-trait-anxiety individuals. *Frontiers in Psychology, 7*, 1039. httpss://doi.org/10.3389/fpsyg.2016.01039

Culcea, I. C., & Freitas, A. L. (2017). Subjective importance as a cue for self-reference. *Personality and Social Psychology Bulletin, 43*(8), 1100-1111. httpss://doi.org/10.1177/0146167217705111

Denkova, E., Wong, G., Dolcos, S., Sung, K., Wang, L., Coupland, N., & Dolcos, F. (2010). The impact of anxiety-inducing distraction on cognitive performance: A combined brain imaging and personality investigation. *PLoS ONE, 5*(11), e14150. httpss://doi.org/10.1371/journal.pone.0014150

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1*(1), 5–11. httpss://doi.org/10.1037/1528-3542.1.1.5

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: Encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective, & Behavioral Neuroscience, 2*(3), 252–263. httpss://doi.org/10.3758/CABN.2.3.252

Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology, 23*(6), 669–694. https://doi.org/10.1080/20445911.2011.594433

Dulas, M. R., & Duarte, A. (2014). Aging affects the interaction between attentional control and source memory: An fMRI study. *Journal of Cognitive Neuroscience, 26*(12), 2653–2669. https://doi.org/10.1162/jocn_a_00663

Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory, 25*(9), 1191-1200. https://doi.org/10.1080/09658211.2017.1282517

Ellis, H. C., Thomas, R. L., & Rodriguez, I. A. (1984). Emotional mood states and memory: Elaborative encoding, semantics processing, and cognitive effort. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(3), 470–482. https://doi.org/10.1037/0278-7393.10.3.470

Erk, S., Kiefer, M., Grothe, J., Wunderlich, A. P., Spitzer, M., & Walter, H. (2003). Emotional context modulates subsequent memory effect. *NeuroImage, 18*(2), 439–447. https://doi.org/10.1016/S1053-8119(02)00015-0

Erk, S., Martin, S., & Walter, H. (2005). Emotional context during encoding of neutral items modulates brain activation not only during encoding but also during recognition. *NeuroImage, 26*(3), 829–838. https://doi.org/10.1016/j.neuroimage.2005.02.045

Erthal, F. S., Oliveira, L., Mocaiber, I., Pereira, M. G., Machado-Pinheiro, W., Volchan, E., & Pessoa, L. (2005). Load-dependent modulation of affective picture processing. *Cognitive, Affective, & Behavioral Neuroscience, 5*(4), 388–395. https://doi.org/10.3758/CABN.5.4.388

Fan, W., Zhong, Y., Li, J., Yang, Z., Zhan, Y., Cai, R., & Fu, X. (2016). Negative emotion weakens the degree of self-reference effect: Evidence from ERPs. *Frontiers in Psychology, 7*, 1408. https://doi.org/10.3389/fpsyg.2016.01408

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. https://doi.org/10.3758/BF03193146

Fenker, D. B., Schott, B. H., Richardson-Klavehn, A., Heinze, H. J., & Düzel, E. (2005). Recapitulating emotional context: Activity of amygdala, hippocampus and fusiform cortex during recollection and familiarity. *European Journal of Neuroscience, 21*(7), 1993–1999. https://doi.org/10.1111/j.1460-9568.2005.04033.x

Ferguson, S. A., Hastroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging, 7*(3), 443–452. https://doi.org/10.1037/0882-7974.7.3.443

Ferré, P., Comesaña, M., & Guasch, M. (2019). Emotional content and source memory for language: Impairment in an incidental encoding task. *Frontiers in Psychology, 10*, 65. https://doi.org/10.3389/fpsyg.2019.00065

Figueira, J. S. B., Oliveira, L., Pereira, M. G., Pacheco, L. B., Lobo, I., Motta-Ribeiro, G. C., & David, I. A. (2017). An unpleasant emotional state reduces working memory capacity: Electrophysiological evidence. *Social Cognitive and Affective Neuroscience, 12*(6), 984–992. https://doi.org/10.1093/scan/nsx030

Fossati, P., Hevenor, S. J., Lepage, M., Graham, S. J., Grady, C., Keightley, M. L., ... Mayberg, H. (2004). Distributed self in episodic memory: Neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage, 22*(4), 1596–1604. https://doi.org/10.1016/j.neuroimage.2004.03.034

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion, 19*(3), 313–332. https://doi.org/10.1080/02699930441000238

Galinha, I. C., & Pais-Ribeiro, J. (2005). Contribuição para o estudo da versão portuguesa da Positive and Negative Affect Schedule (PANAS): II – Estudo psicométrico. *Análise Psicológica, 2*, 219-227.

García-García, I., Kube, J., Gaebler, M., Horstmann, A., Villringer, A., & Neumann, J. (2016). Neural processing of negative emotional stimuli and the influence of age, sex and task-related characteristics. *Neuroscience & Biobehavioral Reviews, 68*, 773-793. https://doi.org/10.1016/j.neubiorev.2016.04.020

García-Pacios, J., Río, D. Del, Villalobos, D., Ruiz-Vargas, J. M., & Maestú, F. (2015). Emotional interference-based forgetting in short-term memory. Cognitive inhibition of pleasant but not unpleasant biologically relevant distractors. *Frontiers in Psychology, 6*, 582. https://doi.org/10.3389/fpsyg.2015.00582

Gärtner, M., & Bajbouj, M. (2014). Encoding-related EEG oscillations during memory formation are modulated by mood state. *Social Cognitive and Affective Neuroscience, 9*(12), 1934–1941. https://doi.org/10.1093/scan/nst184

Gingerich, A. C., & Dodson, C. S. (2013). Sad mood reduces inadvertent plagiarism: Effects of affective state on source monitoring in cryptomnesia. *Motivation and Emotion, 37*(2), 355-371. https://doi.org/10.1007/s11031-012-9309-2

Graf, P., & Yu, M. (2015). Valenced cues and contexts have different effects on event-based prospective memory. *PLoS ONE, 10*(2), e0116953. https://doi.org/10.1371/journal.pone.0116953

Güntekin, B., & Tülay, E. (2014). Event related beta and gamma oscillatory responses during perception of affective pictures. *Brain Research, 1577*, 45–56. https://doi.org/10.1016/j.brainres.2014.06.029

Guo, X., Li, Y., Zhang, Q., Cui, L., & Wei, P. (2018). Weakening of recognition memory caused by task-irrelevant emotional encoding context can be modulated by individuals' inhibitory control. *Personality and Individual Differences, 134*, 201-209. https://doi.org/10.1016/j.paid.2018.06.012

Hamami, A., Serbun, S. J., & Gutchess, A. H. (2011). Self-referencing enhances memory specificity with age. *Psychology and Aging, 26*(3), 636–646. https://doi.org/10.1038/jid.2014.371

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences, 5*(9), 394–400. https://doi.org/10.1016/S1364-6613(00)01707-1

Han, M., Mao, X., Kartvelishvili, N., Li, W., & Guo, C. (2018). Unitization mitigates interference by intrinsic negative emotion in familiarity and recollection of associative memory: Electrophysiological evidence. *Cognitive, Affective, & Behavioral Neuroscience, 18*(6), 1259-1268. https://doi.org/10.3758/s13415-018-0636-y

Iordan, A. D., Dolcos, S., Denkova, E., & Dolcos, F. (2013). Sex differences in the response to emotional distraction: an event-related fMRI investigation. *Cognitive, Affective, & Behavioral Neuroscience, 13*(1), 116–134. https://doi.org/10.3758/s13415-012-0134-6

Iordan, A. D., Dolcos, S., & Dolcos, F. (2013). Neural signatures of the response to emotional distraction: A review of evidence from brain imaging investigations. *Frontiers in Human Neuroscience, 7*, 200. https://doi.org/10.3389/fnhum.2013.00200

Jaeger, A., Johnson, J. D., Corona, M., & Rugg, M. D. (2009). ERP correlates of the incidental retrieval of emotional information: Effects of study-test delay. *Brain Research, 1269*, 105–113. https://doi.org/10.1016/j.brainres.2009.02.082

Jaeger, A., & Rugg, M. D. (2012). Implicit effects of emotional contexts: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience, 12*(4), 748–760. https://doi.org/10.3758/s13415-012-0110-1

JASP Team. (2018). JASP (Version 0.9.0.1) [Computer software].

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28. https://doi.org/10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67–85. https://doi.org/10.1037/0033-295X.88.1.67

Kellermann, T. S., Sternkopf, M. A., Schneider, F., Habel, U., Turetsky, B. I., Zilles, K., & Eickhoff, S. B. (2012). Modulating the processing of emotional stimuli by cognitive demand. *Social Cognitive and Affective Neuroscience, 7*(3), 263-273. https://doi.org/10.1093/scan/nsq104

Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L., & Ingram, R. E. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. *Cognitive Therapy and Research, 11*(3), 289-299. https://doi.org/10.1007/BF01186280

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences, 15*(4), 241–253. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review, 1*(2), 99–113. https://doi.org/10.1177/1754073908100432

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition, 31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America, 101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A, Gutchess, A. H., & Schacter, D. L. (2007). Effects of aging and encoding instructions on emotion-induced memory trade-offs. *Psychology and Aging, 22*(4), 781–795. https://doi.org/10.1037/0882-7974.22.4.781

Kensinger, E. A., & Schacter, D. L. (2006). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *Journal of Neuroscience, 26*(9), 2564-2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

Koenig, S., & Mecklinger, A. (2008). Electrophysiological correlates of encoding and retrieving emotional events. *Emotion*, *8*(2), 162–173. https://doi.org/10.1037/1528-3542.8.2.162

Kohn, N., Falkenberg, I., Kellermann, T., Eickhoff, S. B., Gur, R. C., & Habel, U. (2014). Neural correlates of effective and ineffective mood induction. *Social Cognitive and Affective Neuroscience, 9*(6), 864–872. https://doi.org/10.1093/scan/nst055

Koster, E. H. W., De Raedt, R., Goeleven, E., Franck, E., & Crombez, G. (2005). Mood-congruent attentional bias in dysphoria: Maintained attention to and impaired disengagement from negative information. *Emotion, 5*(4), 446-455. https://doi.org/10.1037/1528-3542.5.4.446

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97*(3), 377–395. https://doi.org/10.1037/0033-295X.97.3.377

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International Affective Picture System (IAPS): Technical manual and affective ratings.* Gainesville, FL: University of Florida, NIMH Center for the Study of Emotion and Attention.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry, 44*(12), 1248-1263. https://doi.org/10.1016/S0006-3223(98)00275-3

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual* (Technical Report No. A-8). Gainesville, FL: University of Florida, NIMH Center for the Study of Emotion and Attention.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Leshikar, E. D., & Duarte, A. (2012). Medial prefrontal cortex supports memory accuracy for self-referenced items. *Social Neuroscience, 7*(2), 126–145. https://doi.org/10.1080/17470919.2011.585242.Medial

Leshikar, E. D., & Duarte, A. (2014). Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience, 14*(1), 236–252. https://doi.org/10.3758/s13415-013-0198-y

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition, 22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Lv, J. Y., Wang, T., Tu, S., Zheng, F., & Qiu, J. (2011). The effect of different negative emotional context on involuntary attention: An ERP study. *Brain Research Bulletin, 86*(1–2), 106–109. https://doi.org/10.1016/j.brainresbull.2011.06.010

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition, 32*(3), 474–488. https://doi.org/10.3758/BF03195840

Macri, A., Pavard, A., & Versace, R. (2018). The beneficial effect of contextual emotion on memory: The role of integration. *Cognition and Emotion, 32*(6), 1355-1361. https://doi.org/10.1080/02699931.2017.1387101

Marchewka, A., Wypych, M., Michałowski, J. M., Sińczuk, M., Wordecha, M., Jednoróg, K., & Nowicka, A. (2016). What is the effect of basic emotions on directed forgetting? Investigating the role of basic emotions in memory. *Frontiers in Human Neuroscience, 10*, 378. https://doi.org/10.3389/fnhum.2016.00378

Martínez-Galindo, J. G., & Cansino, S. (2015). Positive and negative emotional contexts unevenly predict episodic memory. *Behavioural Brain Research, 291*, 89–102. https://doi.org/10.1016/j.bbr.2015.05.018

Martínez-Galindo, J. G., & Cansino, S. (2017). Emotional context during encoding modulates recognition electrophysiological activity. *Experimental Brain Research, 235*(1), 169–179. https://doi.org/10.1007/s00221-016-4780-8

Martins, A., Coelho, R., Ramos, E., & Barros, H. (2000). Administração do BDI-II a adolescentes portugueses: Resultados preliminares. *Revista Portuguesa Psicossomática, 2*(1), 123-132.

Mather, M., Gorlick, M. A., & Nesmith, K. (2009). The limits of arousal's memory-impairing effects on nearby information. *American Journal of Psychology, 122*(3), 349-369.

Mather, M., & Nesmith, K., (2008). Arousal-enhanced location memory for pictures. *Journal of Memory and Language, 58*(2), 449-464. https://doi.org/10.1016/j.jml.2007.01.004

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science, 6*(2), 114-133. https://doi.org/10.1177/1745691611400234

Meinhardt, J., & Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. *Cognition and Emotion, 17*(3), 477–500. https://doi.org/10.1080/02699930244000039

Mitchell, D., Nakic, M., Fridberg, D., Kamel, N., Pine, D. S., & Blair, R. (2007). The impact of processing load on emotion. *NeuroImage, 34*(3), 1299–1309. https://doi.org/10.1016/j.neuroimage.2006.10.012

Mitchell, K. J., Mather, M., Johnson, M. K., Raye, C. L., & Greene, E. R. (2006). A functional magnetic resonance imaging investigation of short-term source and item memory for negative pictures. *NeuroReport, 17*(14), 1543-1547. https://doi.org/10.1097/01.wnr.0000234743.50442.e5

Mickley, K. R., & Kensinger, E. (2009). Phenomenological characteristics of emotional memories in younger and older adults. *Memory, 17*(5), 528-543. https://doi.org/10.1080/09658210902939363

Mogg, K., Bradley, B. P., Williams, R., & Mathews, A. (1993). Subliminal processing of emotional information in anxiety and depression. *Journal of Abnormal Psychology, 102*(2), 304-311. https://doi.org/10.1037/0021-843X.102.2.304

Murphy, N. A., & Isaacowitz, D. M. (2008). Preferences for emotional information in older and younger adults: A meta-analysis of memory and attention tasks. *Psychology and Aging, 23*(2), 263-286. https://doi.org/10.1037/0882-7974.23.2.263

Newsome, R. N., Dulas, M. R., & Duarte, A. (2012). The effects of aging on emotion-induced modulations of source retrieval ERPs: Evidence for valence biases. *Neuropsychologia, 50*(14), 3370–3384. https://doi.org/10.1016/j.neuropsychologia.2012.09.024

Nielen, M. M. A., Heslenfeld, D. J., Heinen, K., Van Strien, J. W., Witter, M. P., Jonker, C., & Veltman, D. J. (2009). Distinct brain systems underlie the processing of valence and arousal of affective pictures. *Brain and Cognition, 71*(3), 387–396. https://doi.org/10.1016/j.bandc.2009.05.007

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy in internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology, 10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

Pereira, M. G., Volchan, E., Souza, G. G. L., Oliveira, L., Campagnoli, R., Pinheiro, W. M., & Pessoa, L. (2006). Sustained and transient modulation of performance induced by emotional picture viewing. *Emotion, 6*(4), 622–634. https://doi.org/10.1037/1528-3542.6.4.622

Pierguidi, L., Righi, S., Gronchi, G., Marzi, T., Caharel, S., Giovannelli, F., & Viggiano, M. P. (2016). Emotional contexts modulate intentional memory suppression of neutral faces: Insights from ERPs. *International Journal of Psychophysiology, 106*, 1–13. https://doi.org/10.1016/j.ijpsycho.2016.05.008

Pinheiro, A. P., Barros, C., Dias, M., & Niznikiewicz, M. (2017). Does emotion change auditory prediction and deviance detection? *Biological Psychology, 127*, 123–133. https://doi.org/10.1016/j.biopsycho.2017.05.007

Pinheiro, A. P., Del Re, E., Nestor, P. G., McCarley, R. W., Gonçalves, O. F., & Niznikiewicz, M. (2013). Interactions between mood and the structure of semantic memory: Event-related potentials evidence. *Social Cognitive & Affective Neuroscience, 8*(5), 579-594. https://doi.org/10.1093/scan/nss035

Pinheiro, A. P., Del Re, E., Nestor, P. G., Mezin, J., Rezaii, N., McCarley, R. W., Gonçalves, O. F., & Niznikiewicz, M. (2015). Abnormal interactions between context, memory structure, and mood in schizophrenia: An ERP investigation. *Psychophysiology, 52*(1), 20-31. https://doi.org/10.1111/psyp.12289

Pratto, F., & John, O. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*(3), 380–391. https://doi.org/0022-3514/91/S3.00

Raye, C. L., & Johnson, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society, 15*(6), 405–408. https://doi.org/10.3758/BF03334572

Ritchey, M., LaBar, K. S., & Cabeza, R. (2011). Level of processing modulates the neural correlates of emotional memory formation. *Journal of Cognitive Neuroscience, 23*(4), 757-771. https://doi.org/10.1162/jocn.2010.21487

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Sakaki, M., Gorlick, M. A., & Mather, M. (2011). Differential interference effects of negative emotional states on subsequent semantic and perceptual processing. *Emotion, 11*(6), 1263–1278. https://doi.org/10.1037/a0026329

Santos, S. F., & Silva, D. R. (1997). Adaptação do State-Trait Anxiety Inventory (STAI) – Form Y para a população portuguesa: Primeiros dados. *Revista Portuguesa de Psicologia, 32*, 85-98.

Schmidt, K., Patnaik, P., & Kensinger, E. A. (2011). Emotion's influence on memory for spatial and temporal context. *Cognition and Emotion, 25*(2), 229-243. https://doi.org/10.1080/02699931.2010.483123

Serbun, S. J., Shih, J. Y., & Gutchess, A. H. (2011). Memory for details with self-referencing. *Memory, 19*(8), 1004–1014. https://doi.org/10.1080/09658211.2011.626429

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition, 106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Silva, D. R., & Campos, R. (1998). Alguns dados normativos do Inventário de Estado Traço de Ansiedade - Forma Y (STAI-Y) de Spielberger, para a população portuguesa. *Revista Portuguesa de Psicologia, 33*, 71-89.

Silva, D. R., & Spielberger, C. D. (2007). *Manual do Inventário de Estado Traço de Ansiedade (STAI)*. USA: Mind Garden, Inc.

Smith, A. P. R., Henson, R. N. A., Dolan, R. J., & Rugg, M. D. (2004). fMRI correlates of the episodic retrieval of emotional contexts. *NeuroImage, 22*(2), 868–878. https://doi.org/10.1016/j.neuroimage.2004.01.049

Smith, A. P. R., Henson, R. N. A., Rugg, M. D., & Dolan, R. J. (2005). Modulation of retrieval processing reflects accuracy of emotional source memory. *Learning & Memory, 12*(5), 472–479. https://doi.org/10.1101/lm.84305

Smith, J. C., Bradley, M. M., & Lang, P. J. (2005). State anxiety and affective physiology: Effects of sustained exposure to affective pictures. *Biological Psychology, 69*(3), 247–260. https://doi.org/10.1016/j.biopsycho.2004.09.001

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Soares, A. P., Pinheiro, A. P., Costa, A., Frade, C. S., Comesaña, M., & Pureza, R. (2015). Adaptation of the International Affective Picture System (IAPS) for European Portuguese. *Behavior Research Methods, 47*(4), 1159-1177. https://doi.org/10.3758/s13428-014-0535-2

Stevens, J. S., & Hamann, S. (2012). Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. *Neuropsychologia, 50*(7), 1578-1593. https://doi.org/10.1016/j.neuropsychologia.2012.03.011

Subramaniam, K., Gill, J., Slattery, P., Shastri, A., Mathalon, D. H., Nagarajan, S., & Vinogradov, S. (2016). Neural mechanisms of positive mood induced modulation of reality monitoring. *Frontiers in Human Neuroscience, 10*, 581. https://doi.org/10.3389/fnhum.2016.00581

Subramaniam, K., Ranasinghe, K. G., Mathalon, D., Nagarajan, S., & Vinogradov, S. (2017). Neural mechanisms of mood-induced modulation of reality monitoring in schizophrenia. *Cortex, 91*, 271-286. https://doi.org/10.1016/j.cortex.2017.01.005

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta- analysis. *Psychological Bulletin, 121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Takashima, A., van der Ven, F., Kroes, M. C. W., & Fernández, G. (2016). Retrieved emotional context influences hippocampal involvement during recognition of neutral memories. *NeuroImage*, *143*, 280–292. https://doi.org/10.1016/j.neuroimage.2016.08.069

Talmi, D., Anderson, A. K., Riggs, L., Caplan, J. B., & Moscovitch, M. (2008). Immediate memory consequences of the effect of emotion on attention to pictures. *Learning & Memory*, *15*(3), 172–182. https://doi.org/10.1101/lm.722908

Van Dillen, L. F., & Koole, S. L. (2007). Clearing the mind: A working memory model of distraction from negative mood. *Emotion, 7*(4), 715-723. https://doi.org/10.1037/1528-3542.7.4.715

Ventura-Bort, C., Dolcos, F., Wendt, J., Wirkner, J., Hamm, A. O., & Weymar, M. (2017). Item and source memory for emotional associates is mediated by different retrieval processes. *Neuropsychologia*, Advance online publication. https://doi.org/10.1016/j.neuropsychologia.2017.12.015

Ventura-Bort, C., Löw, A., Wendt, J., Dolcos, F., Hamm, A. O., & Weymar, M. (2016). When neutral turns significant: Brain dynamics of rapidly formed associations between neutral stimuli and emotional contexts. *European Journal of Neuroscience, 44*(5), 2176–2183. https://doi.org/10.1111/ejn.13319

Ventura-Bort, C., Löw, A., Wendt, J., Moltó, J., Poy, R., Dolcos, F., ...Weymar, M. (2016). Binding neutral information to emotional contexts: Brain dynamics of long-term recognition memory. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(2), 234–247. https://doi.org/10.3758/s13415-015-0385-0

Wagenmakers, E-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ...Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*(1), 58-78. https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ...Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25*(1), 35-57. https://doi.org/10.3758/s13423-017-1343-3

Wang, B. (2018). Retention interval modulates the effect of negative arousing pictures on recognition memory. *Memory, 26*(8), 1105-1116. https://doi.org/10.1080/09658211.2018.1427763

Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070. https://doi.org/10.1037/0022-3514.54.6.1063

Watson, L. A., Dritschel, B., Jentzsch, I., & Obonsawin, M. C. (2008). Changes in the relationship between self-reference and emotional valence as a function of dysphoria. *British Journal of Psychology*, *99*(1), 143–152. https://doi.org/10.1348/000712607X248689

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: An emotional binding account. *Trends in Cognitive Sciences, 19*(5), 259-267. https://doi.org/10.1016/j.tics.2015.02.009

Xie, W., & Zhang, W. (2017). Negative emotion enhances mnemonic precision and subjective feelings of remembering in visual long-term memory. *Cognition*, *166*, 73–83. https://doi.org/10.1016/j.cognition.2017.05.025

Zhang, Q., Liu, X., An, W., Yang, Y., & Wang, Y. (2015). Recognition memory of neutral words can be impaired by task-irrelevant emotional encoding contexts: Behavioral and electrophysiological evidence. *Frontiers in Human Neuroscience*, *9*, 73. https://doi.org/10.3389/fnhum.2015.00073

Zlomuzica, A., Preusser, F., Totzeck, C., Dere, E., & Margraf, J. (2016). The impact of different emotional states on the memory for what, where and when features of specific events. *Behavioural Brain Research*, *298*, 181–187. https://doi.org/10.1016/j.bbr.2015.09.037

CHAPTER V

INTERACTIONS OF EMOTION AND SELF-REFERENCE IN SOURCE MEMORY: AN ERP STUDY

Interactions of emotion and self-reference in source memory: An ERP study[1,2]

Abstract

The way emotional information is encoded (e.g., deciding whether it is self-related or not) has been found to affect source memory. However, few studies have addressed how the emotional quality and self-referential properties of a stimulus interactively modulate brain responses during stimulus encoding and source memory recognition. In the current study, twenty-two participants completed five study-test cycles with negative, neutral and positive words encoded in self-referential vs. non-self-referential conditions, while event-related potentials of the electroencephalogram were recorded. An advantage of self-referential processing in source memory performance, reflected in increased recognition accuracy, was shown for neutral and positive words. At the electrophysiological level, self-referential words elicited increased amplitudes in later processing stages during encoding (700 – 1200 ms) and were associated with the emergence of old/new effects in the 300 – 500 ms latency window linked to familiarity effects. In the 500 – 800 ms latency window, old/new effects emerged for all valence conditions except for negative words studied in the non-self-referential condition. Negative self-referential words also elicited a greater mobilization of post-retrieval monitoring processes, reflected in an enhanced mean amplitude in the 800 – 1200 ms latency window. Together, the current findings suggest that valence and self-reference interactively modulate source memory. Specifically, negative self-related information is more likely to interfere with the recollection of source memory features.

Keywords: source memory; valence; self-referential processing; event-related potentials (ERPs); emotion; recognition memory.

The benefits of self-referential processing on memory have been observed not only for the main features of an episode (i.e., item memory; D'Argembeau, Comblain, & Van Der Linden, 2005; Fossati et al., 2004; Kelley et al., 2002; Rogers, Kuiper, & Kirker, 1977; Yin, Ma, Xu, & Yang, 2019), but also for other features associated with its origin (i.e., source memory; Johnson, Hashtroudi, & Lindsay, 1993). Specifically, self-referential processing was found to enhance memory for stimulus location (Yin et al., 2019) and encoding task (Dulas, Newsome, & Duarte, 2011; Durbin, Mitchell, & Johnson, 2017; Hou, Grilli, & Glisky, 2019; Leshikar & Duarte, 2012; Leshikar, Dulas, & Duarte, 2015; Mao, Wang, Wu, &

[1] Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (submitted). Interactions of emotion and self-reference in source memory: An ERP study.

[2] The Supplementary Materials associated with this chapter are presented in Appendix D.

Guo, 2017; Pereira, Sampaio, & Pinheiro, 2019; Y. Zhang, Pan, Li, & Guo, 2018). Self-referential and emotional information might recruit similar mechanisms that enhance encoding (e.g., attentional prioritization; elaboration; rehearsal), resulting in improved retention and retrieval (Gutchess & Kensinger, 2018). Nonetheless, few studies have probed the interactive effects of self-referential encoding and emotion on source memory and their electrophysiological correlates.

Studies examining the influence of emotion on source memory (see Pereira et al., 2019 for a summary table) have reported mixed findings (impaired performance - e.g., Cook, Hicks, & Marsh, 2007; Ferré, Comesaña, & Guasch, 2019, Experiment 2 and 3; MacKenzie, Powell, & Donaldson, 2015; Maddock & Frein, 2009; enhanced performance - e.g., Kensinger & Corkin, 2003; Kensinger & Schacter, 2006b; null effects - e.g., Ferré et al., 2019, Experiment 1; Gallant, Spaniol, & Yang, 2019; Gallant & Yang, 2014; Kensinger & Schacter, 2006a; Koenig & Mecklinger, 2008; Wang & Fu, 2011). The way information is encoded may be a relevant factor to consider in the interpretation of these discrepant findings (see Leynes & Crawford, 2018; Leynes & Mok, 2017 for a similar argument). Furthermore, emotion effects on memory may vary as a function of stimulus relevance given the current goals and contextual demands (Oschner, 2000; Mather, 2011). Accordingly, self-referential encoding (*vs.* non-self-referential encoding) conditions offer a way to examine how self-relevant goals, such as maintaining a favorable and coherent self-view, shape source memory retrieval. As such, the current study specified whether and how self-referential processing and stimulus valence modulate overt encoding and source memory recognition processes, combining behavioral (performance accuracy) and electrophysiological (event-related potentials - ERPs) measures.

Effects of emotion and self-reference during encoding: ERP evidence

In a typical recognition memory task, there is an initial study phase whose content is later subjected to a memory test. During encoding, variations of the affective properties of target stimuli, including valence (the degree of pleasantness/unpleasantness of the stimuli) and arousal (the degree of calmness/excitement elicited by the stimuli), might be already detected. ERPs of the electroencephalogram (EEG) afford excellent temporal resolution to specify when and how the processing of emotionally salient information diverges from the processing of neutral information (Citron, 2012; Hinojosa, Méndez-Bértolo, & Pozo, 2010; Schupp, Flaisch, Stockburger, & Junghöfer, 2006). The affective properties of a stimulus were found to modulate distinct ERP components: the early posterior negativity (EPN), emerging between 150 - 300 ms post-stimulus onset over occipito-temporo-parietal electrode sites (Citron, 2012; Schupp et al., 2006); the late positive complex (LPC), peaking around 500

- 800 ms post-stimulus onset and maximal over parietal electrode sites (Citron, 2012; Foti, Hajcak, & Dien, 2009; Schupp et al., 2006); the slow wave (SW), a late positivity that follows the LPC, lasting for several milliseconds and presenting a more fronto-central topographical distribution (Foti et al., 2009; see Figure 1). The amplitude of these components tends to be consistently larger for emotional compared to neutral stimuli (e.g., Barnacle, Tsivilis, Schaefer, & Talmi, 2018; Codispoti, Ferrari, & Bradley, 2007; Diedrich, Naumann, Maier, Becker, & Bartussek, 1997; Dolcos & Cabeza, 2002; Schacht & Sommer, 2009; Schupp et al., 2006). Notwithstanding, whereas differences between emotional and neutral stimuli on the EPN emerge irrespective of task demands (Schacht & Sommer, 2009; but see Bayer, Sommer, & Schacht, 2012; Fischler & Bradley, 2006), emotion effects on the LPC vary as a function of attentional focus (i.e., if participants are explicitly required to attend to the emotional properties of the stimuli) and task demands (i.e., if a task requires stronger perceptual *vs.* semantic processing; Fischler & Bradley, 2006). Specifically, these effects were enhanced when participants were explicitly asked to evaluate the emotional quality of the stimuli and when semantic processing was required (Diedrich et al., 1997; Fischler & Bradley, 2006; Hajcak, Moser, & Simons, 2006; Hinojosa et al., 2010; Schacht & Sommer, 2009).

Self-referential tasks also demand semantic processing, yet they are expected to elicit greater elaboration and organization of information compared to other tasks (Gutchess & Kensinger, 2018; Symons & Johnson, 1997), which should be reflected in the LPC (Citron, 2012; Herbert, Herbert, Ethofer, & Pauli, 2010). If both emotion and self-reference mobilize greater elaborative processes, the LPC is likely to be modulated accordingly (Gutchess & Kensinger, 2018). Whereas some studies revealed increased LPC amplitude for self-related positive stimuli when compared to both negative and neutral stimuli (e.g., Auerbach, Stanton, Proudfit, & Pizzagalli, 2015; Herbert et al., 2010; Shestyuk & Deldin, 2010 considering healthy controls), others demonstrated increased LPC amplitude for self-related negative stimuli (e.g., Grundy, Benarroch, Lebarr, & Shedden, 2015; Herbert, Pauli, & Herbert, 2011; Zhou et al., 2017). Several factors might account for these discrepant findings, including methodological (e.g., implicit *vs.* explicit self-referential tasks; stimulus properties such as self-relevance; Fan et al., 2013; H. Zhang, Guan, Qi, & Yang, 2013) and participant-related variables (e.g., self-esteem; depression; Auerbach et al., 2015; Nowicka, Wójcik, Kotlewska, Bola, & Nowicka, 2018; Shestyuk & Deldin, 2010; H. Zhang et al., 2013). However, it remains unclear how explicit instructions to memorize target stimuli may affect the interactive effects between emotion and self-reference. Specifically, the use of intentional encoding conditions allows emotional, neutral, self-referential and non-self-referential conditions to be task-relevant,

thereby involving similar levels of prioritization and allocation of attentional resources (see Barnacle et al., 2018 for a similar argument).

So far, the existing studies using intentional learning conditions also revealed increased amplitudes of the LPC and SW for negative and positive stimuli when compared to neutral ones (Barnacle et al., 2018; Dolcos & Cabeza, 2002; Palomba, Angrilli, & Mini, 1997; Wirkner, Ventura-Bort, Schulz, Hamm, & Weymar, 2018). Nonetheless, some studies reported valence-specific effects, namely enhanced LPC amplitude for positive relative to both negative and neutral stimuli (Herbert, Kissler, Junghöfer, Peyk, & Rockstroh, 2006) or for negative relative to both positive and neutral stimuli (Gallant, Pun, & Yang, 2018). Notwithstanding, none of these studies probed the interactive effects between valence and self-reference under explicit encoding instructions. Indirect evidence from studies comparing self-referential processing in depressive and healthy participants revealed increased LPC amplitude for self-related positive words compared to negative words in the case of the control group (Shestyuk & Deldin, 2010). Given the paucity of studies testing the interplay of valence and self-reference during memory encoding, the current study examined the effects of these variables on the LPC and SWs in the context of intentional encoding conditions.

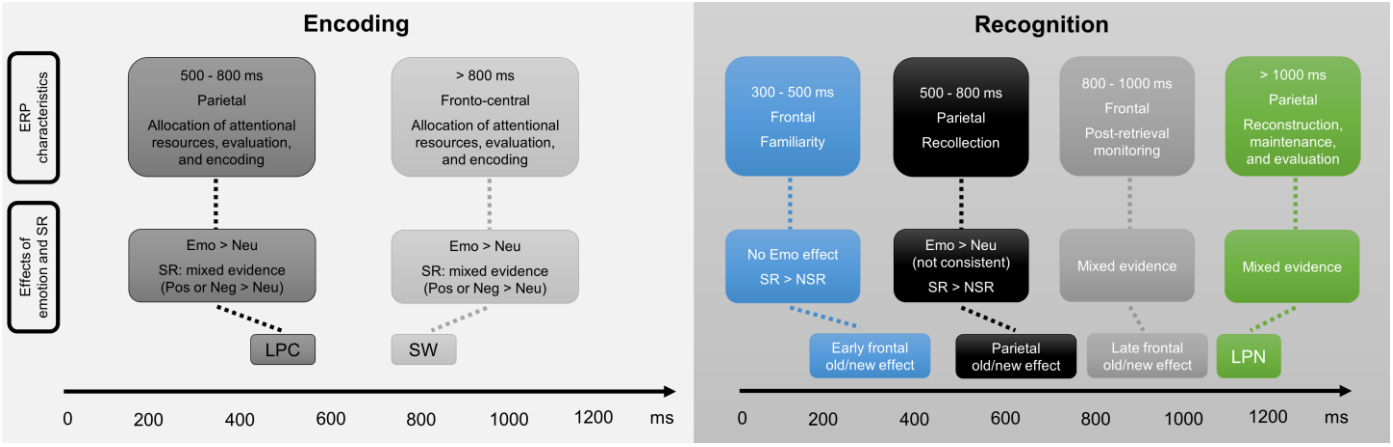


Figure 1. Schematic illustration of the event-related potentials of interest in the current study, including the main findings reported in the literature concerning the effects of emotion and self-reference on the amplitude of each component. Note: Emo = Emotion; ERP = Event-related potential; LPC = Late Positive Complex; LPN = Late Posterior Negativity; Neg = Negative; Neu = Neutral; NSR = Non-self-reference; Pos = Positive; SR = Self-reference; SW = Slow Wave.

Effects of emotion and self-reference on recognition memory: ERP evidence

In recognition memory tasks, the study phase is usually followed by a test phase in which the previously presented stimuli are repeated and intermixed with new items, and participants are invited to make old/new judgments. Given that the study of source memory involves testing memory for details, such as perceptual (e.g., stimulus color or background), contextual (e.g., temporal: list 1 *vs.* list 2; spatial:

location on the screen), semantic (e.g., organization of items into specific thematic units), affective (e.g., emotional reactions during the event), and cognitive operations (e.g., encoding tasks; verbal thoughts; mental images; Johnson et al., 1993), the source memory recognition test requires the discrimination between different details manipulated during encoding.

Differences in brain electrical activity derived from correctly recognized old items and correctly rejected new items were found to be associated with old/new effects characterized by distinct spatiotemporal distributions, polarities, and putative functional roles (see Voss & Paller, 2017; see Figure 1 for a schematic illustration). The first well-described component is the early frontal old/new effect (also known as frontal N400-like component: FN400), which is observed in the 300 - 500 ms latency window post-stimulus onset, with maximum voltage over frontal electrode sites, and characterized by an enhanced negativity for new compared to old items (Curran, 2000). This effect has been proposed as an index of familiarity, i.e., the impression that certain information was previously seen even though no further details from the study episode are available (Curran, 2000; Curran & Dien, 2003; Czernochowski, Mecklinger, Johansson, & Brinkmann, 2005; but see Voss & Federmeier, 2011 for a different proposal). In a subsequent time window (approximately 500 - 800 ms post-stimulus onset), the parietal old/new effect is typically observed, presenting maximum voltage over parietal electrodes: this positivity is enhanced for old compared to new items (Curran, 2000; Leynes & Phillips, 2008). This effect has been related to operations involving the recollection of information from the study episode, since its amplitude is sensitive not only to accurate source responses (*vs.* incorrect source responses and responses attributed to familiarity; Addante, Ranganath, & Yonelinas, 2012; Cansino & Trejo-Morales, 2008; Leynes, Askin, & Landau, 2017; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999; Wilding & Rugg, 1997), but also to the number of recoverable details (e.g., Leynes & Crawford, 2018; Leynes, Crawford, & Bink, 2005; Wilding, 2000; Woroch & Gonsalves, 2010).

The affective properties of target stimuli were found to modulate the parietal old/new effect more strongly than the early frontal old/new effect: whereas the parietal old/new effect seems to be characterized by an increased positivity in response to emotional (negative and/or positive) compared to neutral items (Dietrich et al., 2001; Langeslag & Van Strien, 2008; McNeely, Dywan, & Segalowitz, 2004; Weymar, Löw, & Hamm, 2011; Weymar, Löw, Melzig, & Hamm, 2009; Wirkner et al., 2018), the early frontal old/new effect seems to be relatively unaffected by emotion (Johansson, Mecklinger, & Treese, 2004; Koenig & Mecklinger, 2008; Lavoie & O'Connor, 2013; Maratos, Allan, & Rugg, 2000; Minor & Herzmann, 2019; Newsome, Dulas, & Duarte, 2012; Weymar, Löw, Melzig, & Hamm, 2009; but see Inaba, Nomura, & Ohira, 2005; Mao, You, Li, & Guo, 2015; Meng et al., 2017; Schaefer, Pottage, &

Rickart, 2011; Xu, Zhang, Li, & Guo, 2015). Nonetheless, emotion effects on the parietal old/new effect are not fully consistent across studies (lack of emotion effect: Lavoie & O'Connor, 2013; Maratos et al., 2000; Windmann & Kutas, 2001; enhanced amplitude for negative items: Inaba et al., 2005; Johansson, Mecklinger, & Treese, 2004; Newsome et al., 2012; Weymar et al., 2011). Also, when self-referential encoding conditions were tested, amplitude enhancements of both early frontal and parietal old/new effects have been documented in the context of source memory attributions (Mao et al., 2017; Newsome et al., 2011).

During recognition, two other later electrophysiological responses might emerge, especially when source discrimination is required: the late right frontal old/new effect and the late posterior negativity (LPN; e.g., Cycowicz & Friedman, 2003; Donaldson & Rugg, 1998; Johansson & Mecklinger, 2003; Mecklinger, Rosburg, & Johansson, 2016). The first is a positivity enhanced for old compared to new items, observed around 800 - 1000 ms post-stimulus onset (and lasting for several ms), with maximum voltage over right frontal electrode sites (Friedman & Johnson, 2000). Although its functional role remains to be specified, the late right frontal old/new effect has been related to post-retrieval monitoring processes that allow the evaluation and search of information retrieved from memory (Donaldson & Rugg, 1998; Leynes, 2002; Leynes, Grey, & Crawford, 2006; Trott et al., 1999; Van Petten, Senkfor, & Newberg, 2000; Wilding & Rugg, 1996). The LPN is a negative-going deflection that is observed approximately 1000 ms post-stimulus onset, with maximum voltage over Pz: its amplitude is enhanced for studied compared to unstudied items (Friedman, Cycowicz, & Bersick, 2005; Mecklinger, Johansson, Parra, & Hanslmayr, 2007; Mecklinger et al., 2016). The LPC has been associated with the reconstruction, maintenance and evaluation of item and contextual features of a prior learning episode, which may be required during difficult source memory decisions or when task-relevant details are not readily available (Johansson & Mecklinger, 2003; Leynes et al., 2006; Mecklinger et al., 2016; Wilding & Rugg, 1997).

With respect to the effects of emotion on these later components, the evidence is mixed and scarce. Whereas some studies reported no right frontal old/new effect for emotional stimuli (e.g., Maratos et al., 2000; Newsome et al., 2012), other studies documented an amplitude enhancement in response to emotional stimuli (e.g., Langeslag & Van Strien, 2008; McNeely et al., 2004; Xu et al., 2015). In the case of the LPN, some failed to show modulatory effects of emotion on its amplitude (Koenig & Mecklinger, 2008), whereas other studies suggested that this effect might emerge only under specific emotional conditions (see Newsome et al., 2012 for an example regarding negative stimuli). Taken together, the studies conducted so far demonstrate that the presence or absence of these old/new effects may depend upon the type of source details tested during recognition and on how emotion affects the memory of such

source-specifying details. Thereby, the modulatory effects of emotion and self-reference on the old/new effects in the context of source memory remain to be specified.

The current study and hypotheses

This study aimed to probe the interactive effects of self-reference (self-referential *vs.* non-self-referential encoding) and valence (negative *vs.* neutral *vs.* positive) on both recognition of internal source memory details (the encoding task used in the study phase) and processing of target words during encoding. Specifically, in the self-referential condition, participants were instructed to evaluate if the words described them, while in the non-self-referential condition participants evaluated if the words were perceived as being commonly used by other people in daily living (i.e., common condition).

At the behavioral level, we anticipated that source memory performance would be improved for both positive and neutral compared to negative words in the self-referential condition, following prior studies (e.g., Durbin et al., 2017; Leshikar et al., 2015; Pereira et al., 2019; Y. Zhang et al., 2018). In the case of the non-self-referential condition, source memory performance was expected to be enhanced for neutral compared to negative and positive words (e.g., Cook et al., 2007; Ferré et al., 2019, Experiment 2 and 3; Pereira et al., 2019, Experiment 2). Nonetheless, a lack of effect was also anticipated (e.g., Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006a; Sharot & Yonelinas, 2008; see also Pereira et al., 2019 for a summary table).

During encoding, we predicted that words studied in the self-referential condition would elicit larger LPC and SW mean amplitudes regardless of valence (Nowicka et al., 2018; Zhao et al., 2016), as self-referential encoding is expected to lead to greater elaboration due to its motivational relevance (Symons & Johnson, 1997). Furthermore, we also hypothesized an interaction of self-reference and valence: in the case of self-referential processing, larger LPC and SW amplitudes were anticipated in response to positive relative to negative and neutral words, which would be indicative of a processing bias toward positive information (Herbert et al., 2010, 2006; Shestyuk & Deldin, 2010); in the case of the non-self-referential condition, the amplitudes of those components were expected to be larger for emotional relative to neutral words (Barnacle et al., 2018; Dolcos & Cabeza, 2002; Palomba et al., 1997; Wirkner et al., 2018).

With respect to the old/new electrophysiological effects, we hypothesized distinct outcomes. First, typical early frontal old/new effects were expected for old items associated with accurate source memory responses (*vs.* correct rejections of new items), irrespective of self-reference and valence, considering prior evidence (Johansson et al., 2004; Koenig & Mecklinger, 2008; Lavoie & O'Connor, 2013; Maratos et al., 2000; Minor & Herzmann, 2019; Newsome et al., 2012; Weymar et al., 2009). Second, typical

parietal old/new effects were expected to be reduced in response to negative self-referential source hits compared to both positive and neutral self-referential source hits. This was specifically based on the assumption that the amount of retrieved details might be lower for negative words as participants tend to disregard self-referential negative information and to show a self-positivity bias (e.g., D'Argembeau & Van der Linden, 2008; Derry & Kuiper, 1981; Lewis, Critchley, Rotshtein, & Dolan, 2007; Moran, Macrae, Heatherton, Wyland, & Kelley, 2006; Pauly, Finkelmeyer, Schneider, & Habel, 2013). No valence-related modulations were expected for the amplitude of source hits in the non-self-referential condition as the amount of source features retrieved, even if qualitatively distinct, may not vary as much across valence categories. Due to the use of intentional encoding instructions, emotional and neutral information may receive similar relevance and elaborative resources (Barnacle et al., 2018), thus being similarly prioritized to meet the task demands (see the arousal-biased competition account by Mather & Sutherland, 2011). Third, based on the functional roles of the late right frontal old/new effect and the LPN (e.g., post-retrieval; monitoring; evaluative processes; Johansson & Mecklinger, 2003; Leynes et al., 2006; Mecklinger et al., 2016; Rugg, Allan, & Birch, 2000), differences between experimental conditions were expected to emerge for those more likely associated with difficult source memory decisions, such as words studied in the non-self-referential condition or negative words studied in the self-referential condition. Additionally, such differences were expected to be more pronounced in the late right frontal old/new effect- than in the LPN-time window. Due to the source manipulation adopted here, it is likely that the details of the study episode recovered during the test phase are diagnostic and relevant to source discrimination, resulting in a reduced mobilization of reconstructive processes typically related to the LPN (Johansson & Mecklinger, 2003; Mecklinger et al., 2016).

Methods

Participants

Twenty-two right-handed European Portuguese speakers (19 females) aged between 19 and 37 (M = 25.81, SD = 6.35) years, with an average of 16.14 years of formal education participated in the study. All participants had normal or corrected-to-normal vision and hearing. Also, they reported no current psychiatric or neurological conditions, no history of drug abuse and no psychotropic medication. To rule out the presence of mood disorders, participants completed the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996; Coelho, Martins, & Barros, 2002; M = 3.68, SD = 5.58), and the State-Trait Anxiety Inventory - Form Y (STAI-Y; Silva & Spielberger, 2007; state subscale: M = 27.64, SD = 5.79; trait subscale: M = 28.95, SD = 9.07). Participants provided written informed consent prior to their enrollment

in the study. The study was approved by a local Ethics committee (University of Minho, Braga, Portugal; SECVS 105/2016).

Stimuli

Stimuli were 360 European Portuguese (EP) adjectives (120 neutral, 120 positive, 120 negative), selected from the EP version of the Affective Norms for English Words (ANEW; Bradley & Lang, 1999; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012, and from a pilot study (n = 125 EP native speakers) that aimed to expand the ANEW adjectives set (see Supplementary Table S1). Details regarding the pilot study were described elsewhere (Pereira, Sampaio, & Pinheiro, 2020). Adjectives differed in valence (negative < neutral < positive; all $p < .001$) but not in arousal ($p = .12$). The decision to keep arousal at medium levels across valence categories followed prior literature showing that ERPs obtained during encoding and recognition may be differentially modulated by valence and arousal (e.g., Bayer et al., 2012; Delaney-Busch, Wilkie, & Kuperberg, 2016; Meng et al., 2017; Xu et al., 2015). For instance, an inverted-U relationship was found between the amplitude of the parietal old/new effect and both stimulus valence (Schaefer, Fletcher, Pottage, Alexander, & Brown, 2009) and arousal (Schaefer et al., 2011). Thus, we ensured that only stimulus valence differed across conditions following previous studies (e.g., Gallant et al., 2019; Gallant & Yang, 2014). The words in the three valence conditions were also matched for frequency ($p = .20$), number of letters ($p = .83$) and number of syllables ($p = .61$) (see Supplementary Table S1). Stimuli were randomly divided into 10 lists to create five different study-test cycles. These lists did not differ in relation to mean valence and arousal ratings, frequency, number of letters, and number of syllables ($p > .05$).

Procedure

The experiment took place in a sound-attenuated, dimly lit, electrically shielded room. Participants were instructed that they were about to perform a memory task with words, but first they had to perform two-choice judgments for each stimulus. For some words, they had to evaluate if the word described or related to them in any way by responding “yes” or “no” (self-referential condition). For other words, they had to evaluate if the words were commonly used by other people in their everyday lives, also by responding “yes” or “no” (non-self-referential or common condition). Participants were informed that the self-reference and the common judgment tasks would appear in a mixed fashion. It was also emphasized that the words and how they were evaluated should be memorized for a later memory test (intentional encoding). Moreover, after each word, they were required to rate in a 6-point scale how likely they were

to remember in a later memory test that the study word was evaluated according to the self-reference or the common judgment task (judgments of source - JOSs: 1 = “Sure I will not remember the judgment”; 6 = “Sure I will remember the judgment”). In the memory test, the studied words were intermixed with new words and participants had to select one of four options: “self-description”, “common”, “evaluated, but do not know if self-description/common” or “new”. In both phases, participants were also instructed to respond as fast as possible. During the test phase, response accuracy was highlighted. Participants performed two short study-test cycles with words not used in the main experiment to become familiarized with the task requirements.

The experimental task was composed of five study-test cycles, which were counterbalanced and randomized across participants. Additionally, the word lists assigned to each study and test phase were counterbalanced across participants so that the words used during the study phase for one participant could be new words during the test phase for another participant. Each study phase included 48 trials, half required a self-referential judgment, and the other half required a common judgment. The assignment of each word to one of those tasks was also counterbalanced across participants so that the same word could be evaluated following the common judgment for one participant and the self-referential judgment for another participant. For each encoding task, eight words were negative, eight were neutral, and eight were positive. The words from each experimental condition – self-negative, self-neutral, self-positive, common-negative, common-neutral, common-positive – were intermixed and randomly presented, one by one, in the center of the computer screen (white Arial; font size 88; black background). The structure of an experimental trial and associated details are presented in Figure 2.

Each test phase was composed of 72 trials: 48 studied words randomly mixed with 24 new words (eight from each valence category). Ideally, the number of old and new items should be balanced due to effects of differences in proportion of old/new items on behavioral and brain measures (e.g., Vilberg & Rugg, 2009). However, due to several constraints (e.g., limitations of the available EP pool of stimuli; length of the recognition phase and overall experiment) and since we were particularly interested in the modulatory role of valence on successful self-referential source memory attributions, we decided to increase the signal-to-noise ratio in favor of old items, following previous studies (Mao et al., 2015; Minor & Herzmann, 2019; Newsome et al., 2012). The structure of an experimental trial in the test phase is illustrated in Figure 1. Each trial was composed of a fixation cross (500 - 650 ms), a black blank screen (1000 ms), the old/new word (2000 ms), and a response screen with the prompt “?” containing four response options at the bottom of the screen (“self-description”; “common”; “evaluated, but do not know if self-description/common”; “new”; see Figure 1). Of note, the “do not know” option was provided to

mitigate possible response contaminations due to guesses and biases (e.g., Addante et al., 2012; Dulas & Duarte, 2013; Minor & Herzmann, 2019; Newsome et al., 2012; Yin et al., 2019), which is ideal to isolate the electrophysiological activity underlying accurate source memory responses (e.g., Leynes et al., 2017; Wilding, 2000; Wilding & Rugg, 1997).

At the end of the experiment, participants were debriefed about their performance and the implementation of specific encoding strategies. Overall, the experimental session lasted for approximately two hours. Presentation® software (Version 18.3, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) was used to control stimulus presentation.

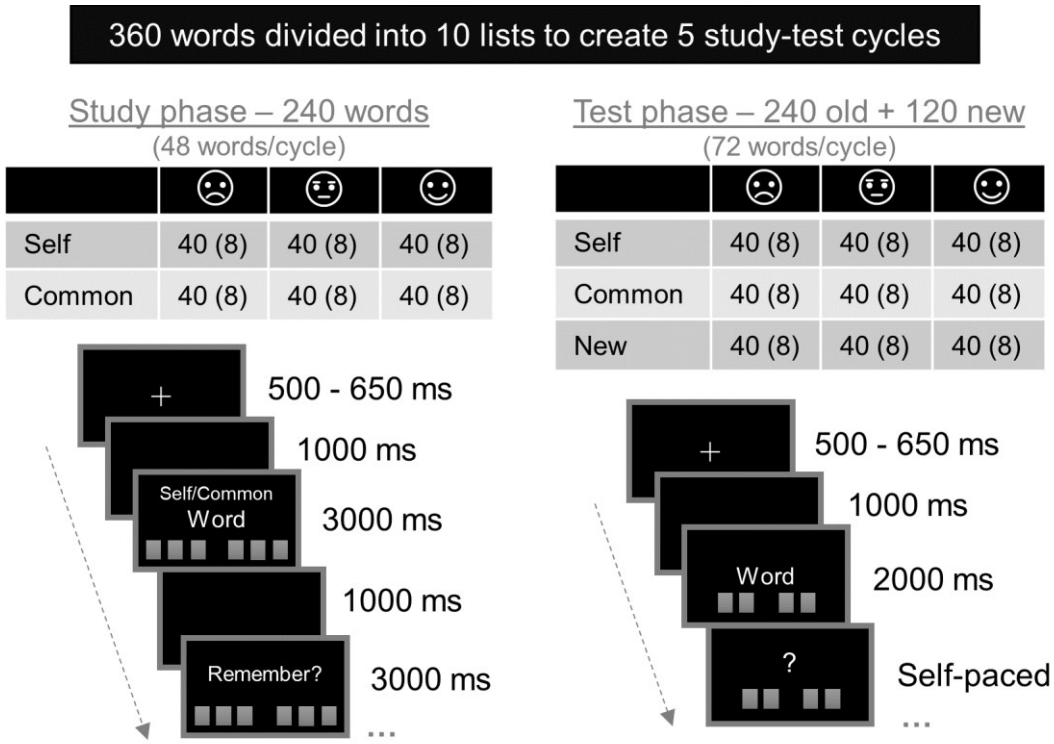


Figure 2. Schematic illustration of an experimental trial from the study and test phases. The number and type of stimuli regarding each encoding task (self *vs*. common) x valence (negative *vs*. neutral *vs*. positive) condition are also summarized. Note: During the study phase, each trial began with a fixation cross (500 - 650 ms) followed by a blank screen (1000 ms). Then, the word appeared together with an instruction at the top of the screen, which informed about the ongoing judgment task (3000 ms). At the bottom of the same screen six small rectangles were shown representing the position of the fingers required for the study phase. Particularly, participants were asked to place the left ring, middle, index fingers on the keyboard keys "Z"(1), "X"(2), "C"(3), and the right ring, middle, index fingers on the keyboard keys "B"(4), "N"(5), "M"(6). These keys were covered with stickers numbered from 1 to 6 to facilitate finger-key mapping and the response during the JOSs ratings, and to minimize cognitive demands. The "yes"/"no" responses were restricted to the index fingers and counterbalanced across participants. Another blank screen was introduced (1500 ms) followed by the JOSs rating which was prompted by the word "remember?" (3000 ms). In the test phase, each trial was composed of a fixation cross (500 - 650 ms), a black blank screen (1000 ms), the old/new word (2000 ms), and a response screen with the prompt "?" containing four response options at the bottom of the screen ("self-description"; "common";

"evaluated, but do not know if self-description/common"; "new"). At this stage, participants used four keys/fingers: "X"(2; left middle), "C"(3; left index), "B"(4; right index), "N"(5; right middle). The finger-response mapping was counterbalanced across participants.

ERP data acquisition and analysis

EEG data were recorded continuously at a digitization rate of 512 Hz using a 64-channel BioSemi Active Two system (Biosemi, Amsterdam, Netherlands; http://www.biosemi.com/products.htm) and stored on hard disk for later analysis. Additional five external electrodes were used: two were placed on the left and right mastoids (reference); two electrodes were placed on the outer canthi of the eyes to register the horizontal electrooculogram (HEOG) and one was placed below the left to record the vertical electrooculogram (VEOG). Electrode offset was kept below 30 mV. The EEG was recorded during the study and test phases of all five experimental blocks.

The offline EEG processing was conducted using two free-open source MATLAB packages suited for EEG/ERP data analysis, the EEGLAB toolbox version 13.5.4b (Delorme & Makeig, 2004; http://sccn.ucsd.edu/eeglab/) and the ERPLAB Toolbox version 7.0.0 (Lopez-Calderon & Luck, 2014; http://erpinfo.org/erplab/). Data were first band-pass filtered at 0.1 - 40 Hz using an IIR Butterworth filter (12 dB/octave roll-of slope) and referenced to the averaged of the right and left mastoids. An independent component analysis (ICA; runica algorithm) was run to detect and correct eye-related artifacts. To guide the rejection of components, three EEGLAB data processing extensions were used: SASICA (Chaumon, Bishop, & Busch, 2015; https://github.com/dnacombo/SASICA), ADJUST (Mognon, Jovicich, Bruzzone, & Buiatti, 2011; https://www.nitrc.org/docman/view.php/739/2101/TutorialADJUST1_1.pdf), and ICLabel (Pion-Tonachini, Makeig, & Kreutz-Delgado, 2017; https://labeling.ucsd.edu/tutorial/overview). On average, 2.82 components (*SD* = 0.66; range 2-4) were removed across participants. Noisy channels were interpolated using the spherical interpolation method (*M* = 3.95; *SD* = 1.29; range 0-5). Event-based epochs were created including a baseline of 200 ms prior to stimulus onset and 1200 ms (word encoding) or 2000 ms (recognition phase) post-stimulus onset. All epochs were baseline corrected considering the 200 ms period prior to stimulus onset. Epochs with ± 100μV amplitude were discarded from the analysis prior to averaging.

Following prior studies (Barnacle et al., 2018; Dolcos & Cabeza, 2002; Fischler & Bradley, 2006; Gallant et al., 2018; Herbert et al., 2006; Koenig & Mecklinger, 2008), mean amplitude (μV) during word encoding was extracted in the following time windows: 500 – 700 (LPC); 700 – 950; 950 – 1200 ms (SWs). We were only able to run the comparison between experimental conditions for general encoding

due to the insufficient number of trials associated with incorrect responses or hits in some participants (see Supplementary Table S2).

Regarding the recognition stage, the mean amplitude of well-described ERP components was extracted for both source hits and correct rejections of new items (e.g., Cansino, Hernández-Ramos, & Trejo-Morales, 2012; Dulas et al., 2011; Koenig & Mecklinger, 2008; Mao et al., 2015; Newsome et al., 2012; Wilding, 1999): early frontal old/new effect (300 – 500 ms), parietal old/new effect (500 – 800 ms), right frontal old/new effect (800 – 1200; 1200 – 1800 ms), LPN (800 – 1200; 1200 – 1800 ms). Again, the ERP responses related to incorrect responses were not examined further due to the low number of trials. Mean amplitude was extracted from distinct regions-of-interest (ROIs) covering anterior to posterior and left to right locations, following previous studies (e.g., Dulas et al., 2011; Langeslag & Van Strien, 2008): left frontal (F3/F5), middle frontal (F1/Fz/F2), right frontal (F4/F6), left fronto-central (FC3/FC5), middle fronto-central (FC1/FCz/FC2), right fronto-central (FC4/FC6), left central (C3/C5), middle central (C1/Cz/C2), right central (C4/C6), left centro-parietal (CP3/CP5), middle centro-parietal (CP1/CPz/CP2), right centro-parietal (CP4/CP6), left parietal (P3/P5), middle parietal (P1/Pz/P2), right parietal (P4/P6).

Statistical analysis

Behavioral data.

Following Snodgrass and Corwin (1988), the recognition measure Pr = ([p(hits) – p(false alarms)]) was used as an index of item recognition accuracy, wherein p(hits) was obtained by collapsing the "do not know" responses with both correct and incorrect source responses. To index response bias, the Br measure was used - Br = ([p(false alarms)/(1 – Pr)]) - wherein values above 0.50 indicate a more liberal response criterion, i.e., in cases of uncertainty the decision is biased to choose "old"; whilst values below 0.50 indicate a more conservative response criterion, that is, the decision is biased to select "new" in the face of uncertainty. Regarding source memory recognition, we computed the difference between source hits and incorrect source responses: [p(correct source) – p (incorrect source)] (Dulas & Duarte, 2012, 2014; Hou et al., 2019; Leshikar et al., 2015; Newsome et al., 2012). Data were submitted to a 2 (self-reference: self *vs.* common) x 3 (valence: negative *vs.* neutral *vs.* positive) repeated-measures analysis of variance (RMANOVA).

ERP data.

Mean amplitudes during encoding were subjected to a 2 (self-reference: self *vs.* common) x 3 (valence: negative *vs.* neutral *vs.* positive) x 5 (ROI: frontal *vs.* fronto-central *vs.* central *vs.* centro-parietal *vs.* parietal) x 3 (laterality: left *vs.* middle *vs.* right) RMANOVA. In the recognition phase, a 3 (self *vs.* common *vs.* new) x 3 (valence) x 5 (ROI) x 3 (laterality) RMANOVA was computed. Statistically significant effects were described and followed up if they included the experimental conditions (self-reference and/or valence). For all the RMANOVAS, Bonferroni-corrected post-hoc tests were run to qualify main and interaction effects, and the Greenhouse-Geisser correction was considered when violations of sphericity were present. Partial eta squared (η_p^2) was used as a measure of effect size[3].

Results

Table 1 presents a summary of the main descriptive statistics regarding the behavioral performance.

Behavioral results (recognition accuracy)

Item memory.

The 2 x 3 RMANOVA revealed a main effect of self-reference, $F(1, 21) = 44.68$, $p < .001$, $\eta_p^2 = .68$, and valence, $F(2, 42) = 10.49$, $p < .001$, $\eta_p^2 = .33$, as well as an interaction effect between factors, $F(2, 42) = 7.98$, $p = .001$, $\eta_p^2 = .28$. Specifically, the interaction revealed that, regardless of valence, there was an item memory benefit for words studied self-referentially when compared to words studied in the common condition (negative: $p = .019$, 95% CI_{diff} [0.01, 0.09]; neutral: $p < .001$, 95% CI_{diff} [0.10, 0.20]; positive: $p < .001$, 95% CI_{diff} [0.08, 0.16]). Whereas no differences in item recognition were observed between negative ($M = 0.59$, $SE = 0.03$), neutral ($M = 0.63$, $SE = 0.03$), and positive words ($M = 0.59$, $SE = 0.02$) in the common condition, there were statistically significant differences between all valence categories for words in the self-referential condition: neutral ($M = 0.78$, $SE = 0.03$; $p < .001$, 95% CI_{diff} [0.09, 0.18]) and positive words ($M = 0.71$, $SE = 0.03$; $p = .019$, 95% CI_{diff} [0.01, 0.12]) were better recognized than negative words ($M = 0.64$, $SE = 0.03$), and neutral words were also better recognized than positive words ($p = .014$, 95% CI_{diff} [0.01, 0.13]; see Figure 3A).

[3] In the Supplementary Material section, we present additional analyses, namely: implementation of a multinomial model approach considering item and source memory recognition; one sample *t*-tests computed for item and source recognition measures; differences between experimental conditions with respect to the proportion of misses, false alarms, correct rejections, "do not know" responses, incorrect source responses, "yes" responses, and mean response times collected during the word encoding and JOSs ratings; mean JOSs ratings, electrophysiological responses, and Goodman-Kruskal gamma correlations computed for each experimental condition.

Table 1

Descriptive Statistics of the Behavioral Performance as a Function of Encoding Task and Valence

	Encoding task / Valence					
	Self – *M* (*SD*)			Common – *M* (*SD*)		
Measure	Negative	Neutral	Positive	Negative	Neutral	Positive
Source hit	0.60 (0.15)	0.74 (0.13)	0.75 (0.14)	0.55 (0.14)	0.59 (0.16)	0.53 (0.12)
Incorrect source	0.12 (0.08)	0.10 (0.07)	0.08 (0.07)	0.10 (0.07)	0.06 (0.04)	0.12 (0.09)
DNK source	0.11 (0.07)	0.06 (0.04)	0.06 (0.05)	0.13 (0.09)	0.10 (0.06)	0.12 (0.07)
Source miss	0.17 (0.10)	0.10 (0.09)	0.11 (0.09)	0.22 (0.11)	0.25 (0.14)	0.23 (0.10)
Item Pr measure	0.64 (0.12)	0.78 (0.12)	0.71 (0.13)	0.59 (0.13)	0.63 (0.15)	0.59 (0.11)
Item Br measure	0.54 (0.21)	0.54 (0.30)	0.64 (0.28)	0.48 (0.20)	0.34 (0.21)	0.45 (0.24)
SM measure	0.48 (0.21)	0.64 (0.18)	0.67 (0.20)	0.46 (0.18)	0.53 (0.18)	0.41 (0.19)
JOSs ratings	3.95 (0.77)	3.98 (0.73)	4.34 (0.76)	3.61 (0.74)	3.72 (0.77)	3.82 (0.77)
Gamma (JOSs)	0.25 (0.38)	0.32 (0.45)	0.18 (0.41)	0.13 (0.44)	0.03 (0.39)	0.13 (0.50)
RT - Word encoding	1772 (341)	1932 (277)	1770 (236)	1804 (259)	1793 (275)	1767 (309)
RT - JOSs	928 (251)	928 (295)	844 (273)	903 (257)	923 (223)	903 (260)

	New negative	New neutral	New positive
CR	0.81 (0.10)	0.88 (0.08)	0.81 (0.12)
FA	0.19 (0.10)	0.12 (0.08)	0.19 (0.12)

Note. CR = Correct Rejections; DNK = Do Not Know; FA = False Alarms; JOSs = Judgments of Source; RT = Response Time; SM = Source Memory. The gamma (JOSs) refers to the Goodman-Kruskal gamma correlations that inform about the degree of association between predictions during encoding and objective memory performance during retrieval (see Supplementary Materials section for more details).

In the case of the response bias measures, the 2 x 3 RMANOVA showed a main effect of self-reference, $F(1, 21) = 27.05$, $p < .001$, $\eta_p^2 = .56$, and valence, $F(2, 42) = 4.98$, $p = .019$, $\eta_p^2 = .19$, $\varepsilon = .79$, as well as an interaction effect, $F(2, 42) = 5.64$, $p = .007$, $\eta_p^2 = .21$. The Br value was higher in the self-referential condition than in the common condition for neutral ($p < .001$, 95% CI_{diff} [0.11, 0.30]) and positive words ($p < .001$, 95% CI_{diff} [0.10, 0.28]), but not for negative words ($p = .064$, 95% CI_{diff} [- 0.004, 0.13]). This indicates that for both neutral and positive words participants used a more conservative response criterion in the common condition than in the self-referential condition. In fact, on average, the Br values were above 0.50 in the self-referential condition, which is indicative of a more liberal response criterion, whilst the same was not observed in the common condition (see Table 1; Snodgrass & Corwin, 1988). Regarding the self-referential condition, the response bias was higher for positive ($M = 0.64$, $SE = 0.06$) than for negative words ($M = 0.54$, $SE = 0.04$; $p = .046$, 95% CI_{diff} [0.001, 0.19]). For the common condition, the response criterion was more conservative for neutral words ($M = 0.34$, $SE = 0.05$) relative to both negative ($M = 0.48$, $SE = 0.04$; $p = .004$, 95% CI_{diff} [0.04, 0.25]) and positive words ($M = 0.45$, $SE = 0.05$; $p = .031$, 95% CI_{diff} [0.01, 0.21]).

Source memory.

The main effects and the interaction effect were statistically significant [self-reference: $F(1, 21) = 13.34$, $p = .001$, $\eta_p^2 = .39$; valence: $F(2, 42) = 10.71$, $p < .001$, $\eta_p^2 = .34$; self-reference x valence: $F(2, 42) = 10.62$, $p < .001$, $\eta_p^2 = .34$]. The planned comparisons revealed that for neutral ($p = .005$, 95% CI_{diff} [0.04, 0.19]) and positive words ($p < .001$, 95% CI_{diff} [0.15, 0.36]), source memory recognition was improved in the self-referential as compared to the common condition (see Table 1). Furthermore, whereas self-referential processing benefited source memory for neutral ($M = 0.64$, $SE = 0.04$; $p < .001$, 95% CI_{diff} [0.07, 0.25]) and positive words ($M = 0.67$, $SE = 0.04$; $p < .001$, 95% CI_{diff} [0.09, 0.28]) compared to negative words ($M = 0.48$, $SE = 0.04$), the source memory for words studied in the common condition was improved when words had a neutral ($M = 0.53$, $SE = 0.04$) compared to a positive valence ($M = 0.41$, $SE = 0.04$; $p = .007$, 95% CI_{diff} [0.03, 0.21]; see Figure 3B), with no differences in the case of negative words ($M = 0.46$, $SE = 0.04$).

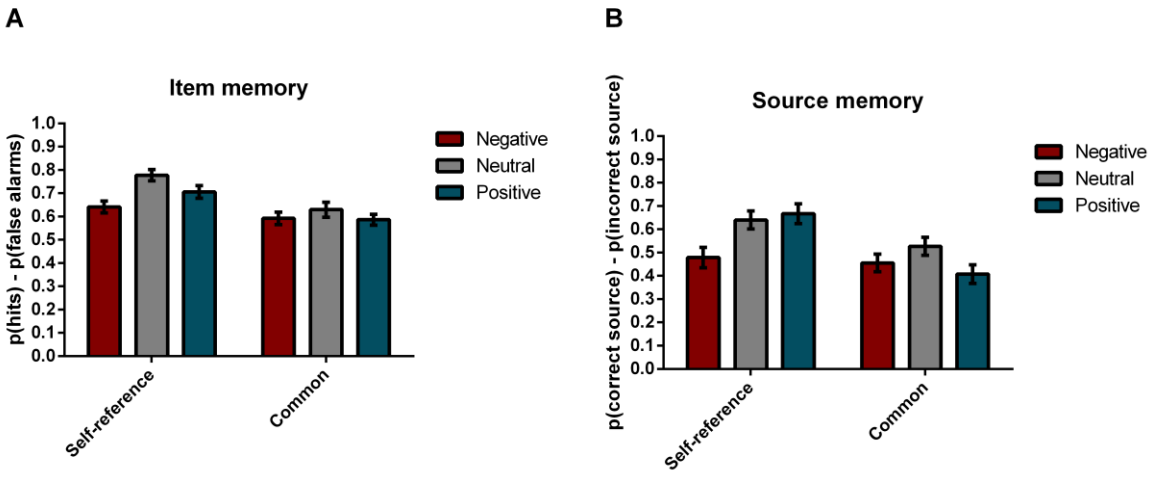


Figure 3. Behavioral results for item memory and source memory are shown in panel A and panel B, respectively. The recognition scores are plotted on the y-axis as a function of encoding task (self-reference *vs*. common) and valence (negative *vs*. neutral *vs*. positive). The error bars represent the standard error of the mean.

ERP results

Word encoding.

500 - 700 ms. No self-reference- and/or valence-related modulations were found in this time window (see Figure 4).

700 - 950 and 950 - 1200 ms. In the 700 - 950 ms latency window, two interaction effects including self-reference were observed: self-reference x ROI: $F(4, 84) = 5.45$, $p = .005$, $\eta_p^2 = .21$, $\varepsilon = .57$; self-reference x ROI x laterality: $F(8, 168) = 3.24$, $p = .002$, $\eta_p^2 = .13$. Specifically, the mean amplitude obtained for words studied self-referentially was more positive-going than the mean amplitude for words studied in the common condition irrespective of valence (all $p < .05$; see Figure 4) and for most ROIs (exceptions were detected in the left and middle parietal locations). In the 950 - 1200 ms latency window, similar interaction effects emerged [self-reference x ROI: $F(4, 84) = 21.73$, $p < .001$, $\eta_p^2 = .51$, $\varepsilon = .43$; x ROI x laterality: $F(8, 168) = 3.71$, $p = .001$, $\eta_p^2 = .15$]. The mean amplitude for words studied in the self-referential condition was again more positive than for words studied in the common condition for most ROIs (all $p < .05$; exceptions were detected in right centro-parietal and all parietal electrode sites; see Figure 4).

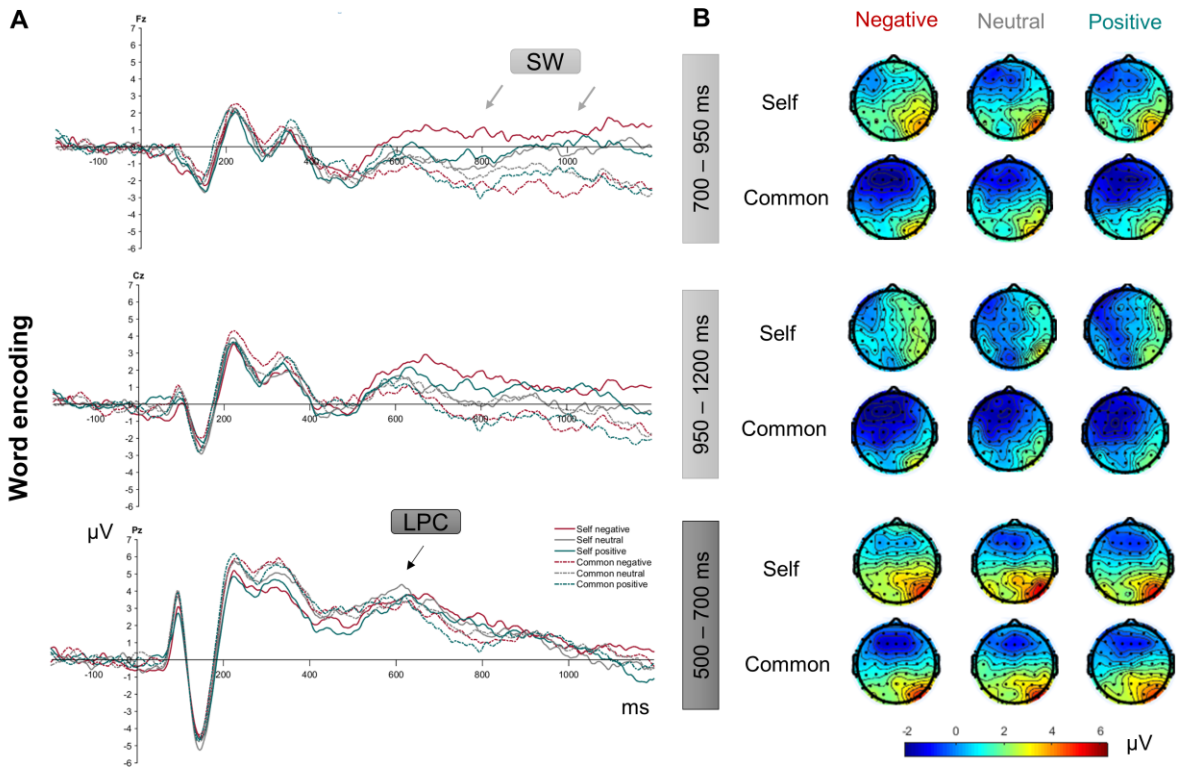


Figure 4. Grand average waveforms obtained during word encoding (panel A) together with their respective topographic maps (panel B) considering each encoding task (self *vs*. common) x valence (negative *vs*. neutral *vs*. positive) condition. Note: LPC = Late Positive Complex; SW = Slow Wave.

Test phase.

300 - 500 ms. The analysis considering self-source hits, common-source hits and correct rejections yielded a main effect of self-reference: $F(2, 42) = 5.22$, $p = .009$, $\eta_p^2 = .20$. The mean amplitude for self-source hits ($M = 1.51$, $SE = 0.57$) was larger than for correct rejections ($M = 0.51$, $SE = 0.75$; $p = .014$, 95% CI_{diff} [0.17, 1.82]), whereas the mean amplitude associated with common-source hits did not differ from the mean amplitude of both self-source hits and correct rejections ($p > .05$). Thereby, this early time window showed a general differentiation between old self-referential and new words, characterized by a more positive-going amplitude for old compared to new items (see Figure 5).

500 - 800 ms. The analysis yielded a self-reference x valence x laterality interaction, $F(8, 168) = 2.89$, $p = .005$, $\eta_p^2 = .12$. The follow-up of the three-way interaction indicated that, in the case of negative words, the mean amplitude for self-source hits (left: $M = 3.93$, $SE = 0.68$; middle: $M = 4.08$, $SE = 0.99$; right: $M = 3.90$, $SE = 0.89$) was larger than the mean amplitude for correct rejections irrespective of laterality (left: $M = 2.33$, $SE = 0.68$; middle: $M = 2.30$, $SE = 0.98$; right: $M = 2.39$, $SE = 0.79$; all $p < .05$). Considering neutral words, the mean amplitude for both self- and common-source hits (self left: $M = 4.40$, $SE = 0.69$; self middle: $M = 4.26$, $SE = 1.03$; common left: $M = 3.77$, $SE = 0.61$; common

middle: *M* = 3.87, *SE* = 0.96) was larger than for correct rejections over left and middle electrode sites (new left: *M* = 2.37, *SE* = 0.65; new middle: *M* = 2.36, *SE* = 0.98; all $p < .05$). In the case of positive words, the mean amplitude for both self- (*M* = 3.51, *SE* = 0.79) and common-source hits (*M* = 3.95, *SE* = 1.02) was larger than for correct rejections over right electrode sites (*M* = 2.09, *SE* = 0.79), whereas only the mean amplitude for common-source hits (left: *M* = 4.12, *SE* = 0.91; middle: *M* = 4.45, *SE* = 1.23) was larger than for correct rejections over left and middle electrode sites (left: *M* = 2.51, *SE* = 0.66; middle: *M* = 2.29, *SE* = 0.93; all $p < .05$; see Figure 5).

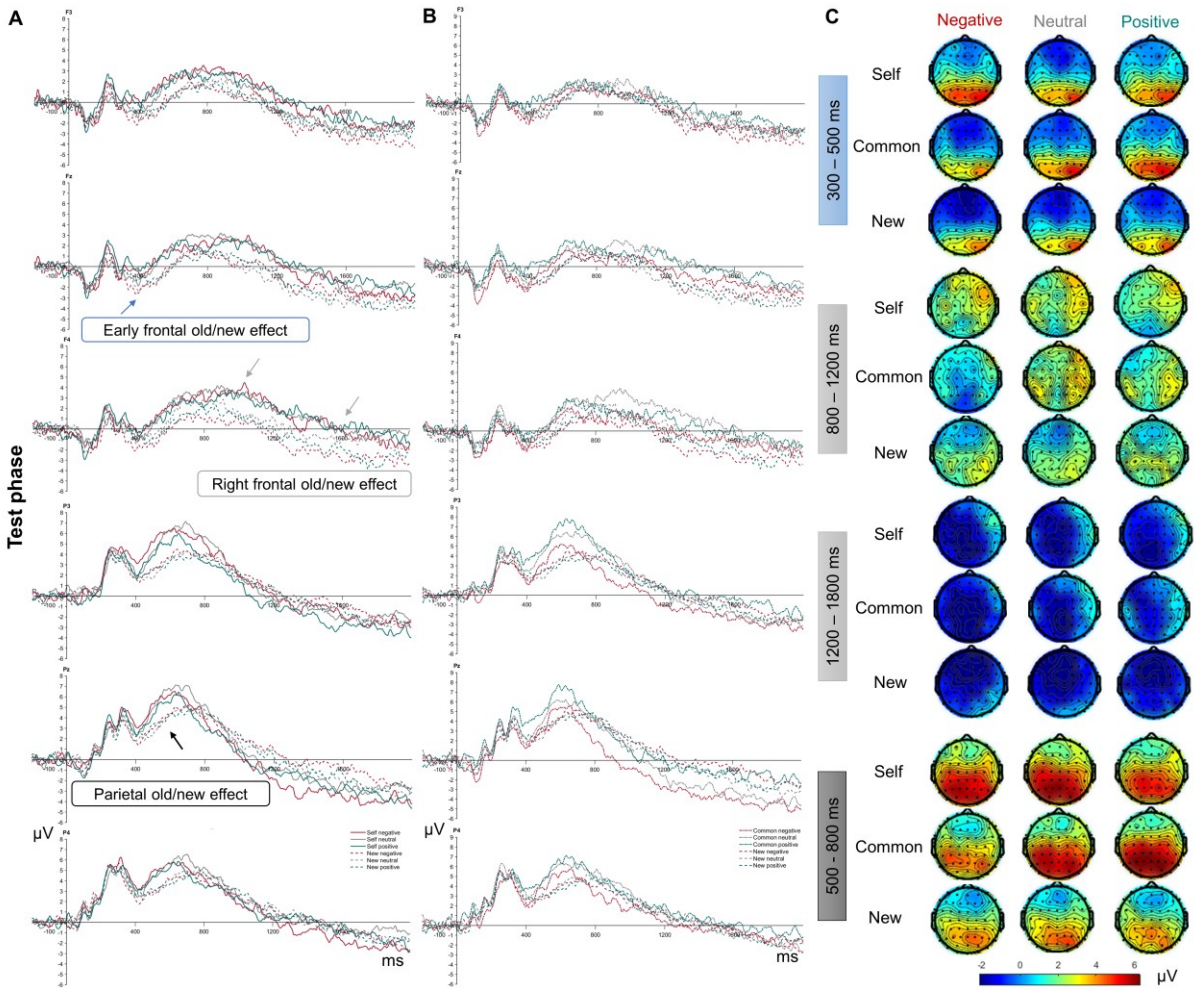


Figure 5. Grand average waveforms and topographical maps obtained during the test phase considering the four-time windows analyzed. The panel A includes self-source hits (old items) and correct rejections (new items) as a function of valence (negative *vs.* neutral *vs.* positive). Likewise, the panel B includes common-source hits and correct rejections. In the panel C, the topographical maps are presented as function of self-source hits *vs.* common-source hits *vs.* correct rejections and valence.

800 - 1200 ms. Two interaction effects including the experimental factors of interest were observed: self-reference x valence x laterality: $F(8, 168) = 2.00$, $p = .049$, $\eta_p^2 = .09$; self-reference x ROI x

laterality: $F(16, 336) = 3.33$, $p < .001$, $\eta_p^2 = .14$. When considering the self-reference x valence x laterality interaction, post-hoc tests showed that, over middle electrode sites, the mean amplitude for negative correct rejections ($M = 1.38$, $SE = 0.84$) and negative self-source hits ($M = 2.05$, $SE = 1.01$) was larger than the mean amplitude for negative common-source hits ($M = 0.33$, $SE = 0.95$; all $p < .05$). At right electrode sites, only the difference between negative common- ($M = 1.42$, $SE = 0.85$) and self-source hits ($M = 3.02$, $SE = 0.86$) reached statistical significance ($p = .031$); additionally, considering the same electrode sites and common-source hits, the mean amplitude for neutral words ($M = 3.24$, $SE = 0.81$) was larger than for negative words ($M = 1.42$, $SE = 0.85$; $p = .030$; see Figure 5). When decomposing the self-reference x ROI x laterality effect, a typical right frontal old/new effect was found, showing a larger mean amplitude for both self- ($M = 3.35$, $SE = 0.72$) and common-source hits ($M = 2.49$, $SE = 0.87$) when compared to correct rejections ($M = 1.27$, $SE = 0.86$; all $p < .05$; see Figure 5).

1200 - 1800 ms. In this time window, a significant self-reference x ROI x laterality interaction effect was observed. The decomposition of the interaction revealed a typical right frontal old/new effect, which was detected over middle/right frontal and right fronto-central sites, showing that the mean amplitude for correct rejections (middle frontal: $M = -2.65$, $SE = 0.96$; right frontal: $M = -1.31$, $SE = 0.94$; right fronto-central: $M = -1.22$, $SE = 0.82$) was less positive than the mean amplitude for both self- (middle frontal: $M = -0.94$, $SE = 0.87$; right frontal: $M = 0.90$, $SE = 0.70$; right fronto-central: $M = 0.80$, $SE = 0.89$) and common-source hits (middle frontal: $M = -0.73$, $SE = 1.00$; right frontal: $M = 0.45$, $SE = 0.92$; right fronto-central: $M = 0.28$, $SE = 0.85$; all $p < .05$; see Figure 5).

Correlational analyses

Given that the main ERP modulations associated with the encoding task and valence during recognition occurred in the 500 – 800 ms and 800 – 1200 ms latency windows, we conducted an exploratory correlational analysis between the amplitude and the proportion of correct source memory responses for each experimental condition. These correlations were employed with the aim of aiding with the interpretation of results and they are depicted in Figure 6. After applying Bonferroni correction for multiple comparisons ($p \leq .004$), only a statistically significant negative correlation between the mean amplitude for neutral words studied self-referentially and the proportion of correct source responses was found ($r = -.59$, $p = .004$). No other correlations reached statistical significance ($p > .004$).

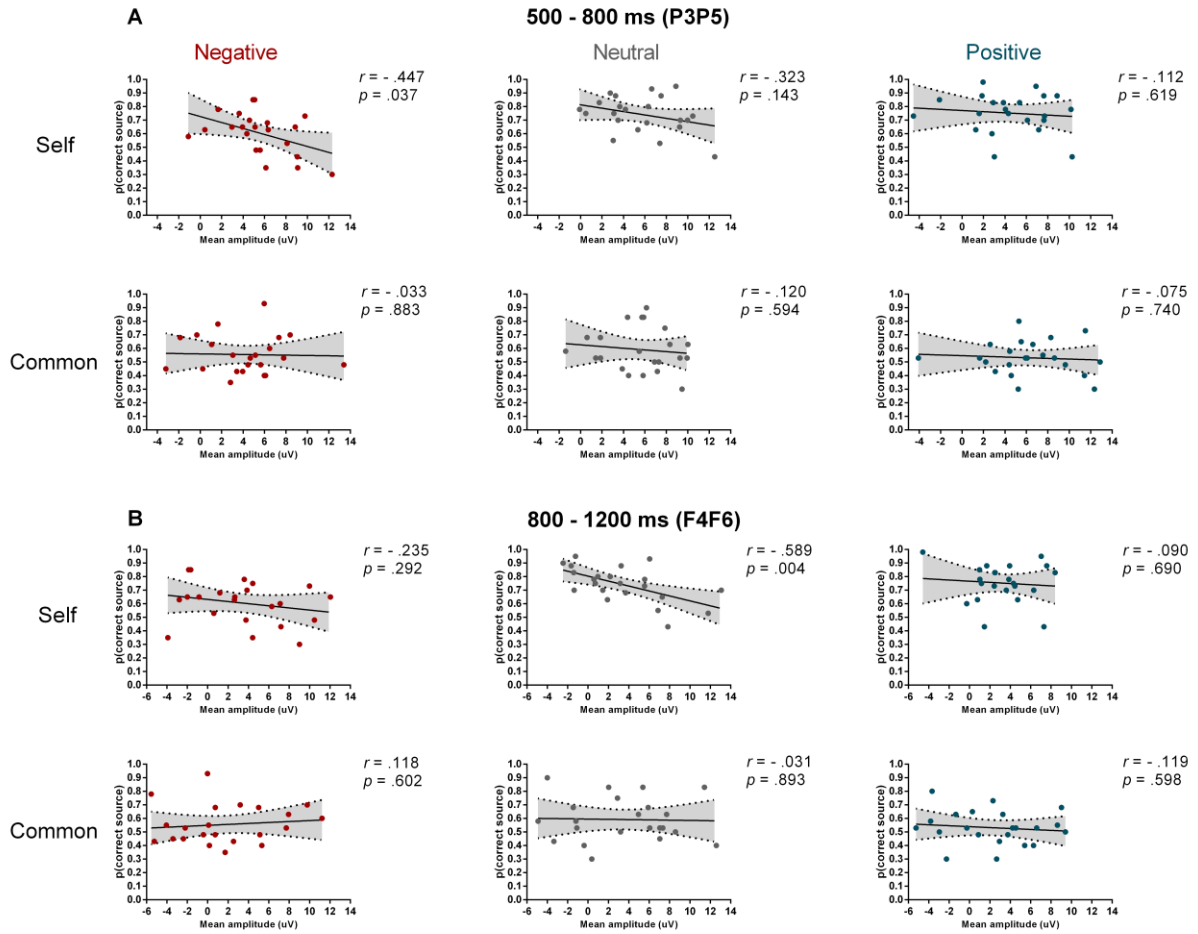


Figure 6. Correlation analyses between the mean amplitude obtained in the 500 – 800 ms (panel A) and in the 800 – 1200 ms (panel B) latency windows and the proportion of correct source responses for each encoding task (self *vs*. common) and valence (negative *vs*. neutral *vs*. positive) conditions. Bonferroni correction was applied to adjust for multiple comparisons. Accordingly, correlations are statistically significant if $p \leq .004$.

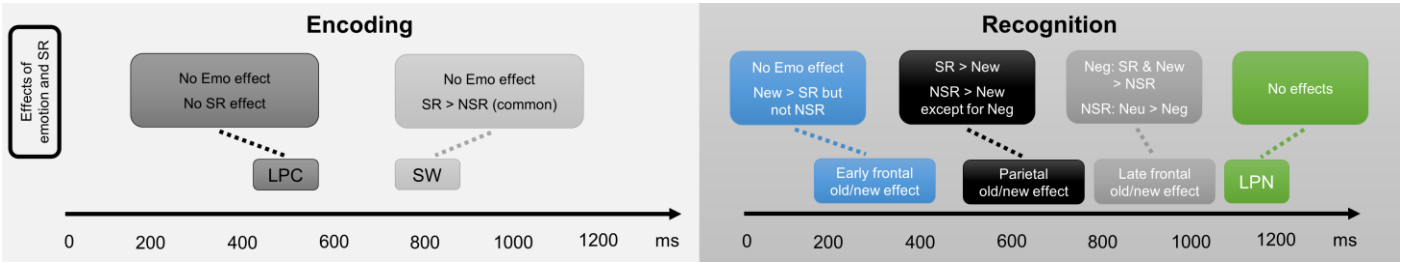


Figure 7. Schematic illustration of the event-related potentials of interest in the current study, including its main findings. Note: Emo = Emotion; ERP = Event-related potential; LPC = Late Positive Complex; LPN = Late Posterior Negativity; Neg = Negative; Neu = Neutral; NSR = Non-self-reference; Pos = Positive; SR = Self-reference; SW = Slow Wave.

Discussion

The main goal of the current study was two-fold: to probe whether and how valence and self-reference interactively modulate internal source memory, combining behavioral and electrophysiological measures (old/new effects) during recognition; to specify how these variables may also influence stimulus processing during intentional encoding, reflected in the LPC and SW amplitudes. With respect to behavioral performance, we observed that self-referential encoding granted an advantage not only in terms of general recognition (item memory) but, more importantly, in terms of source memory. Also, the effects of valence on source memory differed as a function of the encoding task (see Figure 3). With respect to electrophysiological measures, self-referential conditions yielded larger SW amplitudes during encoding, regardless of valence. During recognition, the advantage of self-referential encoding was observed in an earlier time window typically associated with familiarity (300 - 500 ms), and irrespective of valence. Interactive effects began to emerge in the 500 - 800 ms time window as old/new effects were observed for most conditions except for negative words studied in the common condition. Also, in the 800 - 1200 ms time window, the mean amplitude for negative words studied self-referentially was larger than for negative words studied in the common condition (see Figure 7).

Behavioral findings

When positive and neutral words were encoded self-referentially, both item and source memories were boosted as compared to negative words, corroborating our initial hypothesis and prior findings (item memory: Leshikar et al., 2015; Pauly et al., 2013; Yang, Truong, Fuss, & Bislimovic, 2012; source memory: Durbin et al., 2017; Leshikar et al., 2015; Pereira et al., 2019; Y. Zhang et al., 2018). Such results support the notion that while neutral and especially positive information is more likely to be deemed as self-referential, the same does not occur for negative information, whose content is less likely to fit into the current self-view (D'Argembeau & Van der Linden, 2008; Derry & Kuiper, 1981; Lewis, Cairney, Manning, & Critchley, 2011; Moran et al., 2006; Pauly et al., 2013). Other behavioral indices examined in the current study support the idea that negative words were perceived as less self-relevant (see Supplementary Materials section): self-reference effects on source memory were restricted to neutral and positive stimuli; the source of negative words was regarded as less memorable than of positive words considering both encoding tasks; the proportion of responses that matched the self-referential task during encoding (i.e., “yes” responses) was the lowest for negative words encoded self-referentially, whilst the proportion of misses during recognition was the highest for these words. Also, the improved source memory for neutral and positive words might be explained according to the notion that successful access

to information stored in memory is modulated by how well this information is integrated into prior knowledge structures and how well it fits long-term goals (Conway & Pleydell-Pearce, 2000).

In the case of the non-self-referential (common) condition and following a similar study (Pereira et al., 2019, Experiment 2), valence was expected to impair source memory. This hypothesis was partially confirmed as source memory performance for neutral words was improved when compared to positive words. However, the performance for negative words was no different from neutral words, which stands in contrast with prior studies testing memory for non-self-referential and semantic encoding tasks (e.g., Cook et al., 2007; Ferré et al., 2019, Experiment 2; Mao et al., 2015; Newsome et al., 2012; Otani, Jaffa, Libkuman, Goernert, & Kato, 2012; Otani, Libkuman, et al., 2012). Notwithstanding, the lack of differences in internal source memory for emotional *vs.* neutral stimuli was also reported in prior studies (e.g., Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006a; Pereira et al., 2019, Experiment 1; Sharot & Yonelinas, 2008). According to the arousal-biased competition theory (Mather & Sutherland, 2011) and the assumption that intentional encoding conditions may lead participants to prioritize information that meets contextual task demands, this should be translated into a null effect of valence on source memory. Even though emotion did not affect item memory and no differences were observed between negative and both neutral and positive words in source memory performance, the difference between neutral and positive items warrants further clarification. In this context, the self-positivity bias may offer a plausible explanation. In fact, a supplementary analysis ran on the proportion of incorrect source memory attributions revealed that the source of positive words studied in the common condition was more often mistaken as self-referential when compared to neutral words studied in the same condition (see Supplementary Materials section). Given that the source memory measure computed here included incorrect source responses, it is possible that the former bias accounted for the difference between neutral and positive words.

Emotion and self-reference effects during encoding

Contrary to our initial predictions, the LPC and SW amplitudes were not modulated by valence. Rather, differences in word processing only emerged in the SW latency windows, with words studied in the self-referential condition displaying a more positive-going amplitude than words studied in the common condition. Although this effect shows that self-reference grants a processing advantage in late processing stages (e.g., Herbert et al., 2011; Mu & Han, 2010; Nowicka et al., 2018; Zhao et al., 2016), the LPC amplitude was also expected to be sensitive to stimulus valence (Gutchess & Kensinger, 2018). The observation that the LPC was not modulated by valence indicates that similar attentional, evaluative

and elaboration processes were allocated to all words irrespective of valence or encoding task. In this context, intentional learning instructions might have contributed to the null finding. Notwithstanding, previous studies that have also used intentional conditions reported emotion effects on both LPC and SW (e.g., Barnacle et al., 2018; Dolcos & Cabeza, 2002; Palomba et al., 1997; Wirkner et al., 2018). Thus, other experimental factors such as the use of words (*vs.* pictures) and of similar arousal levels across valence categories may offer a more plausible account.

Specifically, several authors have argued that verbal stimuli do not convey the same emotional intensity as pictures and, as a result, null findings are more likely to occur when words are used as experimental stimuli (Citron, 2012; Dolcos & Cabeza, 2002; Herbert et al., 2006; Mathews, Ridgeway, & Holmes, 2013). Indeed, emotional pictures appear to be characterized by enhanced biological relevance and physiological arousal when compared to emotional words (Hinojosa, Carretié, Valcárcel, Méndez-Bértolo, & Pozo, 2009). Furthermore, valence and arousal properties may distinctly modulate ERP components as a function of attention focus. For example, Delaney-Busch and colleagues (2016) demonstrated that, when participants were overtly instructed to attend to the affective properties of the stimuli, the LPC was modulated by stimulus valence but not arousal. On the contrary, when the affective properties of the stimuli were task-irrelevant (as in the current study), the LPC was modulated by arousal but not by valence. Given that in the current study words were controlled for arousal, the lack of a valence effect might be accounted for by the implicit (*vs.* explicit) focus on the affective properties of the stimuli. This possibility agrees with the notion that the impact of emotion on the LPC and subsequent late components is likely to be attenuated when the attention of the participants is diverted away from the affective content of the stimuli (e.g., Hajcak et al., 2006; Hajcak & Nieuwenhuis, 2006). This also explains why the only difference observed during encoding was between encoding tasks as participants were overtly instructed to pay attention to the stimulus and encoding task, which were consistently highlighted and tested in distinct study-test cycles.

Emotion and self-reference effects during recognition

In the early frontal old/new effect time window, which has been typically associated with familiarity (e.g., Curran, 2000; Curran & Dien, 2003; Czernochowski et al., 2005), no valence-related modulations were observed, supporting previous studies (Johansson et al., 2004; Koenig & Mecklinger, 2008; Lavoie & O'Connor, 2013; Maratos et al., 2000; Minor & Herzmann, 2019; Newsome et al., 2012; Weymar et al., 2009). Nevertheless, the current study showed a self-reference effect, as old/new distinctions were only seen for words studied in the self-referential condition. Additionally, the ERP response associated

with common-source hits was similar to both self-source hits and correct rejections. This indicates that such response was sensitive to both familiar and unfamiliar items, whereas most items studied self-referentially probably elicited a familiarity effect at this stage. Even though prior studies showed that familiarity processes are sensitive to the type of encoding task (Nyhus & Curran, 2009; Peters & Daum, 2009) and source details (Mollison & Curran, 2012), this result was unexpected due to the use of intentional encoding conditions and considering prior studies using encoding tasks with deep processing demands (Dulas et al., 2011; Marzi & Viggiano, 2010). Instead, our results support the view that self-referential processing enhances familiarity processes associated with source monitoring (Mao et al., 2017). This advantage might be accounted for by unitization processes, which occur when item and source features are combined and integrated into a meaningful unit of information during encoding, resulting in a facilitated recognition of source features driven by familiarity processes (Diana, Van Den Boom, Yonelinas, & Ranganath, 2011; Diana, Yonelinas, & Ranganath, 2008; Kuo & Van Petten, 2006). The presence of unitization processes in self-referential conditions also dovetails with the proposal that self-reference promotes a better integration of information into preexistent structures of knowledge, namely self-schema (Rogers et al., 1977). Nonetheless, more evidence is required to support these claims as in the current study the electrophysiological response in the first latency window (300 - 500 ms) did not differ as a function of self-relevance (i.e., was similar for self- and common-source hits).

When considering the parietal old/new effect, typically associated with the recollection of information from the study episode including source features (Addante et al., 2012; Allan & Rugg, 1998; Cansino & Trejo-Morales, 2008; Leynes et al., 2017; MacLeod & Donaldson, 2017; Peters & Daum, 2009; Trott et al., 1999; Wilding & Rugg, 1997; Woroch & Gonsalves, 2010), the expected advantage for words encoded self-referentially was not observed. Based on the assumption that the amplitude of the parietal old/new effect is sensitive to variations in the amount of recoverable information from the study episode (e.g., Leynes & Crawford, 2018; Leynes et al., 2005; Wilding, 2000; Woroch & Gonsalves, 2010), this result suggests that the amount of information retrieved was similar across encoding tasks, regardless of whether it was or not source-specifying (i.e., diagnostic) information (see Leynes & Crawford, 2018; Leynes & Mok, 2017).

Nonetheless, whereas old/new effects were observed for neutral and positive words irrespective of encoding task, the old/new effect in the case of negative words was only detected in the self-referential condition. The lack of the old/new effect for negative words studied in the common condition demonstrated that emotional information is not always associated with enhanced recollection, as already documented by prior studies (Koenig & Mecklinger, 2008; Lavoie & O'Connor, 2013; Mao et al., 2015;

Maratos et al., 2000). Instead, old/new effects may or not occur for distinct emotional conditions depending on the type of source features tested during recognition (Koenig & Mecklinger, 2008). Specifically, negative valence might hinder the recollection of extrinsic source details, i.e., attributes that are not integrated in stimulus processing (see the object-based binding theory by Mather, 2007), such as the encoding task used during the study phase (Mao et al., 2015). A similar effect was not observed for negative words encoded self-referentially as stimulus and encoding task were likely unitized during encoding, thus favoring the recollection of contextual information.

Importantly, the analysis of right-frontal old/new effect (800 - 1200 ms) revealed that self-referential negative words and negative correct rejections elicited larger mean amplitudes than negative words studied in the common condition. Furthermore, in the non-self-referential (common) condition, the mean amplitude for neutral words was larger than for negative words. Considering that the old/new effect has been associated with post-retrieval monitoring processes that seem sensitive to retrieval demands (Donaldson & Rugg, 1998; Leynes, 2002; Leynes et al., 2006; Trott et al., 1999; Van Petten et al., 2000; Wilding & Rugg, 1996), including difficult memory decisions (e.g., Rugg et al., 2000), the former differences may indicate that decisions associated with non-self-referential neutral words and self-referential negative words more strongly engaged post-retrieval monitoring resources.

The difference between negative self- and common-source hits may be related to the fact that self-referential conditions were likely associated with the retrieval of details containing less source-specifying or diagnostic of the encoding task. Thus, negative stimuli in such conditions demanded a greater involvement of monitoring processes (see Koenig & Mecklinger, 2008; Newsome et al., 2012 for similar arguments). Considering that source memory performance for negative self-referential (*vs.* neutral and positive self-referential) words was impaired (yet at similar levels when compared to the common condition), it could also be the case that participants conducted a secondary memory search to resolve possible uncertainties regarding the association between negative information and the self (Dulas & Duarte, 2013; Senkfor & Van Petten, 1998). Such uncertainties are likely to arise if we consider that negative information tends to be less associated with the self (D'Argembeau & Van der Linden, 2008; Derry & Kuiper, 1981; Lewis et al., 2011; Moran et al., 2006; Pauly et al., 2013). Indeed, motivational goals associated with self-image and self-beliefs may interfere with retrieval processes by favoring the access to information that fits with the current self-view (e.g., neutral and positive self-referential information), while hindering the access to less relevant information (e.g., negative self-referential information; Conway & Pleydell-Pearce, 2000; Conway, Singer, & Tagini, 2004; D'Argembeau & Van der Linden, 2008). Additionally, the fact that negative words were generally associated with more difficult

memory decisions during recognition might explain the larger ERP amplitude in response to negative new words, in the sense that participants strategically monitored these items to ensure no prior links to the self.

An LPN effect was not observed in the 1200 - 1800 ms latency window, which suggests that the discrimination between the two source features did not impose further evaluation or reconstruction of task-relevant elements from the study episode. This outcome was expected given the use of intentional encoding conditions that emphasized both stimulus and encoding task and the fact that these elements were also targeted during recognition, which may facilitate memory retrieval according to the encoding specificity principle (Tulving & Thompson, 1973). In such experimental conditions, it is likely that source-specifying information is readily available for most retrieval attempts, thus resulting in an absent LPN effect (Mecklinger et al., 2016).

Limitations

One of the main limitations of this study was the number of stimuli per experimental condition, which restricted several aspects of the analysis during encoding and recognition. For instance, it was not possible to compare electrophysiological activity between remembered and forgotten items during word encoding (e.g., difference due to subsequent memory - Dm effects: Dolcos & Cabeza, 2002; Leynes & Crawford, 2018). Importantly, it was not possible to contrast ERP responses elicited by correct *vs.* incorrect source attributions, which should be examined in future studies. This would clarify how valence and self-referential encoding influence source memory recognition in accurate (containing source-specifying features) *vs.* erroneous source decisions (characterized by retrieval attempts containing other features which are not source-specifying; Cansino et al., 2012; Cansino & Trejo-Morales, 2008; Leynes & Mok, 2017; Mao et al., 2015). Together with fact that the sample was mainly composed by female participants (see Glaser, Mendrek, Germain, Lakis, & Lavoie, 2012 for an example of sex-related memory variations), these factors limit the generalization of the current findings. Additionally, given that self-referential conditions grant an advantage in terms of memory performance, this directly influenced the number of artifact-free trials available for the ERP analysis during recognition, creating an imbalance between encoding tasks (see Supplementary Table S2). In this regard, future studies may use a self-referential encoding strategy and probe its impact on other source features such as spatial location, stimulus color, or even other encoding tasks (e.g., Hou et al., 2019; Yin et al., 2019).

Conclusion

The current study revealed that valence and self-reference interactively modulate both behavioral and ERP indices of source memory. Specifically, self-referential (*vs.* non self-referential) processing resulted in improved internal source memory for positive and neutral words. Electrophysiologically, the self-reference advantage was supported by processes occurring during encoding and recognition. During encoding, self-referential conditions led to amplitude enhancements (700 - 950 and 950 - 1200 ms latency windows), possibly associated with a greater mobilization of sustained elaboration and rehearsal resources. During recognition, self-reference effects were manifested in an early processing stage typically related to familiarity, supporting previous claims that familiarity can contribute to source memory recognition (e.g., Diana et al., 2011, 2008; Kuo & Van Petten, 2006; Mao et al., 2015). Nonetheless, interactions between self-reference and valence were particularly observed in subsequent latency windows: while old/new effects were detected for negative words studied self-referentially in the 500 - 800 ms time window, such effect did not emerge for the common condition; when participants had to decide if negative words were previously studied in the self-referential condition, post-retrieval and monitoring processes were more strongly engaged in the 800 - 1200 ms time window. Such findings suggest that negative valence may affect the retrieval of extrinsic source features such as the task used during encoding (Mao et al., 2015). Additionally, they indicate that motivational goals, including the maintenance of a positive self-view, may favor the access to information that matches these goals, while hindering the access to less relevant information (Conway & Pleydell-Pearce, 2000; Conway et al., 2004; D'Argembeau & Van der Linden, 2008). Considering that changes in these positive attributional patterns may be associated with psychiatric disorders such anxiety and depression (Auerbach et al., 2015; Pauly et al., 2013; Shestyuk & Deldin, 2010), unraveling how emotional self-referential information affects distinct source memory features (e.g., perceptual; spatio-temporal; semantic; affective; cognitive operations; see Johnson et al., 1993) may contribute to a better understanding of the neurocognitive mechanisms underpinning memory dysfunction in these clinical conditions.

References

Addante, Richard J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. *NeuroImage*, *62*(1), 439–450. https://doi.org/10.1016/j.neuroimage.2012.04.031

Allan, K., & Rugg, M. D. (1998). Neural correlates of cued recall with and without retrieval of source memory. *NeuroReport*, *9*(15), 3463–3466. https://doi.org/10.1097/00001756-199810260-

00023

Auerbach, R. P., Stanton, C. H., Proudfit, G. H., & Pizzagalli, D. A. (2015). Self-referential processing in depressed adolescents: A high- density ERP study. *Journal of Abnormal Psychology*, *124*(2), 233–245. https://doi.org/10.1037/abn0000023

Barnacle, G. E., Tsivilis, D., Schaefer, A., & Talmi, D. (2018). Local context influences memory for emotional stimuli but not electrophysiological markers of emotion-dependent attention. *Psychophysiology*, 55(4), e13014. https://doi.org/10.1111/psyp.13014

Bayer, M., Sommer, W., & Schacht, A. (2012). P1 and beyond: Functional separation of multiple emotion effects in word recognition. *Psychophysiology*, *49*(7), 959–969. https://doi.org/10.1111/j.1469-8986.2012.01381.x

Beck, A. T., Steer, R., & Brown, G. (1996). *Beck Depression Inventory-II manual*. New York: The Psychological Corporation.

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Cansino, S., Hernández-Ramos, E., & Trejo-Morales, P. (2012). Neural correlates of source memory retrieval in young, middle-aged and elderly adults. *Biological Psychology*, *90*(1), 33–49. https://doi.org/10.1016/j.biopsycho.2012.02.004

Cansino, S., & Trejo-Morales, P. (2008). Neurophysiology of successful encoding and retrieval of source memory. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 85–98. https://doi.org/10.3758/CABN.8.1.85

Chaumon, M., Bishop, D. V. M., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, *250*, 47–63. https://doi.org/10.1016/j.jneumeth.2015.02.025

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*(3), 211–226. https://doi.org/10.1016/j.bandl.2011.12.007

Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and event-related potentials: Distinguishing early and late processes in affective picture perception. *Journal of Cognitive Neuroscience*, *19*(4), 577–586. https://doi.org/ 10.1162/jocn.2007.19.4.577

Coelho, R., Martins, A., & Barros, H. (2002). Clinical profiles relating gender and depressive symptoms among adolescents ascertained by the Beck Depressive Inventory II. *European Psychiatry, 17*(4), 222-226. https://doi.org/10.1016/S0924-9338(02)00663-6

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Pscyhological Review*, *107*(2), 261–288. https://doi.org/10.1037//0033-295X.107.2.261

Conway, M. A., Singer, J. F., & Tagini, A. (2004). The self and autobiographical memory: Correspondence and coherence. *Social Cognition*, *22*(5), 491–529. https://doi.org/10.1521/soco.22.5.491.50768

Cook, G. I., Hicks, J. L., & Marsh, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition*, *35*(2), 222–230. https://doi.org/10.3758/BF03193443

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923–938. https://doi.org/10.3758/BF03209340

Curran, T., & Dien, J. (2003). Differentiating amodal familiarity from modality-specific memory processes: An ERP study. *Psychophysiology*, *40*(6), 979–988. https://doi.org/10.1111/1469-8986.00116

Cycowicz, Y., & Friedman, D. (2003). Source memory for the color of pictures: Event-related brain potentials (ERPs) reveal sensory-specific retrieval-related activity. *Psychophysiology*, *40*(3), 455–464. https://doi.org/10.1111/1469-8986.00047

Czernochowski, D., Mecklinger, A., Johansson, M., & Brinkmann, M. (2005). Age-related differences in familiarity and recollection: ERP evidence from a recognition memory study in children and young adults. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(4), 417–433. https://doi.org/10.3758/CABN.5.4.417

D'Argembeau, A., Comblain, C., & Van Der Linden, M. (2005). Affective valence and the self-reference effect: Influence of retrieval conditions. *British Journal of Psychology*, *96*(4), 457–466. https://doi.org/10.1348/000712605X53218

D'Argembeau, A., & Van der Linden, M. (2008). Remembering pride and shame: Self-enhancement and the phenomenology of autobiographical memory. *Memory*, *16*(5), 534–547. https://doi.org/10.1080/09658210802010463

Delaney-Busch, N., Wilkie, G., & Kuperberg, G. (2016). Vivid: How valence and arousal influence word processing under different task demands. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(3), 415–432. https://doi.org/10.3758/s13415-016-0402-y

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG

dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, *90*(4), 286–297. https://doi.org/10.1037/0021-843X.90.4.286

Diana, R. A., Van Den Boom, W., Yonelinas, A. P., & Ranganath, C. (2011). ERP correlates of source memory: Unitized source information increases familiarity-based retrieval. *Brain Research*, *1367*, 278–286. https://doi.org/10.1016/j.brainres.2010.10.030

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 730–740. https://doi.org/10.1037/0278-7393.34.4.730

Diedrich, O., Naumann, E., Maier, S., Becker, G., & Bartussek, D. (1997). A frontal positive slow wave in the ERP associated with emotional slides. *Journal of Psychophysiology, 11*(1), 71-84.

Dietrich, D. E., Waller, C., Johannes, S., Wiering, B. M., Emrich, H. M., & Münte, T. F. (2001). Differential effects of emotional content on event-related potentials in word recognition memory. *Neuropsychobiology*, *43*(2), 96–101. https://doi.org/10.1159/000054874

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: Encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(3), 252–263. https://doi.org/10.3758/CABN.2.3.252

Donaldson, D. I., & Rugg, M. D. (1998). Recognition memory for new associations: Electrophysiological evidence for the role of recollection. *Neuropsychologia*, *36*(5), 377–395. https://doi.org/10.1016/S0028-3932(97)00143-7

Dulas, M. R., & Duarte, A. (2012). The effects of aging on material-independent and material-dependent neural correlates of source memory retrieval. *Cerebral Cortex*, *22*(1), 37–50. https://doi.org/10.1093/cercor/bhr056

Dulas, M. R., & Duarte, A. (2013). The influence of directed attention at encoding on source memory retrieval in the young and old: An ERP study. *Brain Research*, *1500*, 55–71. https://doi.org/10.1016/j.brainres.2013.01.018

Dulas, M. R., & Duarte, A. (2014). Aging affects the interaction between attentional control and source memory: An fMRI study. *Journal of Cognitive Neuroscience*, *26*(12), 2653–2669. https://doi.org/10.1162/jocn_a_00663

Dulas, M. R., Newsome, R. N., & Duarte, A. (2011). The effects of aging on ERP correlates of source memory retrieval for self-referential information. *Brain Research*, *1377*, 84–100. https://doi.org/10.1016/j.brainres.2010.12.087

Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory, 25*(9), 1191-1200. https://doi.org/10.1080/09658211.2017.1282517

Fan, W., Chen, J., Wang, X. Y., Cai, R., Tan, Q., Chen, Y., ...Zhong, Y. (2013). Electrophysiological correlation of the degree of self-reference effect. *PLoS ONE*, *8*(12), e80289. https://doi.org/10.1371/journal.pone.0080289

Ferré, P., Comesaña, M., & Guasch, M. (2019). Emotional content and source memory for language: Impairment in an incidental encoding task. *Frontiers in Psychology*, *10*, 65. https://doi.org/10.3389/fpsyg.2019.00065

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: Words, phrases, and task effects. *Progress in Brain Research*, *156*, 185–203. https://doi.org/10.1016/S0079-6123(06)56009-1

Fossati, P., Hevenor, S. J., Lepage, M., Graham, S. J., Grady, C., Keightley, M. L., ...Mayberg, H. (2004). Distributed self in episodic memory: Neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage*, *22*(4), 1596–1604. https://doi.org/10.1016/j.neuroimage.2004.03.034

Foti, D., Hajcak, G., & Dien, J. (2009). Differentiating neural responses to emotional pictures: Evidence from temporal-spatial PCA. *Psychophysiology*, *46*(3), 521–530. https://doi.org/10.1111/j.1469-8986.2009.00796.x

Friedman, D., Cycowicz, Y. M., & Bersick, M. (2005). The late negative episodic memory effect: The effect of recapitulating study details at test. *Cognitive Brain Research*, *23*(2–3), 185–198. https://doi.org/10.1016/j.cogbrainres.2004.10.005

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique*, *51*(1), 6–28. https://doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R

Gallant, S. N., Pun, C., & Yang, L. (2018). Age differences in the neural correlates underlying control of emotional memory: An event-related potential study. *Brain Research*, *1697*, 83–92. https://doi.org/10.1016/j.brainres.2018.06.019

Gallant, S. N., Spaniol, J., & Yang, L. (2019). Age differences in cue utilization during prospective and

retrospective memory monitoring. *Psychology and Aging, 34*(4), 545-557. http://doi.org/10.1037/pag0000352

Gallant, S. N., & Yang, L. (2014). Positivity effect in source attributions of arousal-matched emotional and non-emotional information during itembased directed forgetting. *Frontiers in Psychology, 5*, 1334. https://doi.org/10.3389/fpsyg.2014.01334

Glaser, E., Mendrek, A., Germain, M., Lakis, N., & Lavoie, M. E. (2012). Sex differences in memory of emotional images: A behavioral and electrophysiological investigation. *International Journal of Psychophysiology*, *85*(1), 17–26. https://doi.org/10.1016/j.ijpsycho.2012.01.007

Grundy, J. G., Benarroch, M. F. F., Lebarr, A. N., & Shedden, J. M. (2015). Electrophysiological correlates of implicit valenced self-processing in high vs. low self-esteem individuals. *Social Neuroscience*, *10*(1), 100–112. https://doi.org/10.1080/17470919.2014.965339

Gutchess, A., & Kensinger, E. A. (2018). Shared mechanisms may support mnemonic benefits from self-referencing and emotion. *Trends in Cognitive Sciences*, *22*(8), 712–724. https://doi.org/10.1016/j.tics.2018.05.001

Hajcak, G., Moser, J. S., & Simons, R. F. (2006). Attending to affect: Appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion*, *6*(3), 517–522. https://doi.org/10.1037/1528-3542.6.3.517

Hajcak, G., & Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 291–297. https://doi.org/10.3758/CABN.6.4.291

Herbert, C., Herbert, B. M., Ethofer, T., & Pauli, P. (2010). His or mine? The time course of self-other discrimination in emotion processing. *Social Neuroscience*, *6*(3), 277–288. https://doi.org/10.1080/17470919.2010.523543

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., & Rockstroh, B. (2006). Processing of emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology*, *43*(2), 197–206. https://doi.org/10.1111/j.1469-8986.2006.00385.x

Herbert, C., Pauli, P., & Herbert, B. M. (2011). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Social Cognitive and Affective Neuroscience*, *6*(5), 653–661. https://doi.org/10.1093/scan/nsq082

Hinojosa, J. A., Carretié, L., Valcárcel, M. A., Méndez-Bértolo, C., & Pozo, M. A. (2009). Electrophysiological differences in the processing of affective information in words and pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(2), 173–189.

https://doi.org/10.3758/CABN.9.2.173

Hinojosa, J. A., Méndez-Bértolo, C., & Pozo, M. A. (2010). Looking at emotional words is not the same as reading emotional words: Behavioral and neural correlates. *Psychophysiology*, *47*(4), 748–757. https://doi.org/10.1111/j.1469-8986.2010.00982.x

Hou, M., Grilli, M. D., & Glisky, E. L. (2019). Self-reference enhances relational memory in young and older adults. *Aging, Neuropsychology, and Cognition, 26*(1), 105-120. https://doi.org/10.1080/13825585.2017.1409333

Inaba, M., Nomura, M., & Ohira, H. (2005). Neural evidence of effects of emotional valence on word recognition. *International Journal of Psychophysiology*, *57*(3), 165–173. https://doi.org/10.1016/j.ijpsycho.2005.01.002

Johansson, M., & Mecklinger, A. (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. *Biological Psychology*, *64*(1–2), 91–117. https://doi.org/10.1016/S0301-0511(03)00104-2

Johansson, M., Mecklinger, A., & Treese, A. C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience*, *16*(10), 1840–1853. https://doi.org/10.1162/0898929042947883

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28. https://doi.org/10.1037/0033-2909.114.1.3

Kelley, A. W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794. https://doi.org/10.1162/08989290260138672

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A., & Schacter, D. L. (2006a). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *The Journal of Neuroscience*, *26*(9), 2564–2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

Kensinger, E. A, & Schacter, D. L. (2006b). Reality monitoring and memory distortion: effects of negative, arousing content. *Memory & Cognition*, *34*(2), 251–260. https://doi.org/10.3758/BF03193403

Koenig, S., & Mecklinger, A. (2008). Electrophysiological correlates of encoding and retrieving emotional events. *Emotion*, *8*(2), 162–173. https://doi.org/10.1037/1528-3542.8.2.162

Kuo, T. Y., & Van Petten, C. (2006). Prefrontal engagement during source memory retrieval depends on the prior encoding task. *Journal of Cognitive Neuroscience*, *18*(7), 1133–1146. https://doi.org/10.1162/jocn.2006.18.7.1133

Langeslag, S. J. E., & Van Strien, J. W. (2008). Age differences in the emotional modulation of ERP old/new effects. *International Journal of Psychophysiology*, *70*(2), 105–114. https://doi.org/10.1016/j.ijpsycho.2008.07.022

Lavoie, M. E., & O'Connor, K. P. (2013). Effect of emotional valence on episodic memory stages as indexed by event-related potentials. *World Journal of Neuroscience*, *03*(04), 250–262. https://doi.org/10.4236/wjns.2013.34034

Leshikar, E. D., & Duarte, A. (2012). Medial prefrontal cortex supports memory accuracy for self-referenced items. *Social Neuroscience*, *7*(2), 126–145. https://doi.org/10.1080/17470919.2011.585242

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150.

Lewis, P. A., Cairney, S., Manning, L., & Critchley, H. D. (2011). The impact of overnight consolidation upon memory for emotional and neutral encoding contexts. *Neuropsychologia*, *49*(9), 2619–2629. https://doi.org/10.1016/j.neuropsychologia.2011.05.009

Lewis, P. A., Critchley, H. D., Rotshtein, P., & Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral Cortex*, *17*(3), 742–748. https://doi.org/10.1093/cercor/bhk024

Leynes, P. A. (2002). The effect of specific test queries on source-monitoring event-related potentials. *Brain and Cognition*, *50*(2), 218–233. https://doi.org/10.1016/S0278-2626(02)00505-5

Leynes, P. A., Askin, B., & Landau, J. D. (2017). Visual perspective during remembering: ERP evidence of familiarity-based source monitoring. *Cortex*, *91*, 157–168. https://doi.org/10.1016/j.cortex.2016.12.003

Leynes, P. A., & Crawford, C. J. (2018). Event-related potential (ERP) evidence that encoding focus alters recollected features. *Brain and Cognition*, *127*, 42–50. https://doi.org/10.1016/j.bandc.2018.09.005

Leynes, P. A., Crawford, J. T., & Bink, M. L. (2005). Interrupted actions affect output monitoring and event-related potentials (ERPs). *Memory*, *13*(7), 759–772.

https://doi.org/10.1080/09658210444000377

Leynes, P. A., Grey, J. A., & Crawford, J. T. (2006). Event-related potential (ERP) evidence for sensory-based action memories. *International Journal of Psychophysiology*, *62*(1), 193–202. https://doi.org/10.1016/j.ijpsycho.2006.04.003

Leynes, P. A., & Mok, B. A. (2017). Encoding focus alters diagnostic recollection and event-related potentials (ERPs). *Brain and Cognition*, *117*, 1–11. https://doi.org/10.1016/j.bandc.2017.06.011

Leynes, P. A., & Phillips, M. C. (2008). Event-related potential (ERP) evidence for varied recollection during source monitoring. *Journal of Experimental Psychology: Learning, Memory, and, Cognition*, *34*(4), 741–751. https://doi.org/10.1037/0278-7393.34.4.741

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

MacKenzie, G., Powell, T. F., & Donaldson, D. I. (2015). Positive emotion can protect against source memory impairment. *Cognition and Emotion*, *29*(2), 236–250. https://doi.org/10.1080/02699931.2014.911145

MacLeod, C. A., & Donaldson, D. I. (2017). Investigating the functional utility of the left parietal ERP old/new effect: Brain activity predicts within but not between participant variance in episodic recollection. *Frontiers in Human Neuroscience*, *11*, 580. https://doi.org/10.3389/fnhum.2017.00580

Maddock, R. J., & Frein, S. T. (2009). Reduced memory for the spatial and temporal context of unpleasant words. *Cognition and Emotion*, *23*(1), 96–117. https://doi.org/10.1080/02699930801948977

Mao, X., Wang, Y., Wu, Y., & Guo, C. (2017). Self-referential information alleviates retrieval inhibition of directed forgetting effects—an ERP evidence of source memory. *Frontiers in Behavioral Neuroscience*, *11*, 187. https://doi.org/10.3389/fnbeh.2017.00187

Mao, X., You, Y., Li, W., & Guo, C. (2015). Emotion impairs extrinsic source memory-An ERP study. *Biological Psychology*, *110*, 182–189. https://doi.org/10.1016/j.biopsycho.2015.07.005

Maratos, E. J., Allan, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, *38*(11), 1452–1465. https://doi.org/10.1016/S0028-3932(00)00061-0

Marzi, T., & Viggiano, M. P. (2010). Deep and shallow encoding effects on face recognition: An ERP study.

International Journal of Psychophysiology, *78*(3), 239–250. https://doi.org/10.1016/j.ijpsycho.2010.08.005

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competitition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

Mathews, A., Ridgeway, V., & Holmes, E. A. (2013). Feels like the real thing: Imagery is both more realistic and emotional than verbal thought. *Cognition and Emotion*, *27*(2), 217–229. https://doi.org/10.1080/02699931.2012.698252

McNeely, H. E., Dywan, J., & Segalowitz, S. J. (2004). ERP indices of emotionality and semantic cohesiveness during recognition judgments. *Psychophysiology*, *41*(1), 117–129. https://doi.org/10.1111/j.1469-8986.2003.00137.x

Mecklinger, A., Johansson, M., Parra, M., & Hanslmayr, S. (2007). Source-retrieval requirements influence late ERP and EEG memory effects. *Brain Research*, *1172*, 110–123. https://doi.org/10.1016/j.brainres.2007.07.070

Mecklinger, A., Rosburg, T., & Johansson, M. (2016). Reconstructing the past: The late posterior negativity (LPN) in episodic memory studies. *Neuroscience & Biobehavioral Reviews*, *68*, 621–638. https://doi.org/10.1016/j.neubiorev.2016.06.024

Meng, X., Zhang, L., Liu, W., Ding, X. S., Li, H., Yang, J., & Yuan, J. J. (2017). The impact of emotion intensity on recognition memory: Valence polarity matters. *International Journal of Psychophysiology*, *116*, 16–25. https://doi.org/10.1016/j.ijpsycho.2017.01.014

Minor, G., & Herzmann, G. (2019). Effects of negative emotion on neural correlates of item and source memory during encoding and retrieval. *Brain Research*, *1718*, 32–45. https://doi.org/10.1016/j.brainres.2019.05.001

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Mollison, M. V., & Curran, T. (2012). Familiarity in source memory. *Neuropsychologia*, *50*(11), 2546–2565. https://doi.org/10.1016/j.neuropsychologia.2012.06.027

Moran, J. M., Macrae, C. N., Heatherton, T. F., Wyland, C. L., & Kelley, W. M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive*

Neuroscience, *18*(9), 1586–1594. https://doi.org/10.1162/jocn.2006.18.9.1586

Mu, Y., & Han, S. (2010). Neural oscillations involved in self-referential processing. *NeuroImage*, *53*(2), 757–768. https://doi.org/10.1016/j.neuroimage.2010.07.008

Newsome, R. N., Dulas, M. R., & Duarte, A. (2012). The effects of aging on emotion-induced modulations of source retrieval ERPs: Evidence for valence biases. *Neuropsychologia*, *50*(14), 3370–3384. https://doi.org/10.1016/j.neuropsychologia.2012.09.024

Nowicka, M. M., Wójcik, M. J., Kotlewska, I., Bola, M., & Nowicka, A. (2018). The impact of self-esteem on the preferential processing of self-related information: Electrophysiological correlates of explicit self vs. other evaluation. *PLoS ONE*, *13*(7), e0200604. https://doi.org/10.1371/journal.pone.0200604

Nyhus, E., & Curran, T. (2009). Semantic and perceptual effects on recognition memory: Evidence from ERP. *Brain Research*, *1283*, 102–114. https://doi.org/10.1016/j.brainres.2009.05.091

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology General, 129*(2), 242-261. https://doi.org/10.1037//0096-3445.129.2.242

Otani, H., Jaffa, M., Libkuman, T. M., Goernert, P. N., & Kato, K. (2012). Does source memory impairment associated with emotionally arousing stimuli occur at encoding or retrieval? *Journal of Cognitive Psychology*, *24*(7), 796–801. https://doi.org/10.1080/20445911.2012.693468

Otani, H., Libkuman, T. M., Goernert, P. N., Kato, K., Migita, M., Freehafer, S. E., & Landow, M. P. (2012). Emotion, directed forgetting, and source memory. *British Journal of Psychology*, *103*(3), 343–358. https://doi.org/10.1111/j.2044-8295.2011.02078.x

Palomba, D., Angrilli, A., & Mini, A. (1997). Visual evoked potentials, heart rate responses and memory to emotional pictorial stimuli. *International Journal of Psychophysiology*, *27*(1), 55–67. https://doi.org/10.1016/S0167-8760(97)00751-4

Pauly, K., Finkelmeyer, A., Schneider, F., & Habel, U. (2013). The neural correlates of positive self-evaluation and self-related memory. *Social Cognitive and Affective Neuroscience*, *8*(8), 878–886. https://doi.org/10.1093/scan/nss086

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy on internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology*, *10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2020). Is internal source memory recognition modulated

by emotional encoding contexts? Advance online publication. *Psychological Research*.

Peters, J., & Daum, I. (2009). Frontal but not parietal positivity during source recollection is sensitive to episodic content. *Neuroscience Letters*, *454*(3), 182–186. https://doi.org/10.1016/j.neulet.2009.03.019

Pion-Tonachini, L., Makeig, S., & Kreutz-Delgado, K. (2017). Crowd labeling latent Dirichlet allocation. *Knowledge and Information Systems*, *53*(3), 749–765. https://doi.org/10.1007/s10115-017-1053-1

Rogers, T. B., Kuiper, N. A, & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Rugg, M., Allan, K., & Birch, C. (2000). Electrophysiological evidence for the modulation of retrieval orientation. *Journal of Cognitive Neuroscience*, *12*(4), 664–678. https://doi.org/10.1162/089892900562291

Schacht, A., & Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 28–43. https://doi.org/10.3758/CABN.9.1.28

Schaefer, A., Fletcher, K., Pottage, C. L., Alexander, K., & Brown, C. (2009). The effects of emotional intensity on ERP correlates of recognition memory. *NeuroReport*, *20*(3), 319–324. https://doi.org/10.1097/WNR.0b013e3283229b52

Schaefer, A., Pottage, C. L., & Rickart, A. J. (2011). Electrophysiological correlates of remembering emotional pictures. *NeuroImage*, *54*(1), 714–724. https://doi.org/10.1016/j.neuroimage.2010.07.030

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: Event-related brain potential studies. *Progress in Brain Research, 156*, 31-51. https://doi.org/10.1016/S0079-6123(06)56002-9

Senkfor, A. J., & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *24*(4), 1005–1025. https://doi.org/10.1037/0278-7393.24.4.1005

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Shestyuk, A. Y., & Deldin, P. J. (2010). Automatic and strategic representation of the self in major

depression: Trait and state abnormalities. *American Journal of Psychiatry*, *167*(5), 536–544. https://doi.org/10.1176/appi.ajp.2009.06091444

Silva, D. R., & Spielberger, C. D. (2007). *Manual do Inventário de Estado Traço de Ansiedade (STAI)*. USA: Mind Garden, Inc.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. http://doi.org/10.1037/0096-3445.117.1.34

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Trott, C. T., Friedman, D., Ritter, W., Fabiani, M., & Snodgrass, J. G. (1999). Episodic priming and memory for temporal source: Event-related potentials reveal age-related differences in prefrontal functioning. *Psychology and Aging*, *14*(3), 390–413. https://doi.org/10.1037/0882-7974.14.3.390

Tulving, E., & Donald, T. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

Vallesi, A., & Shallice, T. (2006). Prefrontal involvement in source memory: An electrophysiological investigation of accounts concerning confidence and accuracy. *Brain Research*, *1124*(1), 111–125. https://doi.org/10.1016/j.brainres.2006.09.076

Van Petten, C., Senkfor, A. J., & Newberg, W. M. (2000). Memory for drawings in locations: Spatial source memory and event-related potentials. *Psychophysiology*, *37*(4), 551–564. https://doi.org/10.1017/S0048577200990541

Van Strien, J. W., Langeslag, S. J. E., Strekalova, N. J., Gootjes, L., & Franken, I. H. A. (2009). Valence interacts with the early ERP old/new effect and arousal with the sustained ERP old/new effect for affective pictures. *Brain Research*, *1251*, 223–235. https://doi.org/10.1016/j.brainres.2008.11.027

Vilberg, K. L., & Rugg, M. D. (2009). An investigation of the effects of relative probability of old and new test items on the neural correlates of successful and unsuccessful source memory. *NeuroImage*, *45*(2), 562–571. https://doi.org/10.1016/j.neuroimage.2008.12.020

Voss, J. L., & Federmeier, K. D. (2011). FN400 potentials are functionally identical to N400 potentials

and reflect semantic processing during recognition testing. *Psychophysiology*, *48*(4), 532–546. https://doi.org/10.1111/j.1469-8986.2010.01085.x

Voss, J. L., & Paller, K. A. (2017). Neural substrates of remembering: Event-related potential studies. In *Learning and memory: A comprehensive reference* (pp. 81-98). Elsevier Inc. https://doi.org/10.1016/b978-0-12-809324-5.21070-5

Wang, B., & Fu, X. (2011). Time course of effects of emotion on item memory and source memory for Chinese words. *Neurobiology of Learning and Memory*, *95*(4), 415–424. https://doi.org/10.1016/j.nlm.2011.02.001

Weymar, M., Löw, A., & Hamm, A. O. (2011). Emotional memories are resilient to time: Evidence from the parietal ERP old/new effect. *Human Brain Mapping*, *32*(4), 632–640. https://doi.org/10.1002/hbm.21051

Weymar, M., Löw, A., Melzig, C. A., & Hamm, A. O. (2009). Enhanced long-term recollection for emotional pictures: Evidence from high-density ERPs. *Psychophysiology*, *46*(6), 1200–1207. https://doi.org/10.1111/j.1469-8986.2009.00869.x

Wilding, E. L. (1999). Separating retrieval strategies from retrieval success: An event-related potential study of source memory. *Neuropsychologia*, *37*(4), 441–454. https://doi.org/10.1016/S0028-3932(98)00100-6

Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology*, *35*(1), 81–87. https://doi.org/10.1016/S0167-8760(99)00095-1

Wilding, E. L., Rugg, M. D. (1996). An event-related potential study of recognition with and without retrieval of source. *Brain: A Journal of Neurology*, *119*(3), 889–905. https://doi.org/10.1093/brain/119.3.889

Wilding, E. L., & Rugg, M. D. (1997). An event-related potential study of memory for spoken and heard words. *Neuropsychologia*, *35*(9), 1185–1195. https://doi.org/10.1016/S0028-3932(97)00048-1

Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, *13*(5), 577–592. https://doi.org/10.1162/089892901750363172

Wirkner, J., Ventura-Bort, C., Schulz, P., Hamm, A. O., & Weymar, M. (2018). Event-related potentials of emotional and neutral memories: The role of encoding position and delayed testing. *Psychophysiology*, *55*(7), e13069. https://doi.org/10.1111/psyp.13069

Woroch, B., & Gonsalves, B. D. (2010). Event-related potential correlates of item and source memory strength. *Brain Research*, *1317*, 180–191. https://doi.org/10.1016/j.brainres.2009.12.074

Xu, H., Zhang, Q., Li, B., & Guo, C. (2015). Dissociable effects of valence and arousal on different subtypes of old/new effect: Evidence from event-related potentials. *Frontiers in Human Neuroscience*, *9*, 650. https://doi.org/10.3389/fnhum.2015.00650

Yang, L., Truong, L., Fuss, S., & Bislimovic, S. (2012). The effects of ageing and divided attention on the self-reference effect in emotional memory: Spontaneous or effortful mnemonic benefits? *Memory*, *20*(6), 596–607. https://doi.org/10.1080/09658211.2012.690040

Yin, X., Ma, Y., Xu, X., & Yang, H. (2019). The effect of self-referencing on memory for different kinds of source information. *Memory*, *27*(4), 519–527. https://doi.org/10.1080/09658211.2018.1532009

Zhao, Y., Luo, W., Chen, J., Zhang, D., Zhang, L., Xiao, C., ...Tan, S. (2016). Behavioral and neural correlates of self-referential processing deficits in bipolar disorder. *Scientific Reports, 6*, 24075. https://doi.org/10.1038/srep24075

Zhang, H., Guan, L., Qi, M., & Yang, J. (2013). Self-esteem modulates the time course of self-positivity bias in explicit self-evaluation. *PLoS ONE*, *8*(12), e81169. https://doi.org/10.1371/journal.pone.0081169

Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology*, *65*(4), 236–244. https://doi.org/10.1027/1618-3169/a000409

Zhou, H., Guo, J., Ma, X., Zhang, M., Liu, L., Feng, L., ...Zhong, N. (2017). Self-reference emerges earlier than emotion during an implicit self-referential emotion processing task: Event-related potential evidence. *Frontiers in Human Neuroscience*, *11*, 451. https://doi.org/10.3389/fnhum.2017.00451

CHAPTER VI

GENERAL DISCUSSION

"Unidirectionality of time is one of nature's most fundamental laws. It has relentlessly governed all happenings in the universe—cosmic, geological, physical, biological, psychological—as long as the universe has existed. Galaxies and stars are born and they die, living creatures are young before they grow old, causes always precede effects, there is no return to yesterday, and so on and on. Time's flow is irreversible. The singular exception is provided by the human ability to remember past happenings. When one thinks today about what one did yesterday, time's arrow is bent into a loop. The rememberer has mentally traveled back into her past and thus violated the law of the irreversibility of the flow of time. (...). How did Mother Nature do it? We do not yet know, not only because the question has seldom been raised by those who study Mother Nature's ways but also because the brain/mind science that could throw some light on the question is still in its early formative years and many important issues to be explored have not yet been discovered."

(Tulving, 2002)

The main objective of this dissertation was to examine whether and how variations of emotion and encoding task may influence later internal source memory attributions. To contribute to this research problem, four distinct studies were developed and presented here. Of note, there was a fifth study in which functional magnetic resonance imaging (fMRI) data were collected, but it is still under analysis, hence it was not included here. Nonetheless, the behavioral data will be briefly mentioned (see Appendix E) and integrated in this general discussion.

With respect to the main findings of this work, first, it was observed that source memory research in this domain had been characterized by mixed findings and methodological heterogeneity. In this regard, the meta-analysis conducted here revealed that whereas for studies that manipulated stimulus valence (i.e., negative *vs*. neutral; positive *vs*. neutral; negative *vs*. positive), the effects of valence were null to very small (suggesting a deleterious effect for positive stimuli), in the case of studies that focused on stimulus arousal (i.e., high *vs*. low arousal; medium *vs*. low arousal; high *vs*. medium arousal), source memory performance was benefited for stimuli with medium-high (*vs*. low) arousal. Nonetheless, even in the presence of an emotion effect on source memory, the magnitude did not surpass the small interpretation convention. Given the heterogeneity among studies, one relevant finding was that both type of source memory task and type of source memory features emerged as moderators, especially in the context of valence-based studies. Even though other moderators were also identified, the former two were the most consistent. This supports the notion that different factors might modulate the effects of emotion on source memory.

Consistent with this idea, the role of encoding task, especially considering advantageous encoding strategies, was examined in the subsequent studies of this dissertation. Accordingly, it was found that the effects of emotion on internal source memory performance varied as a function of the encoding task. For instance, in the context of Chapter III, whereas a source memory enhancement was observed for both positive and neutral words when compared to negative words in the self-referential condition, both negative and positive words were associated with worse source memory performance in comparison with

neutral words in the common condition (semantic or non-self-referential condition). Furthermore, in both event-related potential (ERP; Chapter IV) and fMRI studies (see Appendix E), wherein a similar self-referential *vs*. common judgments approach was used, the behavioral evidence was in good agreement to what was described earlier, with the exception that no significant differences were found between negative and neutral words encoded in the common condition. Even so, when the encoding strategy was based on the production mode (Chapter III, Experiment 1), neither an effect of encoding task nor an effect of emotion was detected.

As a similar approach was used in three different experiments, with three different independent samples, this gives us the opportunity to implement a meta-analytic procedure as described in Chapter II. In this case, sensitivity analyses were not required as the original correlation coefficients between experimental conditions can be used to compute the sample variance of the effect size. More specifically, we decided to employ two sub-group analyses, one focusing on the effects of emotional information (negative *vs*. neutral; positive *vs*. neutral; negative *vs*. positive) when the encoding was self-referential and other when the encoding was semantic and non-self-referential.

In this regard, the metafor package (version 2.1-0; Viechtbauer, 2010) in RStudio (version 1.1.419; RStudio Team, 2018) was used once again. More specifically, the standard rma() function was employed. The results of the meta-analyses are summarized in Table 1 and the correspondent forest plots are displayed in Figure 1. In the case of the self-referential encoding conditions, it was observed that whereas no statistically significant differences were observed between positive and neutral words, d_{unb} = 0.18, *SE* = 0.11, z = 1.68, p = .092, 95% CI [- 0.03, 0.39], the negative words were associated with a worse source memory performance relative to both positive, d_{unb} = - 0.88, *SE* = 0.17, z = - 5.18, p < .0001, 95% CI [- 1.21, - 0.54] and neutral words, d_{unb} = - 0.70, *SE* = 0.12, z = - 5.18, p < .0001, 95% CI [- 0.94, - 0.46]. This pattern of results seems to fit with the notion that, in typical conditions, one is more likely to remember and attribute neutral and positive words to the self as these words are more likely to fit the current self-scheme and/or the current mood. This claim is supported by evidence showing that people are more willing to accept self-related positive information while they try to shield themselves from negative, threatful information (e.g., D'Argembeau, Comblain, & Van der Linden, 2005; Mezulis, Abramson, Hyde, & Hankin, 2004; Sedikides & Green, 2000, 2004; Watson, Dritschel, Jentzsch, & Obonsawin, 2008). It might be also supported by mood-related studies showing, for example, that people suffering from depression or dysphoria appear to display a memory advantage for negative but not for positive self-referential information (e.g., Denny & Hunt, 1992; Mathews & Bradle, 1983; Nasby, 1994; Watson et al., 2008; Zupan, Žeželj, & AndjelkoviĆ, 2017). Accordingly, if we consider that the participants

who collaborated in this dissertation were not in a negative but rather a positive/neutral mood at the time of the experiment (of note, at least a screening of depressive and anxiety symptoms was done in all studies), it could be the case that the encoding of self-related information was facilitated by the current mood (e.g., Itoh, 2004; Nasby, 1994; Subramaniam et al., 2016), especially considering information in consonance with it (i.e., positive and neutral but not negative information).

Table 1

Main Results of the Meta-analysis for each Valence Comparison and Encoding Task

	N of effects (*k*)	Estimate (d_{unb})	*SE*	*z*	*p*	95% CI [LB, UP]	*Q*, *p*
Neg-Neu							
Self	3	- 0.70	0.12	- 5.78	< .0001	- 0.94, - 0.46	2.31, .315
Common	3	- 0.49	0.12	- 4.15	< .0001	- 0.73, - 0.26	0.94, .626
Pos-Neu							
Self	3	0.18	0.11	1.68	.092	- 0.03, 0.39	2.52, .248
Common	3	- 0.78	0.12	- 6.73	< .0001	- 1.01, - 0.55	1.76, .415
Neg-Pos							
Self	3	- 0.88	0.17	- 5.18	< .0001	- 1.21, - 0.54	3.53, .171
Common	3	0.33	0.13	2.57	.010	0.08, 0.58	0.17, .919

Note. HB = Higher Bound; LB = Lower Bound; Neg = Negative; Neu = Neutral; Pos = Positive.

In the case of common encoding conditions, it was observed that both emotional words were related to worse source memory performance when compared to neutral words (negative *vs.* neutral: d_{unb} = - 0.49, *SE* = 0.12, *z* = - 4.15, *p* < .0001, 95% CI [- 0.73, - 0.26]; positive *vs.* neutral: d_{unb} = - 0.78, *SE* = 0.12, *z* = - 6.73, *p* < .0001, 95% CI [- 1.01, - 0.55]). Also, a difference between negative and positive words was observed, d_{unb} = 0.33, *SE* = 0.13, *z* = 2.57, *p* = .010, 95% CI [0.08, 0.58], with a better source memory performance for negative words, even though the effect size was small. Whereas the difference between neutral and positive stimuli was consistent across studies, the same was not observed for the comparison between neutral and negative stimuli, as the difference only reached the threshold of statistical significance in one of the studies (see Chapter III, Experiment 2). Furthermore, when considering each study separately, no statistically significant difference was detected between negative and positive stimuli, which emerged nonetheless in the meta-analysis. However, when looking at the effect sizes of these less consistent differences across studies (see Table 1 and Figure 1), they can be interpreted as small. This implies that the studies included here were likely underpowered to detect such differences. This seems to be the case at least with respect to the difference between negative and neutral stimuli, as it was detected in the study with the biggest sample size, but not in the other two studies with

smaller samples. From Chapter II, the valence-based subgroup analysis considering studies that used internal source memory tasks and that manipulated the encoding task as source feature yielded small effect sizes, which partially corroborates the previous claims. Furthermore, in this same chapter, it was possible to verify that the effects of emotional information in the context of internal source memory tasks oscillated between null or deleterious effects (Blix & Brennen, 2011; Ferré, Comesaña, & Guasch, 2019; Gallant, Spaniol, & Yang, 2019; Gallant & Yang, 2014; Kensinger & Schacter, 2006a; Mao, You, Li, & Guo, 2015; Otani, Jaffa, Libkuman, Goernert, & Kato, 2012; Otani, Libkuman, et al., 2012; Ritchey, Wang, Yonelinas, & Ranganath, 2019; Sharot & Yonelinas, 2008). This agrees to what was reported here, especially considering that, in comparison with self-referential tasks, common judgment tasks are closer to the type of tasks that have been used in the literature (see Table S5, Appendix A).

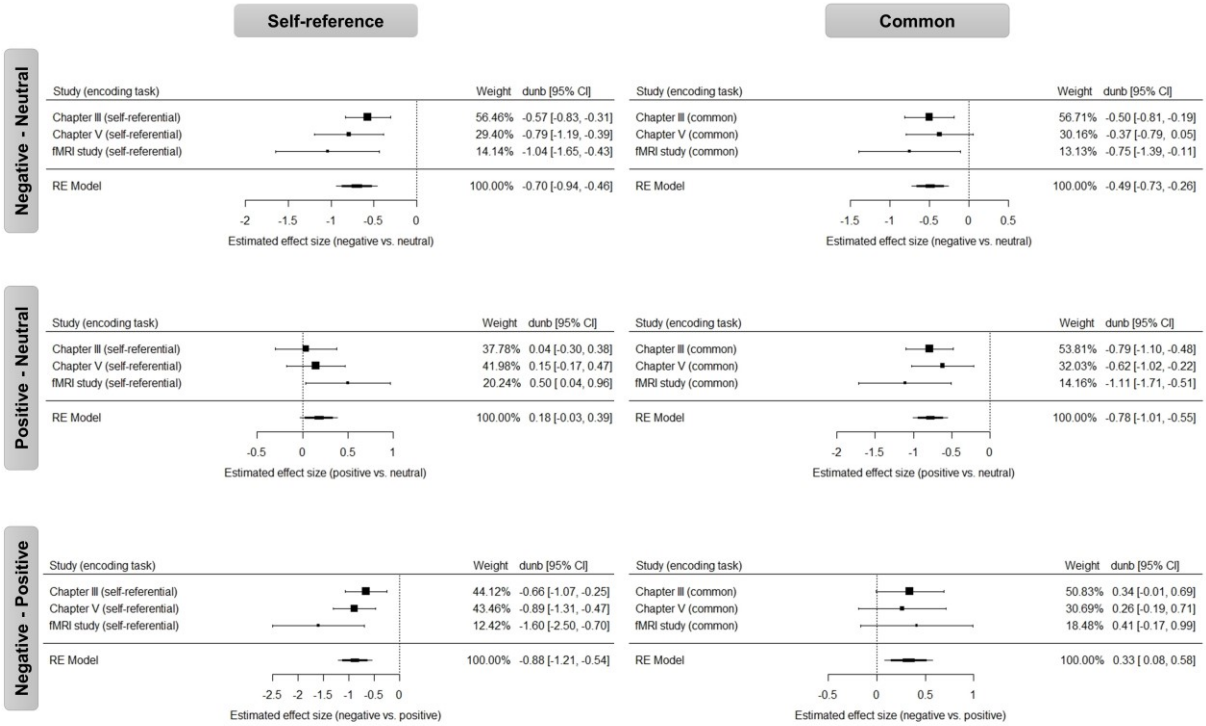


Figure 1. Forest plots for the six comparisons conducted with the behavioral data collected in the context of this dissertation. On the left side, the forest plots correspond to studies that used self-referential encoding tasks, while on the right side the studies correspond to those that used common encoding tasks.

Nevertheless, the difference between negative and positive stimuli reported here is at odds to what was found in the context of the meta-analysis of Chapter II. Indeed, no significant difference between these valence conditions was identified, even considering the sub-group analysis including only internal source memory studies. One possible explanation for the difference between negative and positive stimuli observed here might be related to the encoding tasks. In other words, this difference may result from an

experimental artifact that arises from the selected encoding tasks together with the measure used to compute source memory performance (specifically, the subtraction between correct and incorrect responses). Indeed, the analyses conducted on the proportion of incorrect source memory responses and included in the current dissertation as supplementary materials (see Appendix B and D) revealed that the positive words encoded in the common condition received more incorrect source memory attributions when compared to neutral words. When a similar analysis was conducted with the behavioral data of the fMRI study, the results demonstrated that emotional words studied in the common condition (negative: M = .19, SE = 0.02; p = .017, 95% CI_{diff} [0.01, 0.15]; positive: M = .24, SE = 0.03; p < .001, 95% CI_{diff} [0.07, 0.20]) received a greater proportion of incorrect source memory attributions than neutral words (M = .10, SE = 0.02). Additionally, positive words studied in the common condition (M = .24, SE = 0.03) were associated with more incorrect source memory attributions than positive words studied in the self-referential condition (M = .09, SE = 0.01; p < .001, 95% CI_{diff} [0.08, 0.21]). Taken together, these findings suggest that the source of positive words encoded in the common condition is more likely to be misattributed to the self-referential condition, which also agrees with the previously mentioned idea that people are somehow biased to attribute positive information to themselves. Accordingly, if positive words encoded in the common condition are more likely to attract source memory misattributions, then is plausible to assume that the difference in performance between positive and both negative and neutral words may be accentuated, as it was observed in the results of the meta-analysis presented here.

Continuing with the findings pertaining to source memory performance, a third outcome of this dissertation came from a different manipulation of emotion during encoding. Specifically, in Chapter IV, instead of varying the affective properties of the study items, emotional pictures were presented in the background during the encoding of neutral information. Both mixed and blocked designs were used to test the working hypothesis that emotional encoding conditions, especially negative ones, would interfere with the encoding of neutral information and, as consequence, with the retrieval of source memory information. Nonetheless, we found no robust evidence in favor of this hypothesis. We mainly reported null effects, suggesting that, at least in the context of the experiments described in Chapter IV, the internal source memory was not modulated by emotional encoding conditions. However, it is worth noting that this was one of the few studies employing this type of experimental paradigm in the context of source memory (see Macri, Pavard, & Versace, 2018; Subramaniam et al., 2016; Ventura-Bort et al., 2017; Xie & Zhang, 2017 for other studies), even though several instances can be found with respect to other aspects of episodic memory including item memory (e.g., Baeken et al., 2012; Bridge, Chiao, & Paller, 2010; Brierley, Medford, Shaw, & David, 2007; Erk et al., 2003; Fenker, Schott, Richardson-Klavehn,

Heinze, & Düzel, 2005; Graf & Yu, 2015; Guo, Li, Zhang, Cui, & Wei, 2018; Martínez-Galindo & Cansino, 2015, 2017; Pierguidi et al., 2016; Q. Zhang, Liu, An, Yang, & Wang, 2015). Furthermore, in the experiment with the blocked design, slower response times during encoding were obtained in the negative when compared to the neutral block. Also, from a secondary analysis using the participants' valence ratings, a deleterious effect of negative (*vs.* positive) encoding conditions on source memory performance was observed. Although these findings do not allow to make strong claims in favor of our initial hypothesis, they pinpoint that more studies, perhaps with an adjusted experimental design, are warranted to clarify the effects of emotional encoding contexts on source memory. Drawing a more general comparison between the results obtained in this study and the results reported earlier when emotion was manipulated as an intrinsic property of the stimulus, it seems that the effects of emotion on source memory may also differ as a function of how emotion is manipulated during encoding and this also requires further investigation.

In contrast, one finding that was consistent across the behavioral evidence reported here relates to the memory advantage observed when self-referential processing was used during encoding. This is not surprising given the robustness of this effect in multiple studies (e.g., Bock, 1986; D'Argembeau et al., 2005; Fossati et al., 2004; Ganellen & Carver, 1985; Grilli, Woolverton, Crawford, & Glisky, 2018; Gutchess, Kensinger, Yoon, & Schacter, 2007; Howell & Zelenski, 2017; Kelley et al., 2002; Kuiper & Rogers, 1979; Miall, 1986; Morel et al., 2014; Rogers, Kuiper, & Kirker, 1977; Rudolph, 1993; Symons & Johnson, 1997; Yang, Truong, Fuss, & Bislimovic, 2012; Yin, Ma, Xu, & Yang, 2019). Nevertheless, the studies of this dissertation support more recent evidence indicating that this encoding strategy may also benefit the memory for peripheral and contextual features of an episode (e.g., Dulas, Newsome, & Duarte, 2011; Durbin, Mitchell, & Johnson, 2017; Hou, Grilli, & Glisky, 2019a; Leshikar & Duarte, 2012; Leshikar, Dulas, & Duarte, 2015; Mao, Wang, Wu, & Guo, 2017; Yin et al., 2019; Y. Zhang, Pan, Li, & Guo, 2018). Furthermore, one of the contributions of this dissertation pertains to the examination of the effects of emotion when the encoding is self-referential, not only on source memory performance as already discussed in this chapter, but also with regards to brain mechanisms and prospective metamemory judgments of source.

In this regard, a fourth main finding of this dissertation came from the ERP study (Chapter V), in which the interactive effects of emotional valence and encoding task on source memory were analyzed using not only behavioral but also brain-derived responses. As the experimental session was divided in different study-test cycles and the montage of the electroencephalography (EEG) apparatus was one of the first steps in the data collection, this gave us the opportunity to collect electrophysiological data from

both study and test phases, even though our primary interest was on the processes underlying the retrieval of source memory information. Specifically, during encoding, when participants were asked to memorize emotional and neutral words using two different encoding tasks, it was possible to observe that while no emotion-related modulations emerged in the three time windows analyzed (500 - 700; 700 - 950; 950 - 1200 ms), differences between encoding conditions were observed in the two latest latency windows, with the mean amplitude of self-referenced words being larger than words studied in the common condition, specially over frontal, fronto-central, central, and centro-parietal electrodes. The fact that the difference between encoding tasks emerged later seems to support the idea that words studied in self-referential conditions engaged more sustained elaboration when compared to words studied in the common condition. This is also consistent with the idea that self-referential processing promotes a better organization and elaboration of the information (Ganellen & Carver, 1985; Rogers et al., 1977; Symons & Johnson, 1997). Taking into consideration that internal source memories are more likely to be composed of cognitive operations that occurred during encoding (Ferguson, Hastroudi, & Johnson, 1992; Raye & Johnson, 1980), if participants spent more time elaborating on words presented in self-referential conditions (*vs.* non-self-referential conditions), it is plausible to assume that these words were likely associated with more cues and information that could be later used to inform source memory decisions. This might be one of the mechanisms by which self-referential processing is in advantage relative to semantic processing, which might explain the general memory benefit for both item and source performance documented in the studies of the current dissertation. Nonetheless, as no emotion-related electrophysiological modulations were observed during encoding, this same explanation does not fully account for the interactive effects between emotion and encoding tasks that were later identified during the source memory test.

In this context, the analysis of old/new effects during recognition (see Figure 1, Chapter I and Figure 1, Chapter V, for a schematic overview of these effects) may offer some insight. Considering the first time window under analysis (300 - 500 ms), there was a significant differentiation between self-source hits and correct rejections regardless of location and laterality, indicating a general advantage for words encoded self-referentially in an early stage of recognition. In the 500 - 800 ms time window, the maximum mean amplitudes were registered at central and parietal sites, supporting the appearance of the parietal old/new effect. Importantly, whereas old/new effects were present for neutral and positive words for both self and common conditions, in the case of negative words the old/new effect was only present for the self-referential condition but no for the common condition. In the 800 - 1200 ms time window, both self- and common-source hits were different from correct rejections at right frontal locations,

consistent with the appearance of the late frontal old/new effect. Moreover, whereas in the former window no old/new effects were found for the negative common-source hits, in this window negative common-source hits were associated with lower mean amplitudes in contrast to both negative self-source hits and correct rejections, considering middle and right scalp locations. In the subsequent time window of analysis, 1200 - 1800 ms, old/new effects were registered in middle-right frontal and right fronto-central electrodes, also in agreement with the right frontal old/new effect. However, no LPN effects were observed. Taken together, these findings supplement the behavioral data as they indicated that the advantage of self-referential encoding during retrieval was already visible in an early time window, typically associated with familiarity processes. Furthermore, interactive effects between emotion and encoding task emerged in subsequent windows (500 - 800; 800 - 1200 ms), especially considering negative information. In the case of neutral and positive information, besides the typical old/new effects, only one more effect was observed in the 800 - 1200 ms time window, namely the mean amplitude associated with neutral was larger than negative words studied in the common condition. Although the lack of additional ERP effects does not allow us to presume that similar brain processes were engaged during source memory decisions related to neutral and positive stimuli (Fabiani, Gratton, & Federmeier, 2007), the ERP results obtained here with respect to neutral stimuli encoded in the common condition and especially regarding negative stimuli suggest that these conditions imposed a greater challenge to source monitoring processes. In the context of Chapter V, it is referred that the object-based binding theory (Mather, 2007) might provide a tentative explanation for the results in the case of negative stimuli. Specifically, negative information is anticipated to interfere with the retrieval of source features that, during the study phase, were not a core feature of the stimulus, which includes features associated with encoding tasks.

With respect to prospective metamemory judgments of source, the effects of emotion and encoding task on judgments of source (JOSs) were also assessed as a secondary goal of this dissertation. In this context, while some findings were consistent across studies, others were somehow discrepant even considering that similar manipulations of emotion and encoding task were performed. Specifically, one consistent finding was that JOSs were sensitive to encoding task manipulations, in the sense that both the production and self-reference effects were reflected in JOSs ratings. Words that were read aloud (*vs.* read silently) were rated as more memorable (Chapter III, Experiment 1). By the same token, JOSs ratings were superior for words that were encoded self-referentially when compared to words encoded in the common condition (Chapter III, Experiment 2; Chapter IV). Other consistent finding was that immediate JOSs were not particularly predictive of future source memory performance (see also Carroll, Mazzoni,

Andrews, & Pocock, 1999 for a similar finding using a reality monitoring paradigm). On the contrary, whereas in the ERP study (Chapter V), participants believed that they were more likely to remember the source of positive words when compared to both neutral and negative words in the case of the self-referential condition, and when compared to neutral words in the case of the common condition, in the Experiment 2 of Chapter III such pattern of findings was not found. Indeed, only a marginal effect toward positivity (*vs.* neutral and negative words) was documented in this experiment. One could argue that differences in the experimental tasks may account for these discrepancies, as the ERP study included more stimuli, as well as more study-test cycles, when compared to the Experiment 2 of Chapter III. By completing more study-test cycles, the participants had more opportunities to adjust their JOSs throughout the study. Perhaps the participants used their perceived memory performance during recognition tests to guide their JOSs in the subsequent study phases. The necessity to tune JOSs is highlighted by the fact that JOSs ratings might be deemed as an unusual task (Kelly, Carroll, & Mazzoni, 2002). Overall and according to the cue-utilization hypothesis (Koriat, 1997), the results presented here support the notion that prospective metamemory judgments are affected by intrinsic qualities of the stimuli such as valence (intrinsic cues), as well as by the learning conditions such as the encoding task used to study the stimuli (extrinsic cues).

It is worth noting, however, that having more opportunities to do JOSs at different study-test cycles did not translate into a better prediction of later source memory performance. Thus, even though both emotion and encoding task modulated both JOSs and source memory performance, their influence was reflected in distinct behavioral patterns, implying the involvement of different cognitive mechanisms. For example, one factor that might help to explain the tendency to perceive positive information as more memorable is the processing fluency, which can be assessed by the response time during stimulus encoding (see Witherby & Tauber, 2018 for a similar argument). If processing fluency was a determinant factor in the current dissertation, then the response time associated with positive words should be faster than the remaining valence conditions regardless of the encoding task. However, as analyzed in the ERP study (see the Supplementary Materials – Chapter V, Appendix D), this was not the case. Nonetheless, fluency may partially account for the higher JOSs ratings obtained in the case of emotional words studied self-referentially, as the response times associated with these stimuli were faster than neutral stimuli only in self-referential conditions (see Figure S1, Appendix D). Other possible explanation pertains to the perceived congruency between the stimulus and the encoding task as the proportion of matches was higher for positive words in both encoding tasks. This behavioral pattern was reported in both studies (see the Supplementary Materials – Chapter III and V, Appendix B and D, respectively).

Lastly, to provide an overview of the main findings of the current dissertation, they are summarized in Figure 2.

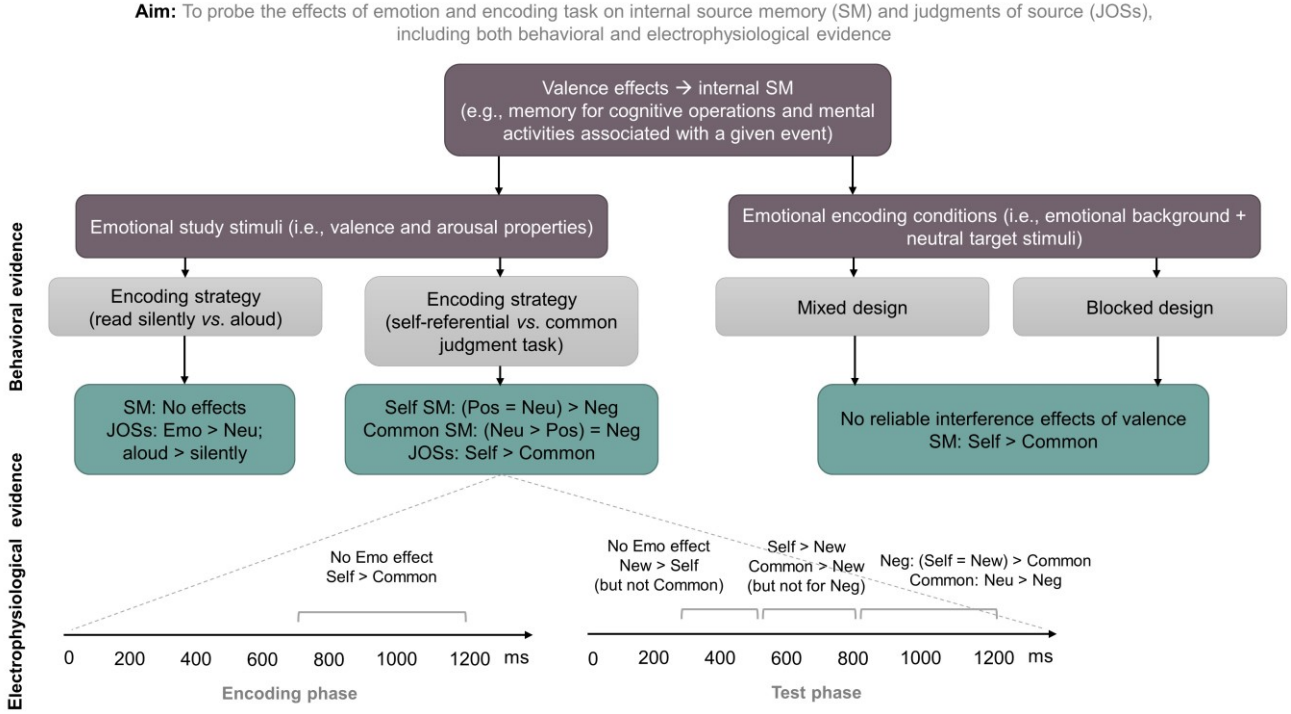


Figure 2. Schematic illustration of the main findings of the studies included in this dissertation. Note: Emo = Emotional; JOSs = Judgments of Source; Neu = Neutral; Neg = Negative; Pos = Positive; SM = Source Memory.

Implications

Regarding possible theoretical implications, we would like to start by highlighting the guidance provided by the source monitoring framework (Johnson, 1997; Johnson, Bush, & Mitchell, 1998; Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981). Throughout this dissertation, different principles of this proposal were adopted, such as the existence of distinct source monitoring processes (i.e., external, internal, and reality monitoring), which are be composed of different features. In fact, the studies presented here were focused on internal source memory that is postulated to integrate features related to cognitive operations and mental activities that occurred during a given study episode. Also, when compared to reality monitoring, internally-derived source features are more likely to overlap and share similarities, hence resulting in challenging source memory decisions (Johnson, 1997; Johnson & Foley, 1984; Raye & Johnson, 1980). Although no other source monitoring processes or comparisons among them were employed in the current dissertation, the former claim finds partial support when observing the average values of source memory accuracy reported here as they were not that high (see Figure S2, Appendix E for an example), especially when compared to item memory accuracy. More

importantly, a relevant tenet of the source monitoring framework is that, besides the evaluation of qualitative features from the study event, source memory also entails the influence of prior knowledge and motivational goals (Johnson et al., 1993). If we are to consider that self-referential encoding conditions provide not only the opportunity to examine the effects of information prioritization through immediate task-related goals (e.g., intentional learning conditions that induce the need to memorize information), but also the opportunity to probe the effects of long-term motivational goals, namely nurturing a positive self-view, then the findings of this dissertation underscore the importance of considering motivational goals when studying source memory. Indeed, there were several instances in which participants showed that source memory-related decisions might be related to a bias toward positive information and a harboring bias against negative information, for example: during encoding the mean proportion of "yes" responses were generally higher for positive words when compared to the other valence categories; also during encoding, JOSs ratings, especially considering the ERP study (Chapter V), revealed that positive self-referent words were perceived as more memorable in contrast to neutral and negative words; during recognition, the source memory accuracy for positive and neutral self-referent words was higher than for negative ones; negative self-referent and new (*vs.* common) information elicited greater post-monitoring processes as indicated by mean amplitudes of the late frontal old/new effect (800 - 1200 ms). Moreover, the fact that only some experimental conditions engaged greater post-monitoring activity while other conditions did not support the notion that source monitoring processes entail not only heuristic but also deliberative processes (Johnson, 1997).

Other implication of this dissertation arises from the theoretical frameworks that have been proposed to account for the synergies between emotion and source memory. From all the theories presented during Chapter II, the object-based binding theory (Mather, 2007) and the arousal-biased competition (ABC) theory (Mather & Sutherland, 2011) were the most used to frame the behavioral evidence obtained here. Specifically, the object-based binding theory (Mather, 2007) provides an explanation for the findings obtained when common encoding conditions were used, in which emotional (*vs.* neutral) information was associated with a detrimental or lack of effect on source memory. As already mentioned, this theory predicts that the memory for extrinsic features, such as the encoding task used during the study phase, is likely to be similar or worse in the case of emotional when compared to neutral information. Nonetheless, this same theory does not accommodate the results when enhancing memory strategies, such as self-referential processing, are employed. Additionally, the distinction between intrinsic and extrinsic features was one of the moderators tested in the context of the meta-analytic review described in Chapter II; however, this factor did not emerge as a statistically significant moderator in any

of the analyses. Moreover, this theory emphasizes arousal as the emotional attribute driving changes on memory processes, when in the current dissertation a deliberate effort was made to examine the role of valence while keeping arousal constant. Notwithstanding, even in these conditions, emotional information affected source memory. In this regard, even though also based on arousal-related effects, the ABC theory (Mather & Sutherland, 2011) represents a more accommodating approach as it suggests that stimuli and features are prioritized as a function of situational demands and current goals of the subjects. In this context, emotion might facilitate memory for information that is perceived as high priority and goal-relevant while hampering memory for information that is low priority and goal-irrelevant. Following the source monitoring framework (Johnson et al., 1993), the former principle also dovetails with the idea that motivational goals play a role in source memory decisions. Nonetheless, it is not possible to draw reliable conclusions about the interactive effects of goal-relevant/goal-irrelevant features and emotional/neutral information on source memory without conducting more studies that overtly address and manipulate these factors (see Ferré et al., 2019 for a similar argument).

Moving on into more practical implications, even though some source memory failures might only cause mild disturbances with no severe implications in our daily life activities (e.g., being wrong about who told us a trivial information), there are situations in which more serious consequences may arise (e.g., confounding the imagined action of turning off the gas with the action itself; Brandt, Bergström, Buda, Henson, & Simons, 2014; see also Mitchell & MacPherson, 2017 for other examples). In this realm, we would like to acknowledge that, as researchers, we intend that the findings from our investigation contribute to a better understanding of the processes under study, as well as the factors that might shape them. Also, we intend that the findings encourage the development of more applied avenues of research or even interventions that respond to day-by-day needs and societal challenges. In this regard, the line of research followed here may be applied to different fields, from which we would like to highlight the following three: affective and mood disorders; auditory hallucinations; aging.

As described earlier in this chapter, distinct information processing biases have been reported in the context of affective and mood disorders. For example, people with depressive symptomology and dysphoria tend to show a processing bias toward negative information (e.g., Herbert, Bendig, & Rojas, 2019; Koster, De Raedt, Goeleven, Franck, & Crombez, 2005; Watson et al., 2008), which is also reflected in a memory advantage for negative self-referent information (e.g., Denny & Hunt, 1992; Mathews & Bradle, 1983; Nasby, 1994; Watson et al., 2008; Zupan et al., 2017). Anxiety disorders are also characterized by a processing bias for negative and threatening information (e.g., Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Mogg, Bradley, Williams, & Mathews,

1993). Even though previous studies have demonstrated that source memory performance is modulated by the presence of depression (e.g., Degl'Innocenti & Bäckman, 1999; Dillon, Dobbins, & Pizzagalli, 2013), the study of source memory and emotion, especially considering self-referential processing conditions, remains less explored. Thus, future studies in this domain might bring valuable contributions to the development of assessment and intervention strategies in the context of affective and mood disorders. For instance, the assessment of source memory could inform which conditions are more susceptible to impairments and which are preserved. Specifically, if in the current dissertation encompassing distinct healthy samples, it was observed that participants demonstrated better source memory for self-referent neutral and positive when compared to negative information, we could hypothesize a reverse pattern in participants in risk for or with depression. Furthermore, due to their temporal resolution, ERPs could be used as markers of change in terms of information processing to monitor intervention-related effects. In this context, the late positive complex (LPC), mentioned in more detail in Chapter V, could be a possible candidate as its amplitude appears to be enhanced for negative self-referent information in depression (*vs.* controls; e.g., Speed, Nelson, Auerbach, Klein, & Hajcak, 2016), being also amenable to change when (re)appraisal strategies targeting negative information are implemented (e.g., Hajcak, Moser, & Simons, 2006; Hajcak & Nieuwenhuis, 2006). These are just some examples, yet they unveil the many avenues of research that may be pursued in future studies.

Auditory hallucinations, more specifically auditory verbal hallucinations, represent one of most prevalent experiences in schizophrenia (Larøi, 2012; Lim et al., 2016). Several models have been proposed to account for these experiences (see Conde, Gonçalves, & Pinheiro, 2016 for a summary), being the reality monitoring account one of them. Indeed, it has been shown that in schizophrenia there is a tendency to attribute self-generated information to externally-generated sources (e.g., Badcock, 2010; Brookwell, Bentall, & Varese, 2013; Fisher, McCoy, Poole, & Vinogradov, 2008; Keefe, Arnold, Bayen, McEvoy, & Wilson, 2002; Pinheiro, Farinha-Fernandes, Roberto, & Kotz, 2019; Pinheiro, Rezaii, Rauber, & Niznikiewicz, 2016; Subramaniam et al., 2012). Patients with schizophrenia, when compared to healthy controls, tend to reveal worse source memory performance with respect to temporal and spatial features (e.g., Brébion, David, Jones, Ohlsen, & Pilowsky, 2007; Diaz-Asper, Malley, Genderson, Apud, & Elvevåg, 2008) and those with auditory hallucinations tend to regard imagined actions as performed in contrast to those without auditory hallucinations and healthy controls (Gaweda, Woodward, Moritz, & Kokoszka, 2013). In this vein, associations between source memory performance and the presence of auditory hallucinations have also been documented (Brébion et al., 2007; Brébion, Ohlsen, Bressan, & David, 2012). The examination of emotion and self-referential processing in the domain of schizophrenia is of

interest as the auditory hallucinations are often negative in content and self-referent, contributing to the impairments and distress experienced in this clinical condition (Freeman & Garety, 2003; Larøi et al., 2019; Nayani & David, 1996). Even though the extant evidence indicates that source memory impairments in schizophrenia occur irrespective of the emotional quality of the stimuli (e.g., Brookwell et al., 2013; Diaz-Asper et al., 2008), more studies tapping different source monitoring processes and source memory features, including self-referential processing, are necessary to provide a clearer picture of the emotion-related source memory functioning in schizophrenia. Once again, this knowledge may guide the development of tailored assessment and intervention strategies to tackle not only source memory bias but also auditory hallucinations. For example, in a study by Subramaniam and colleagues (2012), the authors showed that a computerized cognitive training composed of auditory processing, visual processing, and emotion identification exercises led to improvements in reality monitoring performance accompanied by an increased activity in the medial prefrontal cortex (mPFC), which were not observed in schizophrenia patients that played general computer games (e.g., visuo-spatial puzzles; tetris). Importantly, the neural changes in activity of the mPFC, a region that has been implicated in reality monitoring processes (e.g., Kensinger & Schacter, 2006b; Simons, Henson, Gilbert, & Fletcher, 2008; Turner, Simons, Gilbert, Frith, & Burgess, 2008), were associated with improved social functioning in a follow-up session, six months after the training.

Lastly, during the introductory chapter of this dissertation, it was argued that the development of source memory follows an inverted "U" shape function and so a decline from adulthood to late adulthood is expected (Chua, Chen, & Park, 2006; Dennis et al., 2008; Dulas & Duarte, 2014; Dulas et al., 2011; Glisky, Rubin, & Davidson, 2001). Different non-pharmacological strategies and tools have been proposed to boost episodic memory in the elderly not only to mitigate the impact of age-related memory decline in daily functioning (e.g., Hudes, Rich, Troyer, Yusupov, & Vandermorris, 2019; La Corte, Sperduti, Abichou, & Piolino, 2019), but also to prevent and delay the progression from mild cognitive impairment (MCI) to dementia (e.g., Butler et al., 2018; Irazoki et al., 2020; H. Zhang et al., 2019). If in the previous cases the study of the effects of emotional and self-referential conditions on memory were relevant, the same happens in the case of older adults, even though for different reasons. Specifically, older adults display a superior knowledge about the world in contrast to their younger counterparts (e.g., Park et al., 2002; Salthouse, 2004), as well as a greater ability to regulate their emotions (e.g., Gurera & Isaacowitz, 2019; Lantrip & Huang, 2017; Urry & Gross, 2010). Also, the self-referential effect appears to be reasonably impervious to cognitive and neural age-related alterations, so that it is has been proposed as a potential candidate to enhance memory functioning, including source memory (see Glisky & Marquine, 2009; Grilli

& Glisky, 2011 for a similar argument). Thus, the preservation of both emotional and self-referential processes in the elderly might be used to favor other processes that are subjected to age-related decline. In this regard, it had been shown that self-referential encoding conditions improve the retrieval of details from the study episode in healthy old adults (e.g., Hou et al., 2019a; Hou, Grilli, & Glisky, 2019b; Leshikar et al., 2015), even though this strategy may not be able to increase memory performance to a similar level of the young adults (Dulas et al., 2011). When emotional stimuli were used, both young and old adults revealed better recollection for positive than for negative self-referent stimuli (Leshikar et al., 2015). In the context of MCI, the results have been mixed, as some studies showed no additional benefit of using self-referential (*vs.* semantic) processing conditions when trait adjectives were used, which was not the case for healthy controls (Carson, Shayna Rosenbaum, Moscovitch, & Murphy, 2018). On the contrary, when more ecological materials were explored such as narratives, the beneficial effects of self-referential processing were also observed in participants with MCI, even though an increase in false alarm rates was also documented (Carson, Rosenbaum, Moscovitch, & Murphy, 2019; W. Zhang, Johndro, Budson, & Gutchess, 2020). Taken together, self-referential encoding conditions seem to have positive effects on memory. However, the conditions in which their effects are optimal in both healthy and pathological aging remain to be clarified. In addition, even though self-reference-related benefits have been reported on source memory for spatial (Hou et al., 2019a; Yin et al., 2019) and encoding task features (Dulas et al., 2011; Durbin et al., 2017; Hou et al., 2019b; Leshikar & Duarte, 2012; Leshikar et al., 2015; Mao et al., 2017; Y. Zhang et al., 2018), other types of source features remain less studied. In a recent study by Yin et al., 2019 with young adults, the authors reported that although self-referential encoding conditions benefited memory for spatial features, it was not the case for color-related features. Thus, it might be argued that self-referential processing is useful and valuable in some circumstances, but not in others. For this reason, more studies with healthy and clinical elderly groups are warranted to devise and adopt effective intervention strategies that attend to the needs and characteristics of these groups.

Caveats and future directions

Different limitations were raised throughout this dissertation but, in the current section, we intend to summarize those that are somehow shared across all studies. One of the first challenges faced in this project was related to the process of stimulus selection. Even though many efforts have been made to create, adapt, and validate sets of stimuli to European Portuguese (EP), containing information with respect to different properties, including arousal, valence, frequency of appearance, among others, the available data is still very limited, especially when one is confronted with the necessity to control for

distinct stimulus properties concomitantly. In Portugal, databases such as the EP version of the Affective Norms for English Words (ANEW; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), the EP version of the International Affective Digitized Sounds (IADS-2; Soares et al., 2013), the EP version of the International Affective Picture System (IAPS; Soares et al., 2015), the Minho Word Pool (Soares, Costa, Machado, Comesaña, & Oliveira, 2017), the *Procura-PALavras* (P-PAL; Soares et al., 2010; Soares et al., 2014; Soares et al., 2018) have been valuable contributions, filling a huge gap in the Portuguese context. Nonetheless, more research in this realm is needed. In fact, when we started the studies of this dissertation and considering the tools available at the time, it was not possible to obtain a satisfactory pool of words that complied, for example, with the necessity to control for arousal across valence categories. In face of this situation, we decided to launch a pilot study to enlarge the pool of adjectives available in the EP version of the ANEW (see Appendix C for more details) to obtain a more reasonable pool of stimuli to select from. Although the stimuli used in the context of this dissertation were selected from both the EP version of the ANEW and the former pilot study, in the meanwhile these stimuli were also incorporated in a bigger project that still is in the data collection phase. In this fashion, one of the future ramifications of this dissertation is associated with a contribution to extend the ANEW set. It is worthy to note, however, that this type of investigation is exhaustive, it involves many participants, and it is prolonged in time. Thereby, this leads us to one of the limitations of this dissertation as we were unable to obtain or collect more information concerning other psycholinguistic properties of the words, such as familiarity, concreteness, imageability, inter-item relatedness, age-of-acquisition, among others, meaning that the stimuli used here were not controlled for these properties. Along the same lines, even when pictures were used such in the case of Chapter IV, we did not have enough information to control for physical parameters, including brightness, spatial frequency, color complexity, among others. Despite IAPS being one of the most used databases in the emotion-memory research, some scenes depict old-fashioned and culturally-biased elements, and these might be one more factor interfering with the affective response (Carretié, Tapia, López-Martín, & Albert, 2019).

Other limitation refers to the lack of confirmation of the normative arousal and valence ratings in each sample of participants, as it was partially done in the context of Chapter IV. Even though we followed the most consistently used approach in the literature, some studies had revealed that divergences between the affective ratings provided by the participants and the normative ratings may occur (e.g., Davidson, McFarland, & Glisky, 2006; García-Pacios, Río, Villalobos, Ruiz-Vargas, & Maestú, 2015). There at least two ways to circumvent this issue. One would be to collect individual affective ratings for each stimulus after the experimental task and then study the memory performance as a function of the

individual or sample ratings. When dealing with a large pool of stimuli, it is less feasible to have each participant evaluating each stimulus. Therefore, one possibility could be asking participants to rate only a subset of the stimuli (e.g., Hamann, Ely, Hoffman, & Kilts, 2002). Nevertheless, a possible pitfall of this approach might be the generation of an imbalance across affective categories and, consequently, a reduction in the number of stimuli in certain affective categories, similar to what was verified in Chapter IV. Other possibility would be to collect affective ratings prior to the experimental session and then tailor the stimulus selection for each participant (e.g., Bowen & Kensinger, 2017a, 2017b). Nowadays this can be achieved by online surveys without requiring participants to visit the laboratory. However, to ensure that we have enough stimuli after the individual ratings, this requires a thorough pre-selection to avoid asking participants to rate a copious number of stimuli. A possible drawback of this approach is that the selected stimuli will likely be familiar to the participants during the experimental session, with the possibility of having participants recalling some details from the rating episode. Accordingly, in the case of recognition tests, the list of new stimuli must be derived from other sources, while complying with the criteria already determined for the other stimuli. Even so, a way to bypass the impact of these issues might be to increase, if possible, the time between affective ratings and experimental sessions. Regardless of whether the individual affective ratings are collected before or after the experimental task, one advantage, besides a more rigorous experimental control, is the possibility to specify which information is deemed as neutral for each participant, as in normative ratings the midpoint may represent not only neutrality but also ambivalence of response among participants (Schneider, van Harreveld, Veenstra, Schwarz, & Koole, 2016). Moreover, if individual variations in the affective ratings are likely to influence memory encoding and retrieval, the same might be argued with the respect to the degree of congruency between stimuli and encoding tasks. In this regard, it had been demonstrated that when stimuli match a given category, such as self-descriptiveness or commonness, they are more likely to be remembered when compared to stimuli that mismatch the category (e.g., Glisky & Marquine, 2009; Rogers et al., 1977; Sanquist, Rohrbaugh, Syndulko, & Lindsey, 1980). This might be a confounding factor in our studies and should be addressed in future investigation. Likewise, it had been shown that ERPs are modulated as a function of the self-relevance of the stimuli (Fan et al., 2013; Watson, Dritschel, Obonsawin, & Jentzsch, 2007; Xu et al., 2017), suggesting that self-relatedness properties of the stimuli should be further considered and controlled for in future studies.

With respect to task-related limitations, the immediate delay intervals between study and test phases that characterized the studies included here impose some restrictions to the conclusions of this research project as it is not possible to know if the findings are stable or how they change over time. By

using a short interval, it is possible to acknowledge, at least, that the observed emotion-related effects were likely related to processes taking place during encoding, such as allocation of attentional resources, elaboration, rehearsal (Kensinger & Corkin, 2003); Reisberg, 2006; Reisberg & Heuer, 2004), distinctiveness, and stimulus relatedness (Talmi, 2013; Talmi, Luk, McGarry, & Moscovitch, 2007; Talmi & McGarry, 2012), rather than post-encoding consolidation processes (Quevedo et al., 2003; Schümann, Bayer, Talmi, & Sommer, 2018; Sharot & Yonelinas, 2008; Waring & Kensinger, 2009). Nonetheless, it is not possible to disentangle the contributions of each of the former processes with respect to the results obtained here. Furthermore, few studies have addressed the effects of emotion on source memory considering distinct delay intervals (e.g., Sharot & Yonelinas, 2008; Wang & Fu, 2011; Wang & Ren, 2020; Wang & Sun, 2015). Accordingly, this is an important factor to examine in future studies. Other factor to consider refers to the way source memory measures are computed. Different approaches have been documented in the literature (see Tables S6 and S8, Appendix A, for examples). In principle, these approaches should lead to similar findings but, so far, no study have examined in a systematic manner which measure(s) is/are more adequate to report and interpret source memory performance. This would be a relevant step to promote a more consistent report across studies, also increasing their comparability. In a similar fashion, it is important to recognize that the way emotion is manipulated during encoding varies across studies, which may lead to distinct results and conclusions. In the current dissertation, we included two different approaches, one based on valence/arousal properties of the study stimuli (Chapter III and V) and other based on emotional encoding contexts which were external to the study stimuli (Chapter IV). Yet, there are other methods such as mood induction procedures (e.g., Gingerich & Dodson, 2013; Subramaniam et al., 2016; Subramaniam, Ranasinghe, Mathalon, Nagarajan, & Vinogradov, 2017; Zlomuzica, Preusser, Totzeck, Dere, & Margraf, 2016). In this realm, one prominent challenge is how the current theoretical frameworks may accommodate the diversity of findings in this field. A similar challenge arises when facing the need to integrate different results obtained in the context of emotion effects on different memory processes involving relational binding mechanisms, such as source and associative memory, as already pinpointed by Chiu, Dolcos, Gonsalves, and Cohen, 2013. Other challenge relates to ecological concerns, such as how the materials and source memory tasks mimic the demands required in daily situations. Even though the internal source memory tasks used here might resemble daily activities (e.g., discriminating between something we thought and something we spoke aloud; discriminating between something we thought about ourselves and something we thought about the society in general), it is not possible to ensure that the results obtained here have external validity to be verified in everyday memory situations. In fact, in our daily lives, the learning conditions are usually

incidental, the delay time between learning and remembering is long, the retrieval conditions vary, multiple concurrent sources features must be distinguished, and what is perceived as central/content in a given moment might be considered as peripheral/contextual in the next. Thus, laboratorial experiments, although excellent to isolate the effects of specific variables, cannot encompass all the complexity of everyday source memory decisions.

With regards to participant-based limitations, it is important to note that the participants collaborating in this dissertation can be deemed as members of a WEIRD (western, educated, industrialized, rich, and democratic) society, meaning that we are dealing with a specific subset of participants that is not representative of the general population (Henrich, Heine, & Norenzayan, 2010). Furthermore, even considering the Portuguese context, the samples included here were mostly composed of psychology college students, even though we attempted to recruit students more naïve to psychology and even participants from the general community. These represent other factors constraining the generalization of the findings obtained here. Other limitation to consider is the sample size. In the behavioral studies presented in Chapters III and IV, a priori sample size estimations were computed, given the opportunity to derive reference parameters from previous similar studies. Nonetheless, this was just a general approximation. Moreover, in the case of the ERP and fMRI studies, the samples were small and these studies were likely underpowered to detect some of the small effects that characterize the research in this field, as illustrated in the meta-analytic review of Chapter II and in the subgroup analyses described above. Likewise, it is not possible to rule out that these studies are also underpowered with respect to brain activity. Other participant-related aspects which were not explored in the current dissertation refer to the effects of individual differences (e.g., biological sex; age; personality; genetics) on emotional processing and memory (e.g., Hamann & Canli, 2004), which somehow converge with the abovementioned need to bring individual affective ratings into the equation (e.g., Kuppens, Tuerlinckx, Russell, & Barrett, 2013; Kuppens et al., 2017). Therefore, the role of individual differences could be further pursued in future studies, especially considering that few studies have been conducted in the context of source memory (but see Cansino et al., 2019; D'Argembeau & Van der Linden, 2005; Wang, 2012; Waring, Payne, Schacter, & Kensinger, 2010 for some examples).

To conclude, we would like to add two additional future directions, one associated with the fMRI study and other associated with electrophysiological correlates of JOSs. Regarding the first, even though the interactive effects of valence and self-referential processing on source memory remain relatively unexplored in the literature, these conditions grant us the opportunity to probe the neural mechanisms underlying the source memory benefits for positive and neutral (*vs.* negative) stimuli that are self-

referentially encoded and for neutral (vs. emotional) stimuli encoded in the common condition. Previous works have argued that valence-related effects on memory implies the involvement of controlled and elaborative processes, accompanied by the activation of prefrontal cortex (PFC) and hippocampal regions (Kensinger, 2004; Kensinger & Corkin, 2004). Additionally, negative and positive valence has been associated with distinct neural networks during encoding: negative information engages temporo-occipital regions, whereas positive information involves frontal regions (e.g., Steinmetz & Kensinger, 2009). The mPFC, in its turn, has been implicated not only in memory for both emotional and self-referent information (Gutchess & Kensinger, 2018), but also in source memory for stimuli related to the self during encoding (Leshikar & Duarte, 2012, 2014). Thus, it might be expected that source memory retrieval is accompanied by the activation of multiple brain regions, such as the hippocampus, perirhinal cortex, parahippocampal cortex, and mPFC. Nonetheless, the relative involvement of these and other regions may vary according to the valence and encoding conditions, and this is what we propose to examine in the fMRI study. To illustrate this point and following the binding of item and context (BIC) theory (Diana, Yonelinas, & Ranganath, 2007), when meaningful associations are established between elements during encoding, as it might happen for positive and neutral stimuli related to the self, the retrieval of this association may involve the activation of perirhinal cortex but not the hippocampus or the parahippocampal cortex. This may be a mechanism by which self-referential encoding benefits source memory, even so it requires empirical testing.

Regarding the second, as far as we know, no previous study has explored the ERPs associated with JOSs. Even so, the ERP technique has been already used to identify underlying neural correlates of Judgments of Learning (JOLs) by contrasting, for example, stimuli receiving low *vs*. high JOLs ratings (Skavhaug, Wilding, & Donaldson, 2010) and stimuli receiving prospective metamemory judgments *vs*. a control condition (Müller et al., 2016). Comparisons between distinct affective and encoding conditions remain underexplored in the literature, even so in the ERP study (see the Supplementary Materials – Chapter V, Appendix D), preliminary evidence for these comparisons in the context of JOSs was provided. Specifically, when analyzing for the first time how valence and encoding task modulated the electrophysiological response underlying JOSs, the only difference emerged in the earlier time window (300 - 500 ms) and over parietal electrode sites, suggesting a modulation that likely involves the P300 component. Specifically, positive words studied self-referentially reached a more positive mean amplitude than positive words studied in the common condition. On the one hand, this finding indicates that the increase in JOSs ratings observed for positive self-related words (*vs*. the common condition) was accompanied by a greater allocation of attentional and evaluative resources. On the other hand, the lack

of differences between the remaining experimental conditions suggests one of two possibilities: one is that similar processing resources supported the prospective metamemory judgments, although this claim warrants caution when using the EEG/ERP technique (see Fabiani et al., 2007); other is that the experimental task was not primarily suited to test JOSs-related modulations. Indeed, given that the behavioral data showed that JOSs were sensitive to both valence and encoding task, a plausible explanation is that the electrophysiological responses obtained in this study were not tuned to capture subtle variations on the metamemory judgments. This claim may also be supported by the fact that metamemory decisions were not a central feature of the study, being predictably required after each stimulus encoding. So, it could the case that JOSs-related processes were taking place even during word encoding and prior to the time window dedicated to JOSs. Even so, this was a novel manipulation and future research can build on these findings to devise experiments more sensitive to variations in metamemory judgements.

Concluding remarks

The study of the effects of emotion on source memory had been yielding mixed findings. In this regard, the series of studies developed in the context of this dissertation aimed at contributing to the study of the effects of emotion and encoding task (including self-referential *vs.* semantic processing) on internal source memory attributions, well as on prospective metamemory judgments of source, including both behavioral (e.g., accuracy) and brain activity. Specifically, the encoding tasks adopted here had in mind the principle of enriching memory performance, so that the results would be more easily translatable to situations in which memory is at risk of impairment by age-related factors, incidence of pathological conditions or the overwhelming flood of information we have to deal with on daily basis (e.g., Uncapher, Thieu, & Wagner, 2016; Uncapher & Wagner, 2018), among other factors. Overall, the findings of this dissertation highlight that both source memory performance and JOSs differed as a function of emotion and encoding task, confirming that the way information is encoded plays a role in the explanation of the discrepant findings reported in the emotion-source memory field. Furthermore, self-referential encoding conditions had beneficial effects not only on item memory, but also on source memory performance. Even so, future studies are needed to uncover more potentialities of this encoding strategy beyond internal source memory, considering other source monitoring processes (Johnson et al., 1993) and other source memory features. Ultimately, we hope that the studies of this dissertation encourage further investigations on the factors that shape our memory for details of encoding episodes, expanding our understanding of episodic memory.

References

Badcock, J. C. (2010). The cognitive neuropsychology of auditory hallucinations: A parallel auditory pathways framework. *Schizophrenia Bulletin*, *36*(3), 576–584. https://doi.org/10.1093/schbul/sbn128

Baeken, C., De Raedt, R., Van Schuerbeek, P., De Mey, J., Bossuyt, A., & Luypaert, R. (2012). The influence of emotional priming on the neural substrates of memory: A prospective fMRI study using portrait art stimuli. *NeuroImage*, *61*(4), 876–883. https://doi.org/10.1016/j.neuroimage.2012.03.043

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1-24. httpss://doi.org/10.1037/0033-2909.133.1.1

Blix, I., & Brennen, T. (2011). Intentional forgetting of emotional words after trauma: A study with victims of sexual assault. *Frontiers in Psychology*, *2*, 235. https://doi.org/10.3389/fpsyg.2011.00235

Bock, M. (1986). The influence of emotional meaning on the recall of words processed for form or self-reference. *Psychological Research*, *48*(2), 107–112. https://doi.org/10.1007/BF00309324

Bowen, H. J., & Kensinger, E. A. (2017a). Memory-related functional connectivity in visual processing regions varies by prior emotional context. *NeuroReport*, *28*(13), 808–813. https://doi.org/10.1097/WNR.0000000000000829

Bowen, H. J., & Kensinger, E. A. (2017b). Recapitulation of emotional source context during memory retrieval. *Cortex*, *91*, 142–156. https://doi.org/10.1016/j.cortex.2016.11.004

Brandt, V. C., Bergström, Z. M., Buda, M., Henson, R. N. A., & Simons, J. S. (2014). Did i turn off the gas? Reality monitoring of everyday actions. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 209–219. https://doi.org/10.3758/s13415-013-0189-z

Brébion, G., David, A. S., Jones, H. M., Ohlsen, R., & Pilowsky, L. S. (2007). Temporal context discrimination in patients with schizophrenia: Associations with auditory hallucinations and negative symptoms. *Neuropsychologia*, *45*(4), 817–823. https://doi.org/10.1016/j.neuropsychologia.2006.08.009

Brébion, G., Ohlsen, R. I., Bressan, R. A., & David, A. S. (2012). Source memory errors in schizophrenia, hallucinations and negative symptoms: A synthesis of research findings. *Psychological Medicine*, *42*(12), 2543–2554. https://doi.org/10.1017/S003329171200075X

Bridge, D. J., Chiao, J. Y., & Paller, K. A. (2010). Emotional context at learning systematically biases memory for facial information. *Memory & Cognition*, *38*(2), 125–133.

https://doi.org/10.3758/MC.38.2.125

Brierley, B., Medford, N., Shaw, P., & David, A. S. (2007). Emotional memory for words: Separating content and context. *Cognition and Emotion*, *21*(3), 495–521. https://doi.org/10.1080/02699930600684963

Brookwell, M. L., Bentall, R. P., & Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: A meta-analytic review. *Psychological Medicine*, *43*(12), 2465–2475. https://doi.org/10.1017/S0033291712002760

Butler, M., McCreedy, E., Nelson, V. A., Desai, P., Ratner, E., Fink, H. A., ... Kane, R. L. (2018). Does cognitive training prevent cognitive decline? *Annals of Internal Medicine*, *168*(1), 63–68. https://doi.org/10.7326/M17-1531

Cansino, S., Torres-Trejo, F., Estrada-Manilla, C., Ramírez-Barajas, L., Pérez-Loyda, M., Nava-Chaparro, A., ... Ruiz-Velasco, S. (2019). Predictors of source memory success and failure in older adults. *Frontiers in Aging Neuroscience*, *10*, 17. https://doi.org/10.3389/fnagi.2019.00017

Carretié, L., Tapia, M., López-Martín, S., & Albert, J. (2019). EmoMadrid: An emotional pictures database for affect research. *Motivation and Emotion*, *43*(6), 929–939. https://doi.org/10.1007/s11031-019-09780-y

Carroll, M., Mazzoni, G., Andrews, S., & Pocock, P. (1999). Monitoring the future: Object and source memory for real and imagined events. *Applied Cognitive Psychology*, *13*(4), 373–390. https://doi.org/10.1002/(SICI)1099-0720(199908)13:4<373::AID-ACP605>3.0.CO;2-F

Carson, N., Rosenbaum, R. S., Moscovitch, M., & Murphy, K. J. (2019). Self-referential processing improves memory for narrative information in healthy aging and amnestic Mild Cognitive Impairment. *Neuropsychologia*, *134*, 107179. https://doi.org/10.1016/j.neuropsychologia.2019.107179

Carson, N., Rosenbaum, R., S., Moscovitch, M., & Murphy, K. J. (2018). Self-reference effect and self-reference recollection effect for trait adjectives in amnestic mild cognitive impairment. *Journal of the International Neuropsychological Society*, *24*(8), 821–832. https://doi.org/10.1017/S1355617718000395

Chiu, Y. C., Dolcos, F., Gonsalves, B. D., & Cohen, N. J. (2013). On opposing effects of emotion on contextual or relational memory. *Frontiers in Psychology*, *4*, 103. https://doi.org/10.3389/fpsyg.2013.00103

Chua, H. F., Chen, W., & Park, D. C. (2006). Source memory, aging and culture. *Gerontology*, *52*(5), 306–313. https://doi.org/10.1159/000094612

Conde, T., Gonçalves, O. F., & Pinheiro, A. P. (2016). A cognitive neuroscience view of voice-processing abnormalities in schizophrenia: A window into auditory verbal hallucinations? *Harvard Review of Psychiatry*, *24*(2), 148–163. https://doi.org/10.1097/HRP.0000000000000082

D'Argembeau, A., Comblain, C., & Van der Linden, M. (2005). Affective valence and the self-reference effect: Influence of retrieval conditions. *British Journal of Psychology*, *96*(4), 457–466. https://doi.org/10.1348/000712605X53218

D'Argembeau, A., & Van der Linden, M. (2005). Influence of emotion on memory for temporal information. *Emotion*, *5*(4), 503–507. https://doi.org/10.1037/1528-3542.5.4.503

Davidson, P. S. R., McFarland, C. P., & Glisky, E. L. (2006). Effects of emotion on item and source memory in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 306–322. https://doi.org/10.3758/CABN.6.4.306

Degl'Innocenti, A., & Bäckman, L. (1999). Source memory in major depression. *Journal of Affective Disorders*, *54*(1-2), 205–209. https://doi.org/10.1016/S0165-0327(98)00167-0

Dennis, N. A., Hayes, S. M., Prince, S. E., Madden, D. J., Huettel, S. A., & Cabeza, R. (2008). Effect of aging on the neural correlates of successsful item and source memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 791–808. https://doi.org/10.1037/0278-7393.34.4.791

Denny, E. B., & Hunt, R. R. (1992). Affective valence and memory in depression: Dissociation of recall and fragment completion. *Journal of Abnormal Psychology*, *101*(3), 575–580. https://doi.org/10.1037/0021-843X.101.3.575

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Diaz-Asper, C., Malley, J., Genderson, M., Apud, J., & Elvevåg, B. (2008). Context binding in schizophrenia: Effects of clinical symptomatology and item content. *Psychiatry Research*, *159*(3), 259–270. https://doi.org/10.1016/j.psychres.2007.02.018

Dillon, D. G., Dobbins, I. G., & Pizzagalli, D. A. (2013). Weak reward source memory in depression reflects blunted activation of VTA/SN and parahippocampus. *Social Cognitive and Affective Neuroscience*, *9*(10), 1576–1583. https://doi.org/10.1093/scan/nst155

Dulas, M. R., & Duarte, A. (2014). Aging affects the interactio between attentional control and source memory: An fMRI study. *Journal of Cognitive Neuroscience*, *26*(12), 2653–2669. https://doi.org/10.1162/jocn_a_00663

Dulas, M. R., Newsome, R. N., & Duarte, A. (2011). The effects of aging on ERP correlates of source memory retrieval for self-referential information. *Brain Research*, *1377*, 84–100. https://doi.org/10.1016/j.brainres.2010.12.087

Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory*, *25*(9), 1191–1200. https://doi.org/10.1080/09658211.2017.1282517

Erk, S., Kiefer, M., Grothe, J., Wunderlich, A. P., Spitzer, M., & Walter, H. (2003). Emotional context modulates subsequent memory effect. *NeuroImage*, *18*(2), 439–447. https://doi.org/10.1016/S1053-8119(02)00015-0

Fabiani, M., Gratton, G., & Federmeier, K. D. (2007). Event-related brain potentials: Methods, theory, and applications. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 85–119). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511546396.002

Fan, W., Chen, J., Wang, X. Y., Cai, R., Tan, Q., Chen, Y., … Zhong, Y. (2013). Electrophysiological correlation of the degree of self-reference effect. *PLoS ONE*, *8*(12), 8–13. https://doi.org/10.1371/journal.pone.0080289

Fenker, D. B., Schott, B. H., Richardson-Klavehn, A., Heinze, H. J., & Düzel, E. (2005). Recapitulating emotional context: Activity of amygdala, hippocampus and fusiform cortex during recollection and familiarity. *European Journal of Neuroscience*, *21*(7), 1993–1999. https://doi.org/10.1111/j.1460-9568.2005.04033.x

Ferguson, S. A., Hastroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging*, *7*(3), 443–452. https://doi.org/10.1037/0882-7974.7.3.443

Ferré, P., Comesaña, M., & Guasch, M. (2019). Emotional content and source memory for language: Impairment in an incidental encoding task. *Frontiers in Psychology*, *10*, 65. https://doi.org/10.3389/fpsyg.2019.00065

Fisher, M., McCoy, K., Poole, J. H., & Vinogradov, S. (2008). Self and other in schizophrenia: A cognitive neuroscience perspective. *American Journal of Psychiatry*, *165*(11), 1465–1472. https://doi.org/10.1176/appi.ajp.2008.07111806

Fossati, P., Hevenor, S. J., Lepage, M., Graham, S. J., Grady, C., Keightley, M. L., … Mayberg, H. (2004). Distributed self in episodic memory: Neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage*, *22*(4), 1596–1604. https://doi.org/10.1016/j.neuroimage.2004.03.034

Freeman, D., & Garety, P. A. (2003). Connecting neurosis and psychosis: The direct influence of emotion on delusions and hallucinations. *Behaviour Research and Therapy*, *41*(8), 923–947. https://doi.org/10.1016/S0005-7967(02)00104-3

Gallant, S. N., Spaniol, J., & Yang, L. (2019). Age differences in cue utilization during prospective and retrospective memory monitoring. *Psychology and Aging*, *34*(4), 545–557. https://doi.org/10.1037/pag0000352

Gallant, S. N., & Yang, L. (2014). Positivity effect in source attributions of arousal-matched emotional and non-emotional words during item-based directed forgetting. *Frontiers in Psychology*, *5*, 1334. https://doi.org/10.3389/fpsyg.2014.01334

Ganellen, R. J., & Carver, C. S. (1985). Why does self-reference promote incidental encoding? *Journal of Experimental Social Psychology*, *21*(3), 284–300. https://doi.org/10.1016/0022-1031(85)90021-6

García-Pacios, J., Río, D. Del, Villalobos, D., Ruiz-Vargas, J. M., & Maestú, F. (2015). Emotional interference-based forgetting in short-term memory. Cognitive inhibition of pleasant but not unpleasant biologically relevant distractors. *Frontiers in Psychology*, *6*, 582. https://doi.org/10.3389/fpsyg.2015.00582

Gaweda, Ł., Woodward, T. S., Moritz, S., & Kokoszka, A. (2013). Impaired action self-monitoring in schizophrenia patients with auditory hallucinations. *Schizophrenia Research*, *144*(1–3), 72–79. https://doi.org/10.1016/j.schres.2012.12.003

Gingerich, A. C., & Dodson, C. S. (2013). Sad mood reduces inadvertent plagiarism: Effects of affective state on source monitoring in cryptomnesia. *Motivation and Emotion*, *37*(2), 355–371. https://doi.org/10.1007/s11031-012-9309-2

Glisky, E. L., & Marquine, M. J. (2009). Semantic and self-referential processing of positive and negative trait adjectives in older adults. *Memory*, *17*(2), 144–157. https://doi.org/10.1080/09658210802077405

Glisky, E. L., Rubin, S. R., & Davidson, P. S. R. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1131–1146. https://doi.org/10.1037//0278-7393.27.5.1131

Graf, P., & Yu, M. (2015). Valenced cues and contexts have different effects on event-based prospective memory. *PLoS ONE*, *10*(2), e0116953. https://doi.org/10.1371/journal.pone.0116953

Grilli, M. D., & Glisky, E. L. (2011). The self-imagination effect: Benefits of a self-referential encoding strategy on cued recall in memory-impaired individuals with neurological damage. *Journal of the*

International Neuropsychological Society, *17*(5), 929–933. https://doi.org/10.1017/S1355617711000737

Grilli, M. D., Woolverton, C. B., Crawford, M., & Glisky, E. L. (2018). Self-reference and emotional memory effects in older adults at increased genetic risk of Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 25*(2), 186-199. https://doi.org/10.1080/13825585.2016.1275508

Guo, X., Li, Y., Zhang, Q., Cui, L., & Wei, P. (2018). Weakening of recognition memory caused by task-irrelevant emotional encoding context can be modulated by individuals' inhibitory control. *Personality and Individual Differences*, *134*(105), 201–209. https://doi.org/10.1016/j.paid.2018.06.012

Gurera, J. W., & Isaacowitz, D. M. (2019). Emotion regulation and emotion perception in aging: A perspective on age-related differences and similarities. In *Progress in Brain Research* (1st ed., Vol. 247, pp. 329–351). Elsevier B.V. https://doi.org/10.1016/bs.pbr.2019.02.007

Gutchess, A. H., Kensinger, E. A., Yoon, C., & Schacter, D. L. (2007). Ageing and the self-reference effect in memory. *Memory*, *15*(8), 822–837. https://doi.org/10.1080/09658210701701394

Gutchess, A. H., & Kensinger, E. A. (2018). Shared mechanisms may support mnemonic benefits from self-referencing and emotion. *Trends in Cognitive Sciences*, *22*(8), 712–724. https://doi.org/10.1016/j.tics.2018.05.001

Hajcak, G., Moser, J. S., & Simons, R. F. (2006). Attending to affect: Appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion*, *6*(3), 517–522. https://doi.org/10.1037/1528-3542.6.3.517

Hajcak, G., & Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 291–297. https://doi.org/10.3758/CABN.6.4.291

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinion in Neurobiology*, *14*(2), 233–238. https://doi.org/10.1016/j.conb.2004.03.010

Hamann, S., Ely, T. D., Hoffman, J. M., & Kilts, C. D. (2002). Ecstasy and agony: Activation of the human amygdala in positive and negative emotion. *Psychological Science*, *13*(2), 135–141. https://doi.org/10.1111/1467-9280.00425

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Herbert, C., Bendig, E., & Rojas, R. (2019). My sadness - our happiness: Writing about positive, negative, and neutral autobiographical life events reveals linguistic markers of self-positivity and individual

well-being. *Frontiers in Psychology*, *9*, 2522. https://doi.org/10.3389/fpsyg.2018.02522

Hou, M., Grilli, M. D., & Glisky, E. L. (2019a). Self-Reference enhances memory for multi-element events judged likely to happen in young and older adults. *Memory*, *27*(10), 1451–1461. https://doi.org/10.1080/09658211.2019.1674336

Hou, M., Grilli, M. D., & Glisky, E. L. (2019b). Self-reference enhances relational memory in young and older adults. *Aging, Neuropsychology, and Cognition*, *26*(1), 105–120. https://doi.org/10.1080/13825585.2017.1409333

Howell, G. T., & Zelenski, J. M. (2017). Personality self-concept affects processing of trait adjectives in the self-reference memory paradigm. *Journal of Research in Personality*, *66*, 1–13. https://doi.org/10.1016/j.jrp.2016.12.001

Hudes, R., Rich, J. B., Troyer, A. K., Yusupov, I., & Vandermorris, S. (2019). The impact of memory-strategy training interventions on participant-reported outcomes in healthy older adults: A systematic review and meta-analysis. *Psychology and Aging*, *34*(4), 587–597. https://doi.org/10.1037/pag0000340

Irazoki, E., Contreras-Somoza, L. M., Toribio-Guzmán, J. M., Jenaro-Río, C., Van Der Roest, H., & Franco-Martín, M. A. (2020). Technologies for cognitive training and cognitive rehabilitation for people with mild cognitive impairment and dementia. A systematic review. *Frontiers in Psychology*, *11*, 648. https://doi.org/10.3389/fpsyg.2020.00648

Itoh, M. (2004). Mood-congruent encoding effect in self- and other-referent judgments. *Psychologia*, *47*(3), 145–157. https://doi.org/10.2117/psysoc.2004.145

Johnson, M. K. (1997). Source monitoring and memory distortion. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *352*(1362), 1733–1745. https://doi.org/10.1098/rstb.1997.0156

Johnson, M. K., Bush, J. G., & Mitchell, K. J. (1998). Interpersonal reality monitoring: Judging the sources of other people's memories. *Social Cognition, 16*(2), 199-224. https://doi.org/10.1521/soco.1998.16.2.199

Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues*, *40*(2), 33–50. https://doi.org/10.1111/j.1540-4560.1984.tb01092.x

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85.

https://doi.org/10.1037/0033-295X.88.1.67

Keefe, R. S. E., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophrenia Research*, *57*(1), 51–67. https://doi.org/10.1016/S0920-9964(01)00306-1

Kelley, A. W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794. https://doi.org/10.1162/08989290260138672

Kelly, A., Carroll, M., & Mazzoni, G. (2002). Metamemory and reality monitoring. *Applied Cognitive Psychology*, *16*(4), 407–428. https://doi.org/10.1002/acp.803

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*(4), 241–253. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, *101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., & Schacter, D. L. (2006a). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *The Journal of Neuroscience*, *26*(9), 2564–2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

Kensinger, E. A., & Schacter, D. L. (2006b). Neural processes underlying memory attribution on a reality-monitoring task. *Cerebral Cortex*, *16*(8), 1126–1133. https://doi.org/10.1093/cercor/bhj054

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Koster, E. H. W., De Raedt, R., Goeleven, E., Franck, E., & Crombez, G. (2005). Mood-congruent attentional bias in dysphoria: Maintained attention to and impaired disengagement from negative information. *Emotion, 5*(4), 446-455. https://doi.org/10.1037/1528-3542.5.4.446

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self-other differences. *Journal of Personality and Social Psychology*, *37*(4), 499–514. https://doi.org/10.1037/0022-

3514.37.4.499

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940. https://doi.org/10.1037/a0030811

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The relation between valence and arousal in subjective experience varies with personality and culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

La Corte, V., Sperduti, M., Abichou, K., & Piolino, P. (2019). Episodic memory assessment and remediation in normal and pathological aging using virtual reality: A mini review. *Frontiers in Psychology*, *10*, 173. https://doi.org/10.3389/fpsyg.2019.00173

Lantrip, C., & Huang, J. H. (2017). Cognitive control of emotion in older adults: A review. *Clinical Psychiatry*, *3*(1), 9. https://doi.org/10.21767/2471-9854.100040

Larøi, F. (2012). How do auditory verbal hallucinations in patients differ from those in nonpatients? *Frontiers in Human Neuroscience*, *6*, 25. https://doi.org/10.3389/fnhum.2012.00025

Larøi, F., Thomas, N., Aleman, A., Fernyhough, C., Wilkinson, S., Deamer, F., & McCarthy-Jones, S. (2019). The ice in voices: Understanding negative content in auditory-verbal hallucinations. *Clinical Psychology Review*, *67*, 1–10. https://doi.org/10.1016/j.cpr.2018.11.001

Leshikar, E. D., & Duarte, A. (2012). Medial prefrontal cortex supports memory accuracy for self-referenced items. *Social Neuroscience*, *7*(2), 126–145. https://doi.org/10.1080/17470919.2011.585242

Leshikar, E. D., & Duarte, A. (2014). Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 236–252. https://doi.org/10.3758/s13415-013-0198-y

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Aging, Neuropsychology, and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Lim, A., Hoek, H. W., Deen, M. L., Blom, J. D., Bruggeman, R., Cahn, W., … Wiersma, D. (2016). Prevalence and classification of hallucinations in multiple sensory modalities in schizophrenia spectrum disorders. *Schizophrenia Research*, *176*(2–3), 493–499. https://doi.org/10.1016/j.schres.2016.06.010

Macri, A., Pavard, A., & Versace, R. (2018). The beneficial effect of contextual emotion on memory: The role of integration. *Cognition and Emotion*, *32*(6), 1355–1361.

https://doi.org/10.1080/02699931.2017.1387101

Mao, X., Wang, Y., Wu, Y., & Guo, C. (2017). Self-referential information alleviates retrieval inhibition of directed forgetting effects—an ERP evidence of source memory. *Frontiers in Behavioral Neuroscience*, *11*, 187. https://doi.org/10.3389/fnbeh.2017.00187

Mao, X., You, Y., Li, W., & Guo, C. (2015). Emotion impairs extrinsic source memory-An ERP study. *Biological Psychology*, *110*, 182–189. https://doi.org/10.1016/j.biopsycho.2015.07.005

Martínez-Galindo, J. G., & Cansino, S. (2015). Positive and negative emotional contexts unevenly predict episodic memory. *Behavioural Brain Research*, *291*, 89–102. https://doi.org/10.1016/j.bbr.2015.05.018

Martínez-Galindo, J. G., & Cansino, S. (2017). Emotional context during encoding modulates recognition electrophysiological activity. *Experimental Brain Research*, *235*(1), 169–179. https://doi.org/10.1007/s00221-016-4780-8

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competitition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

Mathews, A., & Bradle, B. (1983). Mood and the self-reference bias in recall. *Behaviour Research and Therapy*, *21*(3), 233–239. https://doi.org/10.1016/0005-7967(83)90204-8

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positiviy bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*(5), 711–747. https://10.1037/0033-2909.130.5.711

Miall, D. S. (1986). Emotion and the self: The context of remembering. *British Journal of Psychology*, *77*(3), 389–397. https://doi.org/10.1111/j.2044-8295.1986.tb02205.x

Mitchell, K. J., & MacPherson, S. E. (2017). The cognitive neuroscience of source memory: Moving the ball forward. *Cortex*, *91*, 1–8. https://doi.org/10.1016/j.cortex.2017.04.010

Mogg, K., Bradley, B. P., Williams, R., & Mathews, A. (1993). Subliminal processing of emotional information in anxiety and depression. *Journal of Abnormal Psychology, 102*(2), 304-311. https://doi.org/10.1037/0021-843X.102.2.304

Morel, N., Villain, N., Rauchs, G., Gaubert, M., Piolino, P., Landeau, B., ... Chételat, G. (2014). Brain activity and functional coupling changes associated with self-reference effect during both

encoding and retrieval. *PLoS ONE*, *9*(3), e90488. https://doi.org/10.1371/journal.pone.0090488

Müller, B. C. N., Tsalas, N. R. H., van Schie, H. T., Meinhardt, J., Proust, J., Sodian, B., & Paulus, M. (2016). Neural correlates of judgments of learning – An ERP study on metacognition. *Brain Research*, *1652*, 170–177. https://doi.org/10.1016/j.brainres.2016.10.005

Nasby, W. (1994). Moderators of mood-congruent encoding: Self-/other-reference and affirmative/nonaffirmative judgement. *Cognition and Emotion*, *8*(3), 259–278. https://doi.org/10.1080/02699939408408941

Nayani, T. H., & David, A. S. (1996). The auditory hallucination: A phenomenological survey. *Psychological Medicine*, *26*(1), 177–189. https://doi.org/10.1017/s003329170003381x

Otani, H., Jaffa, M., Libkuman, T. M., Goernert, P. N., & Kato, K. (2012). Does source memory impairment associated with emotionally arousing stimuli occur at encoding or retrieval? *Journal of Cognitive Psychology*, *24*(7), 796–801. https://doi.org/10.1080/20445911.2012.693468

Otani, H., Libkuman, T. M., Goernert, P. N., Kato, K., Migita, M., Freehafer, S. E., & Landow, M. P. (2012). Emotion, directed forgetting, and source memory. *British Journal of Psychology*, *103*(3), 343–358. https://doi.org/10.1111/j.2044-8295.2011.02078.x

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320. https://doi.org/10.1037/0882-7974.17.2.299

Pierguidi, L., Righi, S., Gronchi, G., Marzi, T., Caharel, S., Giovannelli, F., & Viggiano, M. P. (2016). Emotional contexts modulate intentional memory suppression of neutral faces: Insights from ERPs. *International Journal of Psychophysiology*, *106*, 1–13. https://doi.org/10.1016/j.ijpsycho.2016.05.008

Pinheiro, A. P., Farinha-Fernandes, A., Roberto, M. S., & Kotz, S. A. (2019). Self-voice perception and its relationship with hallucination predisposition. *Cognitive Neuropsychiatry, 24*(4), 237-255, https://doi.org/10.1080/13546805.2019.1621159

Pinheiro, A. P., Rezaii, N., Rauber, A., & Niznikiewicz, M. (2016). Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia. *Cognitive Neuropsychiatry, 21*(4), 335-353. https://doi.org/10.1080/13546805.2016.1208611

Quevedo, J., Sant' Anna, M. K., Madruga, M., Lovato, I., De-Paris, F., Kapczinski, F., ... Cahill, L. (2003). Differential effects of emotional arousal in short- and long-term memory in healthy adults. *Neurobiology of Learning and Memory*, *79*(2), 132–135. https://doi.org/10.1016/S1074-

7427(02)00034-5

Raye, C. L., & Johnson, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society*, *15*(6), 405–408. https://doi.org/10.3758/BF03334572

Reisberg, D. (2006). Memory for emotional episodes: The strengths and limits of arousal-based accounts. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), *Memory and emotion: Interdisciplinary perspectives* (pp. 15-36). Oxford, United Kingdom: Blackwell Publishing.

Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 3-41). New York, USA: Oxford University Press.

Ritchey, M., Wang, S. F., Yonelinas, A. P., & Ranganath, C. (2019). Dissociable medial temporal pathways for encoding emotional item and context information. *Neuropsychologia*, *124*, 66–78. https://doi.org/10.1016/j.neuropsychologia.2018.12.015

Rogers, T. B., Kuiper, N. A, & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

RStudio Team (2018). RStudio: Integrated development for R. RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/

Rudolph, U. (1993). The self-reference effect - methodological issues and implications from a schema-theoretical perspective. *European Journal of Social Psychology*, *23*(4), 331–354. https://doi.org/10.1207/s15328023top2004_11

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140–144. https://doi.org/10.1111/j.0963-7214.2004.00293.x

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsey, D. B. (1980). Electrophysiological signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology*, *17*(6), 568–576. https://doi.org/10.1111/j.1469-8986.1980.tb02299.x

Schneider, I. K., van Harreveld, F., Veenstra, L., Schwarz, N., & Koole, S. L. (2016). Let's not be indifferent about neutrality: Neutral ratings in the international affective picture system (IAPS) mask mixed affective responses. *Emotion*, *16*(4), 426–430. https://doi.org/10.1037/emo0000164

Schümann, D., Bayer, J., Talmi, D., & Sommer, T. (2018). Dissociation of immediate and delayed effects of emotional arousal on episodic memory. *Neurobiology of Learning and Memory*, *148*, 11–19. https://doi.org/10.1016/j.nlm.2017.12.007

Sedikides, C., & Green, J. D. (2000). On the self-protective nature of inconsistency-negativity

management: Using the person memory paradigm to examine self-referent memory. *Journal of Personality and Social Psychology*, *79*(6), 906–922. https://doi.org/10.1037/0022-3514.79.6.906

Sedikides, C., & Green, J. D. (2004). What I don't recall can't hurt me: Information negativity versus information inconsistency as determinants of memorial self-defense. *Social Cognition*, *22*(1), 4–29. https://doi.org/10.1521/soco.22.1.4.30987

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Simons, J. S., Henson, R. N. A., Gilbert, S. J., & Fletcher, P. C. (2008). Separable forms of reality monitoring supported by anterior prefrontal cortex. *Journal of Cognitive Neuroscience*, *20*(3), 447–457. https://doi.org/10.1162/jocn.2008.20036

Skavhaug, I. M., Wilding, E. L., & Donaldson, D. I. (2010). Judgments of learning do not reduce to memory encoding operations: Event-related potential evidence for distinct metacognitive processes. *Brain Research*, *1318*, 87–95. https://doi.org/10.1016/j.brainres.2009.11.047

Soares, A. P., Comesaña, M., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., ... Machado, J. (2010). P-PAL: Uma base lexical com índices psicolinguísticos do Português Europeu [P-PAL: A European Portuguese lexical database]. *Linguamática, 2*(3), 67-72.

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Soares, A. P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. M. (2017). The Minho Word Pool: Norms for imageability, concreteness, and subjective frequency for 3,800 Portuguese words. *Behavior Research Methods*, *49*(3), 1065–1081. https://doi.org/10.3758/s13428-016-0767-4

Soares, A. P., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., França, P., ... Comesaña, M. (2014). Procura-PALavras (P-PAL): A new measure of word frequency for contemporary European Portuguese. *Psicologia: Reflexão e Crítica, 27*(1), 110-123. https://doi.org/10.1590/S0102-79722014000100013

Soares, A. P., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., Machado, J., ...Perea, M. (2018). Procura-PALavras (P-PAL): A web-based interface for a new European Portuguese lexical database. *Behavior Research Methods, 50*(4), 1461-1481. https://doi.org/10.3758/s13428-018-1058-z

Soares, A. P., Pinheiro, A. P., Costa, A., Frade, C. S., Comesaña, M., & Pureza, R. (2013). Affective auditory stimuli: Adaptation of the International Affective Digitized Sounds (IADS-2) for European Portuguese. *Behavior Research Methods, 45*(4), 1168-1181. https://doi.org/10.3758/s13428-012-0310-1

Soares, A. P., Pinheiro, A. P., Costa, A., Frade, C. S., Comesaña, M., & Pureza, R. (2015). Adaptation of the International Affective Picture System (IAPS) for European Portuguese. *Behavior Research Methods, 47*(4), 1159-1177. https://doi.org/10.3758/s13428-014-0535-2

Speed, B. C., Nelson, B. D., Auerbach, R. P., Klein, D. N., & Hajcak, G. (2016). Depression risk and electrocortical reactivity during self-referential emotionanl processing in 8 to 14 year-old girls. *Journal of Abnormal Psychology*, *125*(5), 607–619. https://doi.org/10.1037/abn0000173

Steinmetz, K. R. M., & Kensinger, E. A. (2009). The effects of valence and arousal on the neural activity leading to subsequent memory. *Psychophysiology*, *46*(6), 1190–1199. https://doi.org/10.1111/j.1469-8986.2009.00868.x

Subramaniam, K., Gill, J., Slattery, P., Shastri, A., Mathalon, D. H., Nagarajan, S., & Vinogradov, S. (2016). Neural mechanisms of positive mood induced modulation of reality monitoring. *Frontiers in Human Neuroscience*, *10*, 581. https://doi.org/10.3389/fnhum.2016.00581

Subramaniam, K., Luks, T. L., Fisher, M., Simpson, G. V., Nagarajan, S., & Vinogradov, S. (2012). Computerized cognitive training restores neural activity within the reality monitoring network in schizophrenia. *Neuron*, *73*(4), 842–853. https://doi.org/10.1016/j.neuron.2011.12.024

Subramaniam, K., Ranasinghe, K. G., Mathalon, D., Nagarajan, S., & Vinogradov, S. (2017). Neural mechanisms of mood-induced modulation of reality monitoring in schizophrenia. *Cortex*, *91*, 271–286. https://doi.org/10.1016/j.cortex.2017.01.005

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Talmi, D. (2013). Enhanced emotional memory. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Talmi, D., Luk, B. T. C., McGarry, L. M., & Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *Journal of Memory and Language*, *56*(4), 555–574. https://doi.org/10.1016/j.jml.2007.01.002

Talmi, D., & McGarry, L. M. (2012). Accounting for immediate emotional memory enhancement. *Journal of Memory and Language*, *66*(1), 93–108. https://doi.org/10.1016/j.jml.2011.07.009

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*(1), 1–25.

https://doi.org/10.1146/annurev.psych.53.100901.135114

Turner, M. S., Simons, J. S., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2008). Distinct roles for lateral and medial rostral prefrontal cortex in source monitoring of perceived and imagined events. *Neuropsychologia*, *46*(5), 1442–1453. https://doi.org/10.1016/j.neuropsychologia.2007.12.029

Uncapher, M. R., Thieu, M. K., & Wagner, A. D. (2016). Media multitasking and memory: Differences in working memory and long-term memory. *Psychonomic Bulletin & Review*, *23*(2), 483–490. https://doi.org/10.3758/s13423-015-0907-3

Uncapher, M. R., & Wagner, A. D. (2018). Minds and brains of media multitaskers: Current findings and future directions. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(40), 9889–9896. https://doi.org/10.1073/pnas.1611612115

Urry, H. L., & Gross, J. J. (2010). Emotion regulation in older age. *Current Directions in Psychological Science*, *19*(6), 352–357. https://doi.org/10.1177/0963721410388395

Ventura-Bort, C., Dolcos, F., Wendt, J., Wirkner, J., Hamm, A. O., & Weymar, M. (2017). Item and source memory for emotional associates is mediated by different retrieval processes. *Neuropsychologia*. Advance online publication. https://doi.org/10.1016/j.neuropsychologia.2017.12.015

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wang, B. (2012). Females' superiority in item memory, but not source memory for neutral and emotional Chinese words. *Personality and Individual Differences*, *52*(8), 925–929. https://doi.org/10.1016/j.paid.2012.02.005

Wang, B., & Fu, X. (2011). Time course of effects of emotion on item memory and source memory for Chinese words. *Neurobiology of Learning and Memory*, *95*(4), 415–424. https://doi.org/10.1016/j.nlm.2011.02.001

Wang, B., & Ren, Y. (2020). Time-dependent effects of discrete post-encoding emotions on item memory and source memory. *Memory*, *28*(3), 417–440. https://doi.org/10.1080/09658211.2020.1729384

Wang, B., & Sun, B. (2015). Time-limited effects of emotional arousal on item and source memory. *Quarterly Journal of Experimental Psychology*, *68*(11), 2274–2290. https://doi.org/10.1080/17470218.2015.1013043

Waring, J. D., & Kensinger, E. A. (2009). Effects of emotional valence and arousal upon memory trade-offs with aging. *Psychology and Aging*, *24*(2), 412–422. https://doi.org/10.1037/a0015526

Waring, J. D., Payne, J. D., Schacter, D. L., & Kensinger, E. A. (2010). Impact of individual differences upon emotion-induced memory trade-offs. *Cognition and Emotion*, *24*(1), 150–167. https://doi.org/10.1080/02699930802618918

Watson, L. A., Dritschel, B., Jentzsch, I., & Obonsawin, M. C. (2008). Changes in the relationship between self-reference and emotional valence as a function of dysphoria. *British Journal of Psychology*, *99*(1), 143–152. https://doi.org/10.1348/000712607X248689

Watson, L. A., Dritschel, B., Obonsawin, M. C., & Jentzsch, I. (2007). Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Research*, *1152*(1), 106–110. https://doi.org/10.1016/j.brainres.2007.03.049

Witherby, A. E., & Tauber, S. K. (2018). Monitoring of learning for emotional faces: how do fine-grained categories of emotion influence participants' judgments of learning and beliefs about memory? *Cognition and Emotion*, *32*(4), 860–866. https://doi.org/10.1080/02699931.2017.1360252

Xie, W., & Zhang, W. (2017). Negative emotion enhances mnemonic precision and subjective feelings of remembering in visual long-term memory. *Cognition*, *166*, 73–83. https://doi.org/10.1016/j.cognition.2017.05.025

Xu, K., Li, S., Ren, D., Xia, R., Xue, H., Zhou, A., & Xu, Y. (2017). Importance modulates the temporal features of self-referential processing: An event-related potential study. *Frontiers in Human Neuroscience*, *11*, 470. https://doi.org/10.3389/fnhum.2017.00470

Yang, L., Truong, L., Fuss, S., & Bislimovic, S. (2012). The effects of ageing and divided attention on the self-reference effect in emotional memory: Spontaneous or effortful mnemonic benefits? *Memory*, *20*(6), 596–607. https://doi.org/10.1080/09658211.2012.690040

Yin, X., Ma, Y., Xu, X., & Yang, H. (2019). The effect of self-referencing on memory for different kinds of source information. *Memory*, *27*(4), 519–527. https://doi.org/10.1080/09658211.2018.1532009

Zhang, H., Huntley, J., Bhome, R., Holmes, B., Cahill, J., Gould, R. L., … Howard, R. (2019). Effect of computerised cognitive training on cognitive outcomes in mild cognitive impairment: A systematic review and meta-analysis. *BMJ Open*, *9*, e027062. https://doi.org/10.1136/bmjopen-2018-027062

Zhang, Q., Liu, X., An, W., Yang, Y., & Wang, Y. (2015). Recognition memory of neutral words can be impaired by task-irrelevant emotional encoding contexts: Behavioral and electrophysiological evidence. *Frontiers in Human Neuroscience*, *9*, 73. https://doi.org/10.3389/fnhum.2015.00073

Zhang, W., Johndro, H., Budson, A. E., & Gutchess, A. (2020). Influence of self-referential mode on memory for aMCI patients. *Cognitive Neuropsychology*, *37*(1-2), 46–57. https://doi.org/10.1080/02643294.2019.1657390

Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-Serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology*, *65*(4), 236–244. https://doi.org/10.1027/a000001

Zlomuzica, A., Preusser, F., Totzeck, C., Dere, E., & Margraf, J. (2016). The impact of different emotional states on the memory for what, where and when features of specific events. *Behavioural Brain Research*, *298*, 181–187. https://doi.org/10.1016/j.bbr.2015.09.037

Zupan, Z., Žeželj, I., & AndjelkoviĆ, I. (2017). Memory bias in depression: Effects of self-reference and age. *Journal of Social and Clinical Psychology*, *36*(4), 300–315. https://doi.org/10.1521/jscp.2017.36.4.300

APPENDIX A

SUPPLEMENTARY MATERIALS - CHAPTER II

Table S1

Search Expressions Used in Each Database

Database	Search expression
Web of science	1st search: TS=("source memory" OR "context memory" OR "source monitoring" OR "memory binding" OR "internal source monitoring" OR "external source monitoring" OR "reality monitoring") 2nd search: TS=(emotion* OR affect* OR "arousal" OR valence* OR "mood") 3rd search: combine the former searches with "AND"
Scopus	(TITLE-ABS-KEY("source memory") OR TITLE-ABS-KEY("context memory") OR TITLE-ABS-KEY("source monitoring") OR TITLE-ABS-KEY("memory binding") OR TITLE-ABS-KEY("internal source monitoring") OR TITLE-ABS-KEY("external source monitoring") OR TITLE-ABS-KEY("reality monitoring")) AND (TITLE-ABS-KEY (emotion*) OR TITLE-ABS-KEY(affect*) OR TITLE-ABS-KEY("arousal") OR TITLE-ABS-KEY(valence*) OR TITLE-ABS-KEY("mood"))
Pubmed	(source memory [Title/Abstract] OR context memory [Title/Abstract] OR source monitoring [Title/Abstract] OR memory binding [Title/Abstract] OR internal source monitoring [Title/Abstract] OR external source monitoring [Title/Abstract] OR reality monitoring [Title/Abstract]) AND (emotion* [Title/Abstract] OR affect* [Title/Abstract] OR arousal [Title/Abstract] OR valence* [Title/Abstract] OR mood [Title/Abstract])
PsycInfo (through ProQuest)	ab("source memory" OR "context memory" OR "source monitoring" OR "memory binding" OR "internal source monitoring" OR "external source monitoring" OR "reality monitoring") AND ab(emotion* OR affect* OR "arousal" OR valence* OR "mood")
Psychology and Behavioral Sciences Collection (through EBSCO host)	1st search: "source memory" OR "context memory" OR "source monitoring" OR "memory binding" OR "internal source monitoring" OR "external source monitoring" OR "reality monitoring" 2nd search: emotion* OR affect* OR "arousal" OR valence* OR "mood" 3rd search: combine the former searches with "AND"

Table S2

Formulas Used to Compute the Effect Sizes in the Current Meta-analysis

	Formula
d (equation 11.10; Cumming, 2012)	$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SD_1^2 + SD_2^2}{2}}}$
	Where $\bar{X}_1$ represents the mean of a given valence/arousal condition and SD_1 its standard deviation. The $\bar{X}_2$ and SD_2 represent the same but for the contrast condition.
d_{unb} (equation 11.13; Cumming, 2012)	$\left(1 - \frac{3}{4\,df - 1}\right) \times d$
	Where *d* is the result the former formula and *df* corresponds to the number of participants minus 1 ($n - 1$) as the emotion manipulation in all the studies included here was within-subjects.
v_d (equation 4.28; Borenstein et al., 2009)	$\left(\frac{1}{n} + \frac{d_{unb}^2}{2n}\right) 2(1 - r)$
	Where d_{unb} is obtained as described above, *n* corresponds to number of participants, and *r* is the correlation coefficient between the valence/arousal conditions considered in the analysis.

Table S3

Removed Experiments/Conditions after the Influential Outcomes Analysis

Meta-analysis	Excluded
Negative *vs.* Neutral	Mao et al., 2017 - Experiment 2 (high goal relevance, low familiarity; low goal relevance, high familiarity; low goal relevance, low familiarity)
	Removed = 3
Positive *vs.* Neutral	n/a
	Removed = 0
Negative *vs.* Positive	Zhang et al., 2018 - Experiment 2 (self; other)
	Removed = 2
High *vs.* Low	n/a
	Removed = 0
Medium *vs.* Low	Boywitt, 2005 – Experiment 1 (remember)
	Removed = 1
High *vs.* Medium	Boywitt, 2005 – Experiment 2 (remember)
	Removed = 1

Note. n/a = Not Applicable.

Table S4

Results of the Sensitivity Analysis for each Comparison

	r	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate (d_{unb})	*SE*	*z*	*p*	95% CI [LB, UP]	*Q*, *p*
Neg-Neu	.5	92	52	- 0.11	0.07	- 1.63	.103	- 0.25, 0.02	489.59, < .001
	.3	92	52	- 0.11	0.07	- 1.66	.097	- 0.25, 0.02	349.64, < .001
	.7	92	52	- 0.11	0.07	- 1.59	.112	- 0.25, 0.03	815.69, < .001
Pos-Neu	.5	67	34	- 0.11	0.05	- 2.36	.018	- 0.20, - 0.02	165.55, < .001
	.3	67	34	- 0.11	0.05	- 2.40	.017	- 0.20, - 0.02	118.40, < .001
	.7	67	34	- 0.11	0.05	- 2.34	.019	- 0.21, - 0.02	275.46, < .001
Neg-Pos	.5	75	41	- 0.01	0.05	- 0.29	.775	- 0.10, 0.08	268.93, < .001
	.3	75	41	- 0.01	0.05	- 0.27	.785	- 0.10, 0.08	191.90, < .001
	.7	75	41	- 0.01	0.05	- 0.29	.769	- 0.10, 0.08	447.67, < .001
High-Low	.5	20	14	0.27	0.11	2.42	.016	0.05, 0.48	150.35, < .001
	.3	20	14	0.26	0.11	2.36	.018	0.04, 0.48	106.58, < .001
	.7	20	14	0.27	0.11	2.47	.014	0.06, 0.48	253.54, < .001
Medium-Low	.5	5	3	0.49	0.18	2.65	.008	0.13, 0.85	15.04, .005
	.3	5	3	0.48	0.18	2.67	.008	0.13, 0.84	10.75, .029
	.7	5	3	0.49	0.19	2.64	.008	0.13, 0.86	25.02, < .001
High-Medium	.5	5	3	- 0.12	0.15	- 0.82	.414	- 0.41, 0.17	16.16, .003
	.3	5	3	- 0.13	0.15	- 0.85	.396	- 0.41, 0.16	11.66, .020
	.7	5	3	- 0.12	0.15	- 0.79	.432	- 0.41, 0.17	26.30, < .001

Note. HB = Higher Bound; LB = Lower Bound; Neg = Negative; Neu = Neutral; Pos = Positive.

Table S5

Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based on Valence and Arousal (Organized by Alphabetic Order) – Part I

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* study-test cycles	*N* Study items (total)	Time study (s)
Aupée, 2007	1	Within	Ext	63	33	57.25	n/a	P (IAPS)	Visual	48	16	16	16	2.43	7.51	4.93	5.62	5.36	n/a	N (Neg & Pos > Neu)	Mixed	1	24	6
Blix & Brennen, 2011	1 (to-be-remembered)	Within	Int	23	23	24.08 (CG)	n/a	W	Visual	64	n/a	16	16	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	64	2
	1 (to-be-forgotten)	Within	Int	23	23	24.08 (CG)	n/a	W	Visual	64	n/a	16	16	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	64	2
†Cook et al., 2007	1	Within	RM	34	n/a	n/a	n/a	W (ANEW)	Dual	270	90	90	90	2.31	7.70	5.18	5.89	5.71	5.70	Y	Mixed	1	60	3
	2 (mixed list)	Within	RM	29	n/a	n/a	n/a	W (ANEW)	Dual	270	90	90	90	2.31	7.70	5.18	5.89	5.71	5.70	Y	Mixed	1	60	3
	2 (pure list)	Within	RM	30	n/a	n/a	n/a	W (ANEW)	Dual	270	90	90	90	2.31	7.70	5.18	5.89	5.71	5.70	Y	Pure	1	60	3
	3 (mixed list)	Within	RM	36	n/a	n/a	n/a	W (ANEW)	Dual	270	90	90	90	2.31	7.70	5.18	5.89	5.71	5.70	Y	Mixed	1	60	3
	3 (pure list)	Within	RM	35	n/a	n/a	n/a	W (ANEW)	Dual	270	90	90	90	2.31	7.70	5.18	5.89	5.71	5.70	Y	Pure	1	60	3
	4	Within	RM	33	n/a	n/a	n/a	W (Shaver et al., 1987)	Dual	144	48	80	32	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	72	n/a
	4	Within	Int	35	n/a	n/a	n/a	W (Shaver et al., 1987)	Visual	144	48	80	32	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	72	n/a
†D'Argembeau & Van der Linden, 2004	1 (intentional)	Within	Ext	32	20	22.90	n/a	W (Messina et al., 1989)	Visual	96	24	24	48	≤ 1.90	≥ 4.10	2.70 – 3.30	n/a	n/a	n/a	n/a	Mixed	1	48	2
	1 (incidental source)	Within	Ext	32	20	22.90	n/a	W (Messina et al., 1989)	Visual	96	24	24	48	≤ 1.90	≥ 4.10	2.70 – 3.30	n/a	n/a	n/a	n/a	Mixed	1	48	2
	2 (intentional)	Within	Ext	32	22	21.80	n/a	W (Messina et al., 1989)	Visual	64	16	16	32	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	64	2
	3 (intentional)	Within	Ext	24	17	21.50	n/a	W (Messina et al., 1989)	Visual	96	24	24	48	≤ 1.90	≥ 4.10	2.70 – 3.30	n/a	n/a	n/a	n/a	Mixed	1	48	2

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_f	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
†D'Argembeau & Van der Linden, 2004	3 (incidental source)	Within	Ext	24	17	21.50	n/a	W (Messina et al., 1989)	Visual	96	24	24	48	≤ 1.90	≥ 4.10	2.70 – 3.30	n/a	n/a	n/a	n/a	Mixed	1	48	2
	4 (intentional)	Within	Ext	24	17	22.50	n/a	W (Messina et al., 1989)	Visual	96	24	24	48	≤ 1.90	≥ 4.10	2.70 – 3.30	n/a	n/a	n/a	n/a	Mixed	1	48	2
	4 (incidental source)	Within	Ext	24	17	22.50	n/a	W (Messina et al., 1989)	Visual	96	24	24	48	≤ 1.90	≥ 4.10	2.70 – 3.30	n/a	n/a	n/a	n/a	Mixed	1	48	2
D'Argembeau & Van der Linden, 2005	1 (females)	Within	Ext	24	24	23.70 (all)	n/a	P (IAPS)	Visual	126	42	42	42	2.73	7.40	4.98	5.69	5.39	3.15	N (Neg & Pos > Neu)	Mixed	1	63	2
	1 (males)	Within	Ext	24	0	23.70 (all)	n/a	P (IAPS)	Visual	126	42	42	42	2.73	7.40	4.98	5.69	5.39	3.15	N (Neg & Pos > Neu)	Mixed	1	63	2
†Diaz-Asper et al., 2008	1 (sort)	Within	RM	21	11	39.00 (CG)	16.40	W pairs (ANEW)	Visual	80 (to form pairs)	20	20	40	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	80	18
	1 (session)	Within	Ext	21	11	39.00 (CG)	16.40	W pairs (ANEW)	Visual	80 (to form pairs)	20	20	40	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	80	18
†Doerksen & Shimamura, 2001	1	Within	Ext	24	16	19.60	13.40	W	Visual	128	32	32	64	≤ 2.50	≥ 5.25	> 2.50 & < 5.25	n/a	n/a	n/a	n/a	Mixed	1	64	2
	2	Within	Ext	24	17	19.90	13.30	W	Visual	128	32	32	64	≤ 2.50	≥ 5.25	> 2.50 & < 5.25	n/a	n/a	n/a		Mixed	1	64	2
Dougal et al., 2007	1	Within	Ext	14	9	24.00	n/a	W (ANEW)	Visual	384	96	96	192	2.24	7.79	5.16	6.63	6.48	4.15	N (Neg & Pos > Neu)	Mixed	2	192	4.5
Durbin et al., 2017	1 (self-referential)	Within	Int	24	12	20.60	n/a	W (ANEW)	Visual	432	144	144	144	2.79	7.41	5.10	5.70	5.68	3.84	N (Neg & Pos > Neu)	Mixed	1	216	3

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Durbin et al., 2017	1 (non-self-referential)	Within	Int	24	12	20.60	n/a	W (ANEW)	Visual	432	144	144	144	2.79	7.41	5.10	5.70	5.68	3.84	N (Neg & Pos > Neu)	Mixed	1	216	3
	2 (self-referential)	Within	Int	24	12	20.50	n/a	P (IAPS)	Visual	504	168	168	168	2.90	7.07	5.05	5.34	5.28	3.36	N (Neg & Pos > Neu)	Mixed	1	252	3
	2 (non-self-referential)	Within	Int	24	12	20.50	n/a	P (IAPS)	Visual	504	168	168	168	2.90	7.07	5.05	5.34	5.28	3.36	N (Neg & Pos > Neu)	Mixed	1	252	3
Ferré et al., 2019	1 (spanish)	Within	Int	30	23	22.27	n/a	W (emoFinder and ANEW)	Visual	120	60	n/a	60	2.24	n/a	5.13	6.53	n/a	4.13	N (Neg > Neu)	Mixed	1	60	2
	1 (catalan)	Within	Int	30	23	22.27	n/a	W (emoFinder and ANEW)	Visual	120	60	n/a	60	2.24	n/a	5.13	6.53	n/a	4.13	N (Neg > Neu)	Mixed	1	60	2
	2 (spanish)	Within	Int	30	28	21.37	n/a	W (emoFinder and ANEW)	Visual	120	60	n/a	60	2.24	n/a	5.13	6.53	n/a	4.13	N (Neg > Neu)	Mixed	1	60	n/a
	2 (catalan)	Within	Int	30	28	21.37	n/a	W (emoFinder and ANEW)	Visual	120	60	n/a	60	2.24	n/a	5.13	6.53	n/a	4.13	N (Neg > Neu)	Mixed	1	60	n/a
	3 (spanish)	Within	Int	30	20	24.30	n/a	W (emoFinder and ANEW)	Visual	120	n/a	60	60	n/a	7.46	5.13	n/a	6.02	4.13	N (Pos > Neu)	Mixed	1	60	n/a
	3 (catalan)	Within	Int	30	20	24.30	n/a	W (emoFinder and ANEW)	Visual	120	n/a	60	60	n/a	7.46	5.13	n/a	6.02	4.13	N (Pos > Neu)	Mixed	1	60	n/a
Gallant et al., 2019	1 (to-be-remembered)	Within	Int	36	28	19.25 (YA)	13.40	W (ANEW)	Visual	120	40	40	40	2.64	7.39	5.01	4.45	4.38	4.20	Y	Mixed	1	60	3
	1 (to-be-forgotten)	Within	Int	36	28	19.25 (YA)	13.40	W (ANEW)	Visual	120	40	40	40	2.64	7.39	5.01	4.45	4.38	4.20	Y	Mixed	1	60	3
	2 (gain + 10)	Within	Int	24	22	19.67 (YA)	13.75	W (ANEW)	Visual	120	40	40	40	2.64	7.39	5.01	4.45	4.38	4.20	Y	Mixed	1	60	3
	2 (loss -10)	Within	Int	24	22	19.67 (YA)	13.75	W (ANEW)	Visual	120	40	40	40	2.64	7.39	5.01	4.45	4.38	4.20	Y	Mixed	1	60	3
Gallant & Yang, 2014	1 (to-be-remembered)	Within	Int	36	29	20.22 (YA)	13.90	W (ANEW)	Visual	120	40	40	40	2.64	7.39	5.01	4.45	4.38	4.20	Y	Mixed	1	60	3
	1 (to-be-forgotten)	Within	Int	36	29	20.22 (YA)	13.90	W (ANEW)	Visual	120	40	40	40	2.64	7.39	5.01	4.45	4.38	4.20	Y	Mixed	1	60	3

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Gallo et al., 2011	1 (past autobiographical events)	Within	Int	24	21	20.00	n/a	W (ANEW)	Visual	48	24	24	n/a	2.41	7.60	n/a	6.02	5.97	n/a	Y (Neg = Pos)	Pure	1	24	60
	1 (future autobiographical events)	Within	Int	24	21	20.00	n/a	W (ANEW)	Visual	48	24	24	n/a	2.41	7.60	n/a	6.02	5.97	n/a	Y (Neg = Pos)	Pure	1	24	60
Jia et al., 2018	1	Within	Ext	26	15	22.00	n/a	W (CAWS)	Visual	440	220	n/a	220	3.12	n/a	5.72	6.06	n/a	4.64	N (Neg > Neu)	Mixed	10	200	3
Kaida et al., 2015	1 (total sleep deprivation, immediate)	Within	Ext	14	0	21.40	n/a	P (IAPS)	Visual	180	60	60	60	3.30	6.40	5.20	4.10	4.60	3.30	n/a	Mixed	1	180	2
	1 (total sleep deprivation, 1 week)	Within	Ext	14	0	21.40	n/a	P (IAPS)	Visual	180	60	60	60	3.30	6.40	5.20	4.10	4.60	3.30	n/a	Mixed	1	180	2
	2 (selective REM sleep deprivation, immediate)	Within	Ext	18	0	22.00	n/a	P (IAPS)	Visual	180	60	60	60	4.10	4.60	3.30	4.30	3.60	2.70	n/a	Mixed	1	180	2
	2 (selective REM sleep deprivation, 1 week)	Within	Ext	18	0	22.00	n/a	P (IAPS)	Visual	180	60	60	60	4.10	4.60	3.30	4.30	3.60	2.70	n/a	Mixed	1	180	2
†Kensinger et al., 2011	1 (encoding task, low arousal)	Within	Int	19	11	22.70	15.30	P (Hemera photoset + clipart + images.google.com	Visual	300	60	60	60	3.00	7.10	5.30	3.30	3.10	3.50	Y	Mixed	1	150	2
	1 (encoding task, high arousal)	Within	Ext	19	11	22.70	15.30	P (Hemera photoset + clipart + images.google.com	Visual	300	60	60	60	2.70	7.20	5.30	6.20	6.10	3.50	N (Neg & Pos > Neu)	Mixed	1	150	2
	1 (list, low arousal)	Within	Ext	19	11	22.70	15.30	P (Hemera photoset + clipart + images.google.com	Visual	300	60	60	60	3.00	7.10	5.30	3.30	3.10	3.50	Y	Mixed	1	150	2

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
†Kensinger et al., 2011	1 (list, high arousal)	Within	Ext	19	11	22.70	15.30	P (Hemera photoset + clipart + images.google.com	Visual	300	60	60	60	2.70	7.20	5.30	6.20	6.10	3.50	N (Neg & Pos > Neu)	Mixed	1	150	2
	1 (location, low arousal)	Within	Ext	19	11	22.70	15.30	P (Hemera photoset + clipart + images.google.com	Visual	300	60	60	60	3.00	7.10	5.30	3.30	3.10	3.50	Y	Mixed	1	150	2
	1 (location, high arousal)	Within	Ext	19	11	22.70	15.30	P (Hemera photoset + clipart + images.google.com	Visual	300	60	60	60	2.70	7.20	5.30	6.20	6.10	3.50	N (Neg & Pos > Neu)	Mixed	1	150	2
Kensinger & Corkin, 2003	2	Within	Ext	18	0	27.70	15.70	W (ANEW)	Visual	280	140	n/a	140	1.78 (set A); 1.91 (set B)	n/a	5.83 (set A); 5.65 (set B)	6.95 (set A); 7.16 (set B)	n/a	3.12 (set A); 3.35 (set B)	N (Neg > Neu)	Mixed	1	140	2
	4	Within	Ext	18	0	21.70	13.50	W	Visual	180	60	n/a	60	1.78	n/a	5.22	3.35	n/a	4.51	Y (Neg = Neu	Mixed	1	90	2
Kensinger et al., 2007	1	Within	RM	30	18	21.50 (YA)	14.10	W + P	Visual	180	60	60	60	< 3.8	> 6.2	> 4 & < 6	> 4	> 4	< 5	N (Neg > Pos > Neu)	Mixed	1	120	6
	2	Within	RM	30	17	20.60 (YA)	15.10	W + P	Dual	108	36	36	36	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg & Pos > Neu)	Mixed	1	72	Self-paced (6 max)
Kensinger & Schacter, 2006a	1 (pictures)	Within	Int	21	10	18-35	n/a	P (IAPS + Neu P from unknown source)	Visual	360	120	120	120	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg & Pos > Neu)	Mixed	1	180	n/a
	1 (words)	Within	Int	21	10	18-35	n/a	W (ANEW)	Visual	360	120	120	120	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg & Pos > Neu)	Mixed	1	180	n/a

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Kensinger & Schacter, 2006b	1 (words, incidental)	Within	RM	16	n/a	21.30 (all)	n/a	W (ANEW + from previous study)	Dual	64	24	n/a	24	3.00	n/a	5.80	6.20	n/a	3.90	N (Neg > Neu)	Mixed	1	48	Self-paced (6 max)
	1 (words, intentional)	Within	RM	16	n/a	21.30 (all)	n/a	W (ANEW + from previous study)	Dual	64	24	n/a	24	3.00	n/a	5.80	6.20	n/a	3.90	N (Neg > Neu)	Mixed	1	48	Self-paced (6 max)
	2 (objects)	Within	RM	20	12	21.50	n/a	W + P (objects from clipart)	Visual	450	225	n/a	225	n/a	n/a	n/a	2.96 (W); 3.07 (P)	n/a	0.59 (W); 0.67 (P)	N (Neg > Neu)	Mixed	1	300	4
	3A (words, seen)	Within	RM	16	8	21.90	n/a	W (ANEW + from previous study)	Dual	64	24	n/a	24	3.00	n/a	5.80	6.20	n/a	3.90	N (Neg > Neu)	Mixed	1	48	Self-paced (6 max)
	3A (words, imagine)	Within	RM	16	8	21.90	n/a	W (ANEW + from previous study)	Dual	64	24	n/a	24	3.00	n/a	5.80	6.20	n/a	3.90	N (Neg > Neu)	Mixed	1	48	Self-paced (6 max)
	3B (objects, seen)	Within	RM	16	n/a	21.30	n/a	W + P (objects from clipart)	Visual	450	225	n/a	225	n/a	n/a	n/a	2.96 (W); 3.07 (P)	n/a	0.59 (W); 0.67 (P)	N (Neg > Neu)	Mixed	1	300	4
	3B (objects, imagine)	Within	RM	16	n/a	21.30	n/a	W + P (objects from clipart)	Visual	450	225	n/a	225	n/a	n/a	n/a	2.96 (W); 3.07 (P)	n/a	0.59 (W); 0.67 (P)	N (Neg > Neu)	Mixed	1	300	4
Koenig & Mecklinger, 2008	1 (time)	Within	Ext	20	10	21.00	n/a	P (IAPS)	Visual	120	40	40	40	2.97	7.13	5.09	5.20	5.04	3.43	N (Neg & Pos > Neu)	Mixed	1	120	2
	1 (location)	Within	Ext	20	10	21.00	n/a	P (IAPS)	Visual	120	40	40	40	2.97	7.13	5.09	5.20	5.04	3.43	N (Neg & Pos > Neu)	Mixed	1	120	2
Le Bigot et al., 2018	1 (self-produced)	Within	RM	42	38	20.86	n/a	W (ANEW)	Dual	104	52	n/a	52	2.28	n/a	4.93	5.20	n/a	3.46	N (Neg > Neu)	Mixed	1	104	1.5
	1 (partner-produced)	Within	RM	42	38	20.86	n/a	W (ANEW)	Dual	104	52	n/a	52	2.28	n/a	4.93	5.20	n/a	3.46	N (Neg > Neu)	Mixed	1	104	1.5
	2 (self-produced)	Within	RM	46	46	21.28	n/a	W (ANEW)	Dual	104	n/a	52	52	n/a	7.89	4.93	n/a	6.36	3.46	N (Pos > Neu)	Mixed	1	104	1.5

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Le Bigot et al., 2018	2 (partner-produced)	Within	RM	46	46	21.28	n/a	W (ANEW)	Dual	104	n/a	52	52	n/a	7.89	4.93	n/a	6.36	3.46	N (Pos > Neu)	Mixed	1	104	1.5
Leshikar et al., 2015	1 (self-referential)	Within	Int	24	14	21.25 (YA)	15.20	W (ANEW + Anderson adjective norms)	Visual	288	144	144	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	192	3
	1 (common)	Within	Int	24	14	21.25 (YA)	15.20	W (ANEW + Anderson adjective norms)	Visual	288	144	144	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	192	3
MacKenzie et al., 2015	1	Within	Ext	32	23	19	n/a	P (IAPS)	Visual	60	20	20	20	2.74	7.92	5.27	4.53	4.53	4.53	Y	Mixed	1	30	2
	2	Within	Ext	32	20	21	n/a	P (IAPS)	Visual	60	20	20	20	2.74	7.92	5.27	4.53	4.53	4.53	Y	Mixed	1	30	2
Maddock & Frein, 2009	1A	Within	Ext	48	34	19.77	n/a	W (Maddock et al., 2003)	Visual	36	12	12	12	-2.74	3.04	0.26	7.13	6.98	2.57	N (Neg & Pos > Neu)	Mixed	1	36	Self-paced
	1B	Within	Ext	48	40	19.08	n/a	W (Maddock et al., 2003)	Visual	36	12	12	12	-2.62	2.36	0.22	6.78	6.67	2.62	N (Neg & Pos > Neu)	Mixed	1	36	Self-paced
	3	Within	Ext	48	29	19.90	n/a	W (Maddock et al., 2003)	Visual	36	12	12	12	-2.74	3.04	0.26	7.13	6.98	2.57	N (Neg & Pos > Neu)	Mixed	1	36	Self-paced
	4	Within	Ext	24	17	20.67	n/a	W (Maddock et al., 2003)	Visual	72	24	24	24	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg & Pos > Neu)	Mixed	2	72	Self-paced
Mao et al., 2017	1 (high goal relevance, high salience)	Within	Ext	32	25	18.78	n/a	P (IAPS + CAPS)	Visual	112	56	n/a	56	2.10	n/a	4.93	6.62	n/a	3.50	N (Neg > Neu)	Mixed	1	112	3
	1 (high goal relevance, low salience)	Within	Ext	32	25	18.78	n/a	P (IAPS + CAPS)	Visual	112	56	n/a	56	2.10	n/a	4.93	6.62	n/a	3.50	N (Neg > Neu)	Mixed	1	112	3
	1 (low goal relevance, high salience)	Within	Ext	32	25	18.78	n/a	P (IAPS + CAPS)	Visual	112	56	n/a	56	2.10	n/a	4.93	6.62	n/a	3.50	N (Neg > Neu)	Mixed	1	112	3

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Mao et al., 2017	1 (low goal relevance, low salience)	Within	Ext	32	25	18.78	n/a	P (IAPS + CAPS)	Visual	112	56	n/a	56	2.10	n/a	4.93	6.62	n/a	3.50	N (Neg > Neu)	Mixed	1	112	3
	2 (high goal relevance, high familiarity)	Within	Ext	32	22	19.67	n/a	P (IAPS + CAPS)	Visual	136	68	n/a	56	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg > Neu)	Mixed	1	80	5
	2 (high goal relevance, low familiarity)	Within	Ext	32	22	19.67	n/a	P (IAPS + CAPS)	Visual	136	68	n/a	56	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg > Neu)	Mixed	1	80	5
	2 (low goal relevance, high familiarity)	Within	Ext	32	22	19.67	n/a	P (IAPS + CAPS)	Visual	136	68	n/a	56	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg > Neu)	Mixed	1	80	5
	2 (low goal relevance, low familiarity)	Within	Ext	32	22	19.67	n/a	P (IAPS + CAPS)	Visual	136	68	n/a	56	n/a	n/a	n/a	n/a	n/a	n/a	N (Neg > Neu)	Mixed	1	80	5
Mao et al., 2015	1	Within	Int	17	10	23.40	n/a	P (IAPS + CAPS)	Visual	1080	360	360	360	3.90	2.12	2.84	3.52	3.04	2.62	N (Neg > Pos > Neu)	Mixed	12	720	1.5
Mathews et al., 2013	1	Within	RM	42	28	n/a	n/a	W + P	Visual	216	108	108	n/a	3.72	2.57	n/a	n/a	n/a	n/a	n/a	Mixed	1	216	1.5
	2	Within	RM	51	47	n/a	n/a	W + P	Visual	288	108	108	n/a	4.10	2.60	n/a	n/a	n/a	n/a	n/a	Mixed	1	216	1.5
	3 (affect)	Within	RM	42	32	n/a	n/a	W + P	Visual	216	108	108	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	216	1.5
	3 (size)	Within	RM	42	32	n/a	n/a	W + P	Visual	216	108	108	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	216	1.5
McKague et al., 2012	1	Within	RM	20	14	20.50 (CG)	n/a	W (ANEW)	Audio	192	96	96	n/a	2.63	7.48	n/a	5.60	5.44	n/a	Y (Neg = Pos)	Mixed	1	96	n/a
Minor & Herzmann, 2019	1	Within	Ext	30	23	19.73	n/a	P (IAPS & Google Images)	Visual	360	120	n/a	120	1.94	n/a	3.27	3.28	n/a	1.64	N (Neg > Neu)	Mixed	1	240	1.5
Morrison & Haddock, 1997	1 (errors attributed to the experimenter)	Within	RM	15	4	39.60 (CG)	n/a	W (Clark et al., 1992; Mathews & Klug, 1993)	Audio	24	8	8	8	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	24	n/a
	1 (errors attributed to self)	Within	RM	15	4	39.60 (CG)	n/a	W (Clark et al., 1992; Mathews & Klug, 1993)	Audio	24	8	8	8	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	24	n/a

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Nashiro & Mather, 2011a	1	Within	Ext	24	13	18-29 (YA)	13.56	P (IAPS + unknown source)	Visual	32	8	8	16	n/a	n/a	n/a	> 5	n/a	n/a	N (Neg & Pos > Neu)	Mixed	1	32	5
Nashiro & Mather, 2011b	1	Within	Ext	18	15	18-25 (YA)	14.61	P (IAPS + unknown source)	Visual	32	8	8	16	n/a	n/a	n/a	> 5	> 5	n/a	N (Neg & Pos > Neu)	Mixed	1	32	5
	2	Within	Ext	24	17	18-21 (YA)	12.88	P (IAPS + unknown source)	Visual	16	4	4	8	n/a	n/a	n/a	> 5	> 5	n/a	N (Neg & Pos > Neu)	Mixed	1	16	5
Newsome et al., 2012	1	Within	Int	16	8	25.00 (YA)	16.00	P (IAPS + Google)	Visual	420	140	140	140	4.25	1.69	2.74	3.65	2.50	2.12	N (Neg > Pos > Neu)	Mixed	1	276	3
Novak & Mather, 2009	1	Within	Ext	36	23	18.81	n/a	P (IAPS + unknown source)	Visual	64	32	n/a	32	1.88	n/a	5.61	7.36	n/a	1.77	N (Neg > Neu)	Mixed	1	64	3
Otani, Jaffa et al., 2012	1 (apple kind)	Within	Int	40	29	n/a	n/a	P (IAPS)	Visual	60	30	30	30	2.10 (set 1); 1.75 (set 2)	5.52 (set 1); 6.31 (set 2)	4.41 (set 1); 4.27 (set 2)	3.88 (set 1); 4.65 (set 2)	4.07 (set 1); 4.39 (set 2)	2.34 (set 1); 2.39 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	5
	1 (orange kind)	Within	Int	40	29	n/a	n/a	P (IAPS)	Visual	60	30	30	30	2.10 (set 1); 1.75 (set 2)	5.52 (set 1); 6.31 (set 2)	4.41 (set 1); 4.27 (set 2)	3.88 (set 1); 4.65 (set 2)	4.07 (set 1); 4.39 (set 2)	2.34 (set 1); 2.39 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	5
Otani, Libkuman et al., 2012	1	Within	Int	40	26	n/a	n/a	P (IAPS)	Visual	60	30	30	30	2.10 (set 1); 1.75 (set 2)	5.52 (set 1); 6.31 (set 2)	4.41 (set 1); 4.27 (set 2)	3.88 (set 1); 4.65 (set 2)	4.07 (set 1); 4.39 (set 2)	2.34 (set 1); 2.39 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	5
Pereira et al., 2019	1 (aloud)	Within	Int	28	23	20.25	13.54	W (ANEW)	Visual	180	60	60	60	2.73	7.16	5.07	5.43	5.14	4.01	N (Neg & Pos > Neu)	Mixed	3	90	3
	1 (silent)	Within	Int	28	23	20.25	13.54	W (ANEW)	Visual	180	60	60	60	2.73	7.16	5.07	5.43	5.14	4.01	N (Neg & Pos > Neu)	Mixed	3	90	3

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Pereira et al., 2019	2 (self-referential)	Within	Int	32	30	22.59	14.45	W (ANEW)	Visual	144	48	48	48	2.88	7.37	5.06	4.83	4.73	4.45	Y	Mixed	2	72	Self-paced
	2 (common)	Within	Int	32	30	22.59	14.45	W (ANEW)	Visual	144	48	48	48	2.88	7.37	5.06	4.83	4.73	4.45	Y	Mixed	2	72	Self-paced
Rimmele et al., 2011	1	Within	Ext	25	13	22.58	n/a	P (IAPS)	Visual	120	60	n/a	60	2.88	n/a	5.58	5.62	n/a	3.87	N (Neg > Neu)	Mixed	1	60	4
Rimmele et al., 2012	1	Within	Ext	25	13	22.52	n/a	P (IAPS)	Visual	120	60	n/a	60	2.88	n/a	5.58	5.62	n/a	3.87	N (Neg > Neu)	Mixed	1	60	4
	2	Within	Ext	26	16	24.35	n/a	P (IAPS)	Visual	120	60	n/a	60	2.88	n/a	5.58	5.62	n/a	3.87	N (Neg > Neu)	Mixed	1	60	4
	3	Within	Ext	32	18	23.90	n/a	P (IAPS)	Visual	120	60	n/a	60	2.88	n/a	5.58	5.62	n/a	3.87	N (Neg > Neu)	Mixed	1	60	4
Ritchey et al., 2019	1	Within	Int	22	13	18-32	n/a	P (NAPS)	Visual	264	132	n/a	132	< 4	n/a	4.50 – 6.50	> 5	n/a	< 5	N (Neg > Neu)	Pure (mini-blocks)	1	192	2.5
Robin et al., 2018	1 (seen)	Within	RM	15	15	41.07 (CG)	14.64	W + P (Kensinger et al., 2007 & IAPS & Photolibre)	Visual	84 W + 30 P	28 W + 10 P	28 W + 10 P	28 W + 10 P	2.62 (W); 2.46 (P)	7.05 (W); 7.26 (P)	4.84 (W); 5.02 (P)	6.08 (W); 6.04 (P)	5.98 (W); 5.93 (P)	3.31 (W); 3.04 (P)	N (Neg & Pos > Neu)	Mixed	1	84	3
	1 (imagined)	Within	RM	15	15	41.07 (CG)	14.64	W + P (Kensinger et al., 2007 & IAPS & Photolibre)	Visual	84 W + 30 P	28 W + 10 P	28 W + 10 P	28 W + 10 P	2.62 (W); 2.46 (P)	7.05 (W); 7.26 (P)	4.84 (W); 5.02 (P)	6.08 (W); 6.04 (P)	5.98 (W); 5.93 (P)	3.31 (W); 3.04 (P)	N (Neg & Pos > Neu)	Mixed	1	84	3
Schmidt et al., 2011	1 (time, low arousal)	Within	Ext	24	14	19.70	n/a	P (IAPS)	Visual	540	108	108	54	3.06	6.98	5.08	< 5.00	< 5.00	3.10	Y	Mixed	1	270	3
	1 (time, high arousal)	Within	Ext	24	14	19.70	n/a	P (IAPS)	Visual	540	108	108	54	3.06	6.98	5.08	> 5.50	> 5.50	3.10	N (Neg & Pos > Neu)	Mixed	1	270	3
	1 (location, low arousal)	Within	Ext	24	14	19.70	n/a	P (IAPS)	Visual	540	108	108	54	3.06	6.98	5.08	< 5.00	< 5.00	3.10	Y	Mixed	1	270	3
	1 (location, high arousal)	Within	Ext	24	14	19.70	n/a	P (IAPS)	Visual	540	108	108	54	3.06	6.98	5.08	> 5.50	> 5.50	3.10	N (Neg & Pos > Neu)	Mixed	1	270	3
	2 (time, low arousal)	Within	Ext	24	13	19.70	n/a	P	Visual	75	25	25	25	< 3.00	> 5.00	3.00 – 5.00	< 4.50	< 4.50	< 4.50	Y	Mixed	3	75	7.5

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Schmidt et al., 2011	2 (time, high arousal)	Within	Ext	24	13	19.70	n/a	P	Visual	75	25	25	25	< 3.00	> 5.00	3.00 – 5.00	> 4.50	> 4.50	< 4.50	N (Neg & Pos > Neu)	Mixed	3	75	7.5
	2 (location, low arousal)	Within	Ext	24	13	19.70	n/a	P	Visual	75	25	25	25	< 3.00	> 5.00	3.00 – 5.00	< 4.50	< 4.50	< 4.50	Y	Mixed	3	75	7.5
	2 (location, high arousal)	Within	Ext	24	13	19.70	n/a	P	Visual	75	25	25	25	< 3.00	> 5.00	3.00 – 5.00	> 4.50	> 4.50	< 4.50	N (Neg & Pos > Neu)	Mixed	3	75	7.5
Sharot & Yonelinas, 2008	1 (5 min delay)	Within	Int	24	n/a	18-22	n/a	P (IAPS + Sharot et al., 2004)	Visual	360	180	n/a	180	7.69	n/a	3.75	6.79	n/a	3.03	N (Neg > Neu)	Mixed	1	120	1
	1 (24h delay)	Within	Int	24	n/a	18-22	n/a	P (IAPS + Sharot et al., 2004)	Visual	360	180	n/a	180	7.69	n/a	3.75	6.79	n/a	3.03	N (Neg > Neu)	Mixed	1	120	1
Wang, 2012	1 (females)	Within	Ext	28	28	22.27 (all)	n/a	W	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
	1 (males)	Within	Ext	24	0	22.27 (all)	n/a	W	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
Wang & Fu, 2011	1 (immediate)	Within	Ext	22	16	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
	1 (19 min delay)	Within	Ext	16	10	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Wang & Fu, 2011	1 (63 min delay)	Within	Ext	16	10	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
	1 (4.9 h delay)	Within	Ext	16	10	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
	1 (8.8 h delay)	Within	Ext	16	10	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
	1 (24 h delay)	Within	Ext	16	10	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
	1 (6 day delay)	Within	Ext	16	10	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
	1 (2 week delay)	Within	Ext	16	10	22.91 (all)	n/a	W (Dictionary of modern Chinese)	Visual	120	40	40	40	2.44 (set 1; 2.42 (set 2)	6.31 (set 1); 6.42 (set 2)	5.06 (set 1); 5.05 (set 2)	5.77 (set 1); 5.79 (set 2)	6.02 (set 1); 6.01 (set 2)	4.95 (set 1); 4.95 (set 2)	N (Neg & Pos > Neu)	Mixed	1	60	2
Yick et al., 2015	1	Within	Ext	28	19	22	n/a	P (IAPS + Google images)	Visual	720	360	n/a	360	1.98	n/a	3.21	3.21	n/a	1.95	N (Neg > Neu)	Mixed	1	360	1.5

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_f	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* Neg	*N* Pos	*N* Neu	V Neg (*M*)	V Pos (*M*)	V Neu (*M*)	A Neg (*M*)	A Pos (*M*)	A Neu (*M*)	Control A?	Mixed *vs.* Pure list	*N* Study-test cycles	*N* Study items (total)	Time study (s)
Zhang et al., 2018	1 (self-referential)	Within	Int	39	27	20.18	n/a	W (Chinese pool of personality -traits adjectives)	Visual	80	40	40	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Y (Neg = Pos)	Mixed	1	80	Self-paced
	1 (other-referential)	Within	Int	39	27	20.18	n/a	W (Chinese pool of personality -traits adjectives)	Visual	80	40	40	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Y (Neg = Pos)	Mixed	1	80	Self-paced
	2 (self-referential)	Within	Int	70	54	20.63	n/a	W (Chinese pool of personality -traits adjectives)	Visual	160	40	40	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Y (Neg = Pos)	Mixed	1	80	Self-paced
	2 (other-referential)	Within	Int	70	54	20.63	n/a	W (Chinese pool of personality -traits adjectives)	Visual	160	40	40	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Y (Neg = Pos)	Mixed	1	80	Self-paced

Note. A = Arousal; ANEW = Affective Norms for English Words; CAPS = Chinese Affective Picture System; CAWS = Chinese Affective Word System; CG = Control Group; CR = Cued Recall; edu = Education; emo = Emotional; Ext = External; FA = False Alarms; FR = Free Recall; IAPS = International Affective Picture System; Int = Internal; Mod. = Modality; N = No; n/a = Not Applicable/Available; NAPS = Nencki Affective Picture System; N_f = Number of Females; P = Pictures; Pos = Positive; Neg = Negative; Neu = Neutral; Recog = Recognition; RM = Reality-Monitoring; SM = Source Memory; V = Valence; W = Words; Y = Yes; YA = Young Adults; ↑ = Increased; † = Experiments initially eligible for the meta-analysis and for which information was extracted, but that ended up excluded from the statistical analysis due to unavailable data required for the computation of effect sizes.

Table S6

Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based on Valence and Arousal (Organized by Alphabetic Order) – Part II

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Aupée, 2007	1	Look attentively and memorize	Intentional	Temporal (list discrimination)	2	Intrinsic	15	Recog	Direct	24	6	Pr = p(inclusion) – p(exclusion)	Neu > Neg = Pos	-
Blix & Brennen, 2011	1 (to-be-remembered)	Remember or forget (directed forgetting paradigm)	Intentional	Encoding task (remember *vs.* forget)	2	Extrinsic	n/a	FR	Direct	0	n/a	Number of correct source attributions	Neu > Pos	Only CG; remember > forget
	1 (to-be-forgotten)	Remember or forget (directed forgetting paradigm)	Intentional	Encoding task (remember *vs.* forget)	2	Extrinsic	n/a	FR	Direct	0	n/a	Number of correct source attributions	Neu > Pos	Only CG; remember > forget
†Cook et al., 2007	1	Study a list of words	Incidental	Modality of presentation (seen *vs.* heard)	2	Intrinsic	n/a	Recog	Direct	30	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neu = Pos > Neg	-
	2 (mixed list)	Study a list of words	Incidental	Modality of presentation (seen *vs.* heard)	2	Intrinsic	n/a	Recog	Direct	30	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neu > Neg	-
	2 (pure list)	Study a list of words	Incidental	Modality of presentation (seen *vs.* heard)	2	Intrinsic	n/a	Recog	Direct	30	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neu > Neg	-
	3 (mixed list)	Study a list of words	Incidental	Modality of presentation (seen *vs.* heard)	2	Intrinsic	n/a	Recog	Direct	30	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neu > Pos	-
	3 (pure list)	Study a list of words	Incidental	Modality of presentation (seen *vs.* heard)	2	Intrinsic	n/a	Recog	Direct	30	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neu > Pos	-

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
†Cook et al., 2007	4	Study a list of words	Incidental	Modality of presentation (seen *vs.* heard)	2	Intrinsic	n/a	Recog	Direct	72	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neu > Neg = Pos	-
	4	Anagram or frequency task	Incidental	Encoding task	2	Extrinsic	n/a	Recog	Direct	72	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neu > Neg = Pos	-
†D'Argembeau & Van der Linden, 2004	1 (intentional)	Remember word and its color	Intentional	Color (yellow *vs.* blue)	2	Intrinsic	1	Recog	Indirect (after FR and old/new recog)	48	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) = Neu	No effect: intentional *vs.* incidental
	1 (incidental source)	Remember word	Incidental	Color (yellow *vs.* blue)	2	Intrinsic	1	Recog	Indirect (after FR and old/new recog)	48	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) = Neu	No effect: intentional *vs.* incidental
	2 (intentional)	Remember word and its color	Intentional	Color (yellow *vs.* blue)	2	Intrinsic	5	Recog	Indirect (after FR and old/new recog)	n/a	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) = Neu	-
	3 (intentional)	Remember word and its color	Intentional	Color (yellow *vs.* blue)	2	Intrinsic	1	Recog	Indirect (after old/new recog)	48	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) = Neu	-
	3 (incidental source)	Remember word	Incidental	Color (yellow *vs.* blue)	2	Intrinsic	1	Recog	Indirect (after old/new recog)	48	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) > Neu	-
	4 (intentional)	Remember word and its color	Intentional	Location	4	Intrinsic	1	Recog	Indirect (after old/new recog)	48	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) > Neu	No effect: intentional *vs.* incidental
	4 (incidental source)	Remember word	Incidental	Location	4	Intrinsic	1	Recog	Indirect (after old/new recog)	48	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) > Neu	No effect: intentional *vs.* incidental
D'Argembeau & Van der Linden, 2005	1 (females)	Look at pictures to recognize them later	Intentional	Temporal (list discrimination)	3	Intrinsic	1	Recog	Indirect (after old/new recog)	63	Self-paced	Proportion of old items whose source was correctly identified (z-score)	Neg > Pos = Neu	No sex effect
	1 (males)	Look at pictures to recognize them later	Intentional	Temporal (list discrimination)	3	Intrinsic	1	Recog	Indirect (after old/new recog)	63	Self-paced	Proportion of old items whose source was correctly identified (z-score)	Neg > Pos = Neu	No sex effect

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
†Diaz-Asper et al., 2008	1 (sort)	Remember the pairs, who sorted the words, and the session	Intentional	Identity (experimenter *vs.* participant)	2	Extrinsic	5	Recog	Indirect (after recog if the pair was intact or rearranged)	0	n/a	Proportion of correct source attributions as a function of the total number of pairs correctly identified (Arscin transformation)	Neg = Neu = Pos	Sort > session for Neg and Neu (set 2) items
	1 (session)	Remember the pairs, who sorted the words, and the session	Intentional	Temporal (session)	2	Intrinsic	5	Recog	Indirect (after recog if the pair was intact or rearranged)	0	n/a	Proportion of correct source attributions as a function of the total number of pairs correctly identified (Arscin transformation)	Neg = Neu = Pos	Sort > session for Neg and Neu (set 2) items
† Doerksen & Shimamura, 2001	1	Read each word silently	Intentional	Color (yellow *vs.* blue)	2	Intrinsic	5	Recog	Indirect (after FR)	64	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) > Neu	-
	2	Read each word silently	Intentional	Background color (yellow *vs.* blue)	2	Extrinsic	5	Recog	Indirect (after FR)	64	n/a	Proportion of old items whose source was correctly identified	Emo (Neg & Pos) > Neu	-
Dougal et al., 2007	1	Generate a mental image of a spatial scene containing the study word and its font color	Incidental	Color (red *vs.* blue)	2	Intrinsic	~1260 - 1560	Recog	Indirect (after old/new recog)	192	n/a	Percentage of correct source attributions	Neg = Pos = Neu	-
Durbin et al., 2017	1 (self-referential)	Self-referential or non-self-referential (story in the current news) judgments	Incidental	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	15	Recog	Direct	216	3	Proportion of old items whose source was correctly identified	Pos (self-referential) > Pos (non-self-referential); Neg & Neu (self-referential) = Neg & Neu (non-self-referential)	-
	1 (non-self-referential)	Self-referential or non-self-referential (story in the current news) judgments	Incidental	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	15	Recog	Direct	216	3	Proportion of old items whose source was correctly identified	Pos (self-referential) > Pos (non-self-referential); Neg & Neu (self-referential) = Neg & Neu (non-self-referential)	-
	2 (self-referential)	Self-referential or non-self-referential (overall perceptual quality of the picture) judgments	Incidental	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	15	Recog	Direct	252	3	Proportion of old items whose source was correctly identified	Neg & Neu (non-self-referential) > Neg & Neu (self-referential); Pos (non-self-referential) = Pos (self-referential)	-

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Durbin et al., 2017	2 (non-self-referential)	Self-referential or non-self-referential (overall perceptual quality of the picture) judgments	Incidental	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	15	Recog	Direct	252	3	Proportion of old items whose source was correctly identified	Neg & Neu (non-self-referential) > Neg & Neu (self-referential); Pos (non-self-referential) = Pos (self-referential)	-
Ferré et al., 2019	1 (Spanish)	Memorize the words and language of appearance	Intentional	Encoding task (Spanish *vs.* Catalan)	2	Intrinsic	5.45	Recog	Direct	60	Self-paced	Unbiased hit rates	Neg = Neu	No effect of language
	1 (Catalan)	Memorize the words and language of appearance	Intentional	Encoding task (Spanish *vs.* Catalan)	2	Intrinsic	5.45	Recog	Direct	60	Self-paced	Unbiased hit rates	Neg = Neu	No effect of language
	2 (Spanish)	Read aloud	Incidental	Encoding task (Spanish *vs.* Catalan)	2	Intrinsic	5.45	Recog	Direct	60	Self-paced	Unbiased hit rates	Neu > Neg	SM better for Spanish than Catalan
	2 (Catalan)	Read aloud	Incidental	Encoding task (Spanish *vs.* Catalan)	2	Intrinsic	5.45	Recog	Direct	60	Self-paced	Unbiased hit rates	Neu > Neg	SM better for Spanish than Catalan
	3 (Spanish)	Read aloud	Incidental	Encoding task (Spanish *vs.* Catalan)	2	Intrinsic	5.45	Recog	Direct	60	Self-paced	Unbiased hit rates	Neu > Pos	No effect of language
	3 (Catalan)	Read aloud	Incidental	Encoding task (Spanish *vs.* Catalan)	2	Intrinsic	5.45	Recog	Direct	60	Self-paced	Unbiased hit rates	Neu > Pos	No effect of language
Gallant et al., 2019	1 (to-be-remembered)	Remember or forget (directed forgetting paradigm)	Intentional (but not for source)	Encoding task (remember *vs.* forget)	2	Extrinsic	2	Recog	Direct	60	Self-paced	Proportion of words attributed to the correct source divided by the total number of words correctly recognized as old	Neg = Pos = Neu	No effect of instruction in YA
	1 (to-be-forgotten)	Remember or forget (directed forgetting paradigm)	Intentional (but not for source)	Encoding task (remember *vs.* forget)	2	Extrinsic	2	Recog	Direct	60	Self-paced	Proportion of words attributed to the correct source divided by the total number of words correctly recognized as old	Neg = Pos = Neu	No effect of instruction in YA

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Gallant et al., 2019	2 (loss -10)	Gain (+ 10) or loss (- 10)	Intentional (but not for source)	Encoding task (gain + 10 *vs.* loss - 10)	2	Extrinsic	2	Recog	Direct	60	Self-paced	Proportion of words attributed to the correct source divided by the total number of words correctly recognized as old	Neg = Pos = Neu	Only YA; no effect of gain or loss cue
Gallant & Yang, 2014	1 (to-be-remembered)	Remember or forget (directed forgetting paradigm)	Intentional (but not for source)	Encoding task (remember *vs.* forget)	2	Extrinsic	2	Recog	Direct	60	Self-paced	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neg = Pos = Neu	No effect of instruction in YA
	1 (to-be-forgotten)	Remember or forget (directed forgetting paradigm)	Intentional (but not for source)	Encoding task (remember *vs.* forget)	2	Extrinsic	2	Recog	Direct	60	Self-paced	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neg = Pos = Neu	No effect of instruction in YA
Gallo et al., 2011	1 (past events)	Generate a past or future event associated with cue words	Intentional (generate and write description)	Encoding task (past *vs.* future)	2	Extrinsic	1440	Recog	Indirect (after CR)	24	Self-paced	Proportion of correct source attributions	Neg = Pos	-
	1 (future events)	Generate a past or future event associated with cue words	Intentional (generate and write description)	Encoding task (past *vs.* future)	2	Extrinsic	1440	Recog	Indirect (after CR)	24	Self-paced	Proportion of correct source attributions	Pos > Neg	-
Jia et al., 2018	1	Memorize association between word and background	Intentional	Background shape (square *vs.* circle)	2	Extrinsic	2	Recog	Direct	120	Self-paced	Ratio of the number of correct source items over the number of the total old items	Neu > Neg	Only the retrieval practice condition was considered
Kaida et al., 2015	1 (total sleep deprivation, immediate)	Memorize picture and frame color	Intentional	Background color (yellow *vs.* blue *vs.* green *vs.* red)	4	Extrinsic	10	FR	Indirect (after old/new recog, confidence/ valence / arousal ratings)	60	n/a	Proportion of old items whose source was correctly identified	n/a	-
	1 (total sleep deprivation, 1 week)	Memorize picture and frame color	Intentional	Background color (yellow *vs.* blue *vs.* green *vs.* red)	4	Extrinsic	10080	FR	Indirect (after old/new recog, confidence/ valence / arousal ratings)	60	n/a	Proportion of old items whose source was correctly identified	n/a	-

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Kaida et al., 2015	2 (selective REM sleep deprivation, immediate)	Memorize picture and frame color	Intentional	Background color (yellow *vs.* blue *vs.* green *vs.* red)	4	Extrinsic	10	FR	Indirect (after old/new recog, confidence/ valence / arousal ratings)	60	n/a	Proportion of old items whose source was correctly identified	n/a	-
	2 (selective REM sleep deprivation, 1 week)	Memorize picture and frame color	Intentional	Background color (yellow *vs.* blue *vs.* green *vs.* red)	4	Extrinsic	10080	FR	Indirect (after old/new recog, confidence/ valence / arousal ratings)	60	n/a	Proportion of old items whose source was correctly identified	n/a	-
†Kensinger et al., 2011	1 (encoding task, low arousal)	Answer a variable question about the object	n/a	Encoding task (store *vs.* living *vs.* common *vs.* fit)	4	Extrinsic	30	Recog	Indirect (after old/new recog)	150	n/a	n/a	n/a	SM was not the focus of the study
	1 (encoding task, high arousal)	Answer a variable question about the object	n/a	Encoding task (store *vs.* living *vs.* common *vs.* fit)	4	Extrinsic	30	Recog	Indirect (after old/new recog)	150	n/a	n/a	n/a	SM was not the focus of the study
†Kensinger et al., 2011	1 (list, low arousal)	Answer a variable question about the object	n/a	Temporal (list discrimination)	4	Intrinsic	30	Recog	Indirect (after old/new recog)	150	n/a	n/a	n/a	SM was not the focus of the study
	1 (list, high arousal)	Answer a variable question about the object	n/a	Temporal (list discrimination)	4	Intrinsic	30	Recog	Indirect (after old/new recog)	150	n/a	n/a	n/a	SM was not the focus of the study
	1 (location, low arousal)	Answer a variable question about the object	n/a	Location	4	Intrinsic	30	Recog	Indirect (after old/new recog)	150	n/a	n/a	n/a	SM was not the focus of the study
	1 (location, high arousal)	Answer a variable question about the object	n/a	Location	4	Intrinsic	30	Recog	Indirect (after old/new recog)	150	n/a	n/a	n/a	SM was not the focus of the study
Kensinger & Corkin, 2003	2	Rate as abstract or concrete	Intentional	Color (blue *vs.* red)	2	Intrinsic	15	Recog	Direct	140	Self-paced	Proportion of old items whose source was correctly identified	Neg > Neu	-
	4	Rate as abstract or concrete	Intentional	Color (blue *vs.* red)	2	Intrinsic	15	Recog	Direct	90	Self-paced	Proportion of old items whose source was correctly identified	Neg > Neu	-
Kensinger et al., 2007	1	See or imagine the object represented by each word, and determine if the object is bigger or smaller than a shoebox	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	2880	Recog	Direct	60	n/a	Proportion of old items whose source was correctly identified	Neg > Pos = Neu	Only YA

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Kensinger et al., 2007	2	See or imagine the object represented by each word and a letter-height decision task	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	0	Recog	Direct	36	Self-paced (max 6)	Proportion of old items whose source was correctly identified	Neg > Neu	Only YA
Kensinger & Schacter, 2006a	1 (pictures)	Animate or common task	Incidental	Encoding task (animate *vs.* common)	2	Extrinsic	30	Recog	Direct	180	n/a	Proportion of correct source attributions	Neg = Pos = Neu	-
	1 (words)	Animate or common task	Incidental	Encoding task (animate *vs.* common)	2	Extrinsic	30	Recog	Direct	180	n/a	Proportion of correct source attributions	Neg = Pos = Neu	-
Kensinger & Schacter, 2006b	1 (words, incidental)	Indicate if the first letter is shorter than the last	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	0	Recog	Direct	16	n/a	Proportion of correct source attributions having the visual condition as reference	Neg > Neu	-
	1 (words, intentional)	Indicate if the first letter is shorter than the last	Intentional	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	0	Recog	Direct	16	n/a	Proportion of correct source attributions having the visual condition as reference	Neg > Neu	-
	2 (objects)	Decide if the object would fit inside a shoebox	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	2880	Recog	Direct	150	n/a	Proportion of correct source attributions having the visual condition as reference	Neg > Neu	-
	3A (words, seen)	Indicate if the first letter is shorter than the last	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	0	Recog	Direct	16	n/a	Proportion of old items whose source was correctly identified	Neg > Neu	-
	3A (words, imagine)	Indicate if the first letter is shorter than the last	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	0	Recog	Direct	16	n/a	Proportion of old items whose source was correctly identified	Neg > Neu	-
	3B (objects, seen)	Decide if the object would fit inside a shoebox	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	1440	Recog	Direct	150	n/a	Proportion of old items whose source was correctly identified	Neg > Neu	SM was better for imagined than seen items
	3B (objects, imagine)	Decide if the object would fit inside a shoebox	Incidental	Modality of presentation (seen *vs.* imagine)	2	Extrinsic	1440	Recog	Direct	150	n/a	Proportion of old items whose source was correctly identified	Neg > Neu	SM was better for imagined than seen items

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Koenig & Mecklinger, 2008	1 (time)	Scene depicts something indoors or outdoors	Intentional	Temporal (list discrimination)	2	Intrinsic	2.5	Recog	Indirect (after old/new recog)	120	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neg = Pos = Neu	-
	1 (location)	Scene depicts something indoors or outdoors	Intentional	Location (booth 1 *vs.* booth 2)	2	Intrinsic	2.5	Recog	Indirect (after old/new recog)	120	n/a	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neg = Pos = Neu	-
Le Bigot et al., 2018	1 (self-produced)	Produce a sentence out loud containing a specific word or hear another participant producing a sentence	Incidental	Identity (self-generated *vs.* partner-generated)	2	Intrinsic	2	FR	Indirect	0	n/a	Proportion of old items whose source was correctly identified	Neu > Neg	No effect of identity
	1 (partner-produced)	Produce a sentence out loud containing a specific word or hear another participant producing a sentence	Incidental	Identity (self-generated *vs.* partner-generated)	2	Intrinsic	2	FR	Indirect	0	n/a	Proportion of old items whose source was correctly identified	Neu > Neg	No effect of identity
	2 (self-produced)	Produce a sentence out loud containing a specific word or hear another participant producing a sentence	Incidental	Identity (self-generated *vs.* partner-generated)	2	Intrinsic	2	FR	Indirect	0	n/a	Proportion of old items whose source was correctly identified	Neu > Pos	No effect of identity
	2 (partner-produced)	Produce a sentence out loud containing a specific word or hear another participant producing a sentence	Incidental	Identity (self-generated *vs.* partner-generated)	2	Intrinsic	2	FR	Indirect	0	n/a	Proportion of old items whose source was correctly identified	Neu > Pos	No effect of identity

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Leshikar et al., 2015	1 (self-referential)	Whether the adjective is self-descriptive (self-referential task) or if it is commonly used (common task)	n/a	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	0	Recog	Indirect (after old/new recog)	96	7	p(correct source) - p(incorrect source)	Neg = Pos	Only YA; only recollection responses for SM
	1 (common)	Whether the adjective is self-descriptive (self-referential task) or if it is commonly used (common task)	n/a	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	0	Recog	Indirect (after old/new recog)	96	7	p(correct source) - p(incorrect source)	Neg = Pos	Only YA; only recollection responses for SM
MacKenzie et al., 2015	1	Memorize picture and border color	Intentional	Background color (red *vs.* yellow)	2	Extrinsic	n/a	Recog	Indirect (after old/new recog)	30	3.75	Proportion of correct source attributions	Neu > Neg = Pos	-
	2	Memorize picture and color	Intentional	Color (red *vs.* cyan)	2	Intrinsic	n/a	Recog	Indirect (after old/new recog)	30	3.75	Proportion of correct source attributions	Pos = Neu > Neg	-
Maddock & Frein, 2009	1A	Rate the valence	Incidental	Location (6x7 grid)	18	Intrinsic	n/a	Recog	Direct	0	n/a	Graded scale of points given the distance from the center to the target cell	Pos = Neu > Neg	-
	1B	Rate the valence	Incidental	Location (6x7 grid)	18	Intrinsic	n/a	Recog	Direct	0	n/a	Graded scale of points given the distance from the center to the target cell	Pos = Neu > Neg	= 1A, but used new set of words
	3	Rate the valence	Incidental	Temporal (session)	n/a	Intrinsic	n/a	Recog	Direct	0	n/a	Number of correct responses	Pos = Neu > Neg	-
	4	Rate the valence	Intentional	Location (2x2 grid)	4	Intrinsic	1	Recog	Direct	0	n/a	Number of correct responses	Pos = Neu > Neg	-
Mao et al., 2017	1 (high goal relevance, high salience)	Categorize as living or non-living	Incidental	Background scene	56	Extrinsic	5	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance
	1 (high goal relevance, low salience)	Categorize as living or non-living	Incidental	Background scene	56	Extrinsic	5	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance
	1 (low goal relevance, high salience)	Categorize as natural or manufactured	Incidental	Background scene	56	Extrinsic	5	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Mao et al., 2017	1 (low goal relevance, low salience)	Categorize as natural or manufactured	Incidental	Background scene	56	Extrinsic	5	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance
	2 (high goal relevance, high familiarity)	Categorize as living or non-living	Incidental	Background scene	56	Extrinsic	3	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance (low familiarity); high familiarity > low familiarity (low goal relevance)
	2 (high goal relevance, low familiarity)	Categorize as living or non-living	Incidental	Background scene	56	Extrinsic	3	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance (low familiarity); high familiarity > low familiarity (low goal relevance)
	2 (low goal relevance, high familiarity)	Categorize as natural or manufactured	Incidental	Background scene	56	Extrinsic	3	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance (low familiarity); high familiarity > low familiarity (low goal relevance)

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Mao et al., 2017	2 (low goal relevance, low familiarity)	Categorize as natural or manufactured	Incidental	Background scene	56	Extrinsic	3	Recog	Indirect (after old/new recog)	56	Self-paced	n/a	Neu > Neg	Low goal relevance > high goal relevance (low familiarity); high familiarity > low familiarity (low goal relevance)
Mao et al., 2015	1	Respond if the stimuli belong to the scene or people category	Intentional	Encoding task (scene *vs.* people)	2	Extrinsic	2	Recog	Direct	360	Self-paced	Number of correct responses with remember or know responses over the number of hits	Neu > Pos = Neg	Based on recollection not familiarity responses
Mathews et al., 2013	1	Look, imagine or generate a sentence	Incidental	Modality of presentation (seen *vs.* imagine *vs.* verbal description)	3	Extrinsic	n/a	Recog	Direct	0	n/a	Number of correct responses (seen pictures)	Neg > Pos	Hits only for pictures as the SM test was related to them
	2	Look, imagine or generate a sentence	Incidental	Modality of presentation (seen *vs.* imagine *vs.* verbal description)	3	Extrinsic	n/a	Recog	Direct	72	n/a	Number of correct responses (seen pictures)	Neg > Pos	Hits only for pictures as the SM test was related to them
	3 (affect)	Size or feel judgments	Incidental	Modality of presentation (seen *vs.* imagine *vs.* verbal description)	3	Extrinsic	n/a	Recog	Direct	0	n/a	Number of correct responses (seen pictures)	Neg > Pos	Hits only for pictures as the SM test was related to them
	3 (size)	Size or feel judgments	Incidental	Modality of presentation (seen *vs.* imagine *vs.* verbal description)	3	Extrinsic	n/a	Recog	Direct	0	n/a	Number of correct responses (seen pictures)	Neg > Pos	Hits only for pictures as the SM test was related to them
McKague et al., 2012	1	Generate or hear a sentence containing a word	Incidental	Identity (self-generated *vs.* other-generated)	2	Extrinsic	0	Recog	Indirect (after old/new recog)	96	n/a	Sensitivity score (d' = z(hits) – z(FA))	Pos > Neg	Only CG

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Minor & Herzmann, 2019	1	Report if approach or avoid the situation depicted in the picture in real-life	Intentional	Temporal (block)	2	Intrinsic	1440	Recog	Indirect (after old/new recog)	120	Self-paced	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neg = Neu	-
Morrison & Haddock, 1997	1 (errors attributed to the experimenter)	Say the first word that comes to mind	n/a	Identity (self-generated *vs.* experimenter-generated)	2	Intrinsic	n/a	Recog	Direct	0	n/a	Square-root transformation of the number of errors	↑ errors for Pos	↑ errors for this condition; only CG
	1 (errors attributed to self)	Say the first word that comes to mind	n/a	Identity (self-generated *vs.* experimenter-generated)	2	Intrinsic	n/a	Recog	Direct	0	n/a	Square-root transformation of the number of errors	↑ errors for Neg	Only CG
Nashiro & Mather, 2011a	1	Different tasks: is the dot blue or red?; is the picture higher or lower than the shape?	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Indirect (after FR)	0	n/a	Proportion of correct source attributions	Neg = Pos	Only YA; only comparison between Neg and Pos
Nashiro & Mather, 2011b	1	Different tasks: is the dot blue or red?; is the picture higher or lower than the shape?	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Indirect (after FR)	0	n/a	Proportion of correct source attributions	Neg = Pos	Only YA; only comparison between Neg and Pos
	2	Different tasks: is the dot blue or red?; is the picture higher or lower than the shape?; is the photograph bigger or smaller than the shape?	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Indirect (after old/new recog)	16	n/a	Proportion of correct source attributions	Neg = Pos	Only YA; only comparison between Neg and Pos
Newsome et al., 2012	1	Common (seen within a typical month) or indoors judgment	Incidental	Encoding task (common *vs.* indoors)	2	Extrinsic	0	Recog	Direct	144	4	p(correct source) - p(incorrect source)	Neu = Pos > Neg	Only YA
Novak & Mather, 2009	1	Learn picture-location conjunctions	Intentional	Location (3x3 grid)	8	Intrinsic	0	FR	Direct	0	n/a	Number of errors made in the source identification	↑ errors for Neu	Only considering data from the first block

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Otani, Jaffa et al., 2012	1 (apple kind)	Remember picture and encoding task (apple or orange kind)	Intentional	Encoding task (apple *vs.* orange)	2	Extrinsic	2	FR	Indirect (after FR)	0	n/a	Conditionalized misattributions errors (i.e., attribute apple to orange or vice-versa)	Errors: Neg > Neu	No effect of encoding task
	1 (orange kind)	Remember picture and encoding task (apple or orange kind)	Intentional	Encoding task (apple *vs.* orange)	2	Extrinsic	2	FR	Indirect (after FR)	0	n/a	Conditionalized misattributions errors (i.e., attribute apple to orange or vice-versa)	Errors: Neg > Neu	No effect of encoding task
Otani, Libkuman et al., 2012	1	Remember or forget (directed forgetting paradigm)	Intentional	Encoding task (remember *vs.* forget)	2	Extrinsic	2	FR	Indirect (after FR)	0	n/a	Conditionalized misattributions errors (i.e., attribute remember to forget or vice-versa)	Errors: Neg > Neu	No effect of encoding task
Pereira et al., 2019	1 (aloud)	Memorize word and study instruction (aloud or silent)	Intentional	Encoding task (aloud *vs.* silent)	2	Extrinsic	0	Recog	Direct	90	Self-paced	p(correct source) - p(incorrect source)	Neg = Pos = Neu	No effect of encoding task
	1 (silent)	Memorize word and study instruction (aloud or silent)	Intentional	Encoding task (aloud *vs.* silent)	2	Extrinsic	0	Recog	Direct	90	Self-paced	p(correct source) - p(incorrect source)	Neg = Pos = Neu	No effect of encoding task
	2 (self-referential)	Memorize word and encoding task (self-referential or common)	Intentional	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	0	Recog	Direct	72	Self-paced	p(correct source) - p(incorrect source)	Pos = Neu > Neg	Self-referential > common task
	2 (common)	Memorize word and encoding task (self-referential or common)	Intentional	Encoding task (self-referential *vs.* non-self-referential)	2	Extrinsic	0	Recog	Direct	72	Self-paced	p(correct source) - p(incorrect source)	Neu > Neg = Pos	Self-referential > common task
Rimmele et al., 2011	1	Judge if the frame color appears in the scene or not	Incidental	Background color of the dots (yellow *vs.* red *vs.* blue *vs.* green)	4	Extrinsic	60	Recog	Indirect (after old/new recog)	60	Self-paced	Proportion of old items whose source was correctly identified	Neu > Neg	-
Rimmele et al., 2012	1	Judge if the color of dots appears elsewhere in the picture	Incidental	Color of the dots (yellow *vs.* red *vs.* blue *vs.* green)	4	Extrinsic	60	Recog	Indirect (after old/new recog)	60	Self-paced	Proportion of old items whose source was correctly identified	Neu > Neg	-
	2	Indicate the corner in which the scene is located	Incidental	Location	4	Intrinsic	60	Recog	Indirect (after old/new recog)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg > Neu	-
	3	Watch the pictures	Incidental	Temporal (block)	3	Intrinsic	n/a (variable)	Recog	Indirect (after old/new recog)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg > Neu	-

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Ritchey et al., 2019	1	Photographer or San Francisco question	Incidental	Encoding task (photographer or San Francisco question)	2	Extrinsic	1440	Recog	Indirect (after remember/ know)	72	n/a	Proportion of correct source attributions	Neg = Neu	-
Robin et al., 2018	1 (seen)	Indicate on a 4-point scale whether the photograph/ imagined image is an appropriate representation of the word	Incidental	Modality of presentation (seen *vs.* imagined)	2	Extrinsic	10080	Recog	Indirect (after confidence ratings)	24	Self-paced	Proportion of correct source attributions divided by the total number of words whose answer was 3 or 4 on the confidence scale	Neg = Pos > Neu	No effect of modality; only CG
	1 (imagined)	Indicate on a 4-point scale whether the photograph/ imagined image is an appropriate representation of the word	Incidental	Modality of presentation (seen *vs.* imagined)	2	Extrinsic	10080	Recog	Indirect (after confidence ratings)	24	Self-paced	Proportion of correct source attributions divided by the total number of words whose answer was 3 or 4 on the confidence scale	Neg = Pos > Neu	No effect of modality; only CG
Schmidt et al., 2011	1 (time, low arousal)	Different encoding tasks: living or common or approach	Incidental	Temporal (list discrimination)	3	Intrinsic	0	Recog	Indirect (after old/new recog; remember/ know)	270	Self-paced	n/a	Neg & Pos (high arousal) > Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	List > location task
	1 (time, high arousal)	Different encoding tasks: living or common or approach	Incidental	Temporal (list discrimination)	3	Intrinsic	0	Recog	Indirect (after old/new recog; remember/ know)	270	Self-paced	n/a	Neg & Pos (high arousal) > Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	List > location task
	1 (location, low arousal)	Different encoding tasks: living or common or approach	Incidental	Location	3	Intrinsic	0	Recog	Indirect (after old/new recog; remember/ know; list discrimination)	270	Self-paced	n/a	Neg & Pos (high arousal) > Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	List > location task
	1 (location, high arousal)	Different encoding tasks: living or common or approach	Incidental	Location	3	Intrinsic	0	Recog	Indirect (after old/new recog; remember/ know; list discrimination)	270	Self-paced	n/a	Neg & Pos (high arousal) > Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	List > location task

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Schmidt et al., 2011	2 (time, low arousal)	Remember background and objects, and formulate a short story about the scene	Intentional	Temporal (order of appearance)	3	Intrinsic	0	CR	Indirect (after CR)	0	Self-paced	n/a	Neg & Pos = Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	-
	2 (time, high arousal)	Remember background and objects, and formulate a short story about the scene	Intentional	Temporal (order of appearance)	3	Intrinsic	0	CR	Indirect (after CR)	0	Self-paced	n/a	Neg & Pos = Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	-
	2 (location, low arousal)	Remember background and objects, and formulate a short story about the scene	Intentional	Location	6	Intrinsic	0	CR	Indirect (after CR)	0	Self-paced	n/a	Neg & Pos = Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	For "perfect" location recall
	2 (location, high arousal)	Remember background and objects, and formulate a short story about the scene	Intentional	Location	6	Intrinsic	0	CR	Indirect (after CR)	0	Self-paced	n/a	Neg & Pos = Neu; Neg & Pos (high arousal) > Neg & Pos (low arousal)	For "perfect" location recall
Sharot & Yonelinas, 2008	1 (5 min delay)	Rate picture for visual complexity considering color or visual detail	Incidental	Encoding task (color *vs.* detail)	2	Extrinsic	5	Recog	Indirect (after remember/ know)	120	8	Proportion of old items whose source was correctly identified	Neg = Neu	5 min > 24 h
	1 (24h delay)	Rate picture for visual complexity considering color or visual detail	Incidental	Encoding task (color *vs.* detail)	2	Extrinsic	1440	Recog	Indirect (after remember/ know)	120	8	Proportion of old items whose source was correctly identified	Neg = Neu	5 min > 24h
Wang, 2012	1 (females)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	n/a	Proportion of old items whose source was correctly identified	Pos > Neg	No sex effect
	1 (males)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	n/a	Proportion of old items whose source was correctly identified	Pos > Neg	No sex effect
Wang & Fu, 2011	1 (immediate)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Wang & Fu, 2011	1 (19 min delay)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay
	1 (63 min delay)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay
	1 (4.9 h delay)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay
	1 (8.8 h delay)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay
	1 (24 h delay)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay
	1 (6 day delay)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay
	1 (2 week delay)	Memorize both word and font color	Intentional	Color (red *vs.* blue)	2	Intrinsic	0	Recog	Indirect (after FR; remember/ know)	60	Self-paced	Proportion of old items whose source was correctly identified	Neg = Pos = Neu	No effect of delay
Yick et al., 2015	1	Rate the emotionality	Intentional	Temporal (block)	2	Intrinsic	1440	Recog	Indirect (after old/new; confidence ratings)	360	Self-paced	Proportion of correct source attributions divided by the sum of correct and incorrect source attributions	Neg > Neu	-
Zhang et al., 2018	1 (self-referential)	Self-referential vs other-referential judgments	Incidental	Encoding task (Self-referential *vs.* other-referential)	2	Extrinsic	5	Recog	Direct	0	n/a	Number of correct responses	Self-referential: Pos > Neg; other-referential: Neg > Pos	-
	1 (other-referential)	Self-referential vs other-referential judgments	Incidental	Encoding task (Self-referential *vs.* other-referential)	2	Extrinsic	5	Recog	Direct	0	n/a	Number of correct responses	Self-referential: Pos > Neg; other-referential: Neg > Pos	-
	2 (self-referential)	Self-referential vs other-referential judgments	Incidental	Encoding task (Self-referential *vs.* other-referential)	2	Extrinsic	2	Recog	Indirect (after old/new)	80	Self-paced	Parameter estimate from a multinomial model using raw frequency responses	Self-referential: Pos > Neg; other-referential: Neg > Pos	-

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Zhang et al., 2018	2 (other-referential)	Self-referential vs other-referential judgments	Incidental	Encoding task (Self-referential *vs.* other-referential)	2	Extrinsic	2	Recog	Indirect (after old/new)	80	Self-paced	Parameter estimate from a multinomial model using raw frequency responses	Self-referential: Pos > Neg; other-referential: Neg > Pos	-

Note. A = Arousal; ANEW = Affective Norms for English Words; CAPS = Chinese Affective Picture System; CAWS = Chinese Affective Word System; CG = Control Group; CR = Cued Recall; edu = Education; emo = Emotional; Ext = External; FA = False Alarms; FR = Free Recall; IAPS = International Affective Picture System; Int = Internal; Mod. = Modality; N = No; n/a = Not Applicable/Available; P = Pictures; Pos = Positive; Neg = Negative; Neu = Neutral; Recog = Recognition; RM = Reality-Monitoring; SM = Source Memory; V = Valence; W = Words; Y = Yes; YA = Young Adults; ↑ = Increased; † = Experiments initially eligible for the meta-analysis and for which information was extracted, but that ended up being excluded from the statistical analysis due to unavailable data required for the computation of effect sizes.

Table S7

Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based on Arousal (Organized by Alphabetic Order) – Part I

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_r	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* LA	*N* MA	*N* HA	A LA (*M*)	A MA (*M*)	A HA (*M*)	V LA (*M*)	V MA (*M*)	V HA (*M*)	Control V?	Mixed *vs.* Pure list	*N* study-test cycles	*N* Study items (total)	Time study (s)
Boywitt, 2005	1 (remember)	Within	Ext	101	85	n/a	n/a	P (IAPS)	Visual	120	40	40	40	2.88	4.82	6.76	5.28	5.10	4.55	Y	Mixed	1	90	4
	1 (know)	Within	Ext	101	85	n/a	n/a	P (IAPS)	Visual	120	40	40	40	2.88	4.82	6.76	5.28	5.10	4.55	Y	Mixed	1	90	4
	2 (remember)	Within	Ext	83	64	n/a	n/a	P (IAPS)	Visual	120	40	40	40	2.88	4.82	6.76	5.28	5.10	4.55	Y	Mixed	1	90	4
	2 (know)	Within	Ext	83	64	n/a	n/a	P (IAPS)	Visual	120	40	40	40	2.88	4.82	6.76	5.28	5.10	4.55	Y	Mixed	1	90	4
Davidson et al., 2006	1	Within	Ext	32	n/a	21.61 (YA)	13.67	W (ANEW)	Audio	48	24	n/a	24	3.17 (list 1); 3.20 (list 2)	n/a	7.11 (list 1); 7.52 (list 2)	5.33 (list 1); 5.66 (list 2)	n/a	4.18 (list 1); 4.10 (list 2)	N (LA > HA)	Mixed	1	24	n/a
	2A	Within	Ext	32	n/a	18.53 (YA)	12.53	W (ANEW)	Visual	48	24	n/a	24	3.17 (list 1); 3.20 (list 2)	n/a	7.11 (list 1); 7.52 (list 2)	5.33 (list 1); 5.66 (list 2)	n/a	4.18 (list 1); 4.10 (list 2)	N (LA > HA)	Mixed	1	24	3
	2B	Within	Ext	23	n/a	18.26 (YA)	12.13	W (ANEW)	Visual	48	24	n/a	24	n/a	n/a	n/a	n/a	n/a	n/a	N	Mixed	1	24	3
†Dutton & Carroll, 2001	1 (central details)	Within	Ext	37	26	24.70	n/a	Films and text	Visual	3	1	1	1	2.29	3.58	5.50	n/a	n/a	n/a	N	Mixed	1	3	n/a
	1 (peripheral details)	Within	Ext	37	26	24.70	n/a	Films and text	Visual	3	1	1	1	2.29	3.58	5.50	n/a	n/a	n/a	N	Mixed	1	3	n/a
Kensinger & Schacter, 2005a	1 (word-picture)	Within	RM	16	9	18-30	n/a	W + P (Hemera tech)	Visual	450 W + 450 P	225	n/a	225	0.59 (W); 0.67 (P)	n/a	2.96 (W); 3.07 (P)	n/a	n/a	n/a	n/a	Mixed	1	300	2
	1 (word-only)	Within	RM	16	9	18-30	n/a	W + P (Hemera tech)	Visual	450 W + 450 P	225	n/a	225	0.59 (W); 0.67 (P)	n/a	2.96 (W); 3.07 (P)	n/a	n/a	n/a	n/a	Mixed	1	300	2
†Kensinger & Schacter, 2005b	1 (word-picture)	Within	RM	16	8	18-30	n/a	W + P (Hemera tech)	Visual	450 W + 450 P	225	n/a	225	< 1	n/a	> 2.5	n/a	n/a	n/a	n/a	Mixed	1	300	2

Study	Experiment (condition)	Design emo	Type of SM task	*N*	N_f	M_{age}	M_{edu}	Type of material (source)	Mod.	*N* items (total)	*N* LA	*N* MA	*N* HA	A LA (*M*)	A MA (*M*)	A HA (*M*)	V LA (*M*)	V MA (*M*)	V HA (*M*)	Control V?	Mixed *vs.* Pure list	*N* study-test cycles	*N* Study items (total)	Time study (s)
†Kensinger & Schacter, 2005b	1 (word-only)	Within	RM	16	8	18-30	n/a	W + P (Hemera tech)	Visual	450 W + 450 P	225	n/a	225	< 1	n/a	> 2.5	n/a	n/a	n/a	n/a	Mixed	1	300	2
Mather & Knight, 2008	5	Within	Ext	56	38	21.21	n/a	P (faces)	Visual	12	6	n/a	6	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Mixed	1	12	2
Mather & Nesmith, 2008	1 (no-lag)	Within	Ext	40	26	19.65	n/a	P (IAPS + unknown source)	Visual	144	72	48	24	2.43 – 2.66	4.50 – 5.88	7.64	5.32 – 5.71	2.71 – 6.62	1.71	No	Mixed	1	72	1
	1 (lag)	Within	Ext	40	26	19.65	n/a	P (IAPS + unknown source)	Visual	144	72	48	24	2.43 – 2.66	4.50 – 5.88	7.64	5.32 – 5.71	2.71 – 6.62	1.71	No	Mixed	1	72	1
	2 (alternating picture and dot)	Within	Ext	16	8	19.00	n/a	P (IAPS + unknown source)	Visual	144	72	n/a	72	2.43 – 2.66	4.50 – 5.88	7.64	5.32 – 5.71	2.71 – 6.62	1.71	No	Mixed	1	72	2
	2 (blocked dots)	Within	Ext	16	8	19.00	n/a	P (IAPS + unknown source)	Visual	144	72	n/a	72	2.43 – 2.66	4.50 – 5.88	7.64	5.32 – 5.71	2.71 – 6.62	1.71	No	Mixed	1	72	2
	3 (seen with arousing picture)	Within	Ext	48	33	18.81	n/a	P (IAPS + unknown source)	Visual	144	72	n/a	72	2.43 – 2.66	4.50 – 5.88	7.64	5.32 – 5.71	2.71 – 6.62	1.71	No	Mixed	1	72	2
	3 (seen with nonarousing picture)	Within	Ext	48	33	18.81	n/a	P (IAPS + unknown source)	Visual	144	72	n/a	72	2.43 – 2.66	4.50 – 5.88	7.64	5.32 – 5.71	2.71 – 6.62	1.71	No	Mixed	1	72	2
	4	Within	Ext	18 6	12 3	23.60	n/a	P (IAPS + unknown source)	Visual	48	24	n/a	24	n/a	n/a	n/a	n/a	n/a	n/a	No	Mixed	1	24	Self-paced
Nashiro & Mather, 2011a	1	Within	Ext	24	13	18 - 29 (YA)	13.56	P (IAPS + unknown source)	Visual	32	16	n/a	16	n/a	n/a	n/a	n/a	n/a	n/a	No	Mixed	1	32	5
Nashiro & Mather, 2011b	1	Within	Ext	18	15	18 – 25 (YA)	14.61	P (IAPS + unknown source)	Visual	32	16	n/a	16	n/a	n/a	n/a	n/a	n/a	n/a	No	Mixed	1	32	5
	2	Within	Ext	24	17	18 – 21 (YA)	12.88	P (IAPS + unknown source)	Visual	16	8	n/a	8	n/a	n/a	n/a	n/a	n/a	n/a	No	Mixed	1	16	5

Note. A = Arousal; ANEW = Affective Norms for English Words; CAPS = Chinese Affective Picture System; CAWS = Chinese Affective Word System; CG = Control Group; CR = Cued Recall; edu = Education; emo = Emotional; Ext = External; FA = False Alarms; FR = Free Recall; HA = High Arousal; IAPS = International Affective Picture System; Int = Internal; LA = Low Arousal; MA = Medium Arousal; Mod. = Modality; N = No; n/a = Not Applicable/Available; N_f = Number of Females; P = Pictures; pos = Positive; neg = Negative; neu = Neutral; Recog = Recognition; RM = Reality-Monitoring; SM = Source Memory; V = Valence; W = Words; Y = Yes; YA = Young Adults; ↑ = Increased; ∩ = Inverted "U" Relationship; † = Experiments initially eligible for the meta-analysis and for which information was extracted, but that ended up being excluded from the statistical analysis due to unavailable data required for the computation of effect sizes.

Table S8

Summary of the Data Extracted from Each Experiment Eligible for the Meta-analysis with Respect to the Effects Based Arousal (Organized by Alphabetic Order) – Part II

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Boywitt, 2005	1 (remember)	The color the frame appeared in the picture?	Intentional	Background color (green *vs.* blue)	2	Extrinsic	5	Recog	Indirect (after old/new; remember/ know)	30	n/a	Proportion of remember items whose source was correctly identified	Low > Medium > High	-
	1 (know)	The color the frame appeared in the picture?	Intentional	Background color (green *vs.* blue)	2	Extrinsic	5	Recog	Indirect (after old/new; remember/ know)	30	n/a	Proportion of know items whose source was correctly identified	Low = Medium = High	-
	2 (remember)	The picture appeared in the right or left side of the screen?	Intentional	Location	2	Intrinsic	5	Recog		30	n/a	Proportion of remember items whose source was correctly identified	∩ (from low to high)	-
	2 (know)	The picture appeared in the right or left side of the screen?	Intentional	Location	2	Intrinsic	5	Recog	Indirect (after old/new; remember/	30	n/a	Proportion of know items whose source was correctly identified	∩ (from low to high)	-
Davidson et al., 2006	1	Indicate if emotional or neutral	Incidental	Identity (female *vs.* male)	2	Intrinsic	0	Recog	Direct (two-alternative forced choice)	0	n/a	Proportion of correct source attributions	High > Low (normative ratings); High = Low (subjective ratings)	Only YA
	2A	Indicate if emotional or neutral	Incidental	Color (purple *vs.* orange)	2	Intrinsic	0	Recog	Direct (two-alternative forced choice)	0	n/a	Proportion of correct source attributions	High = Low (irrespective of normative or subjective ratings)	Only YA
	2B	Remember word and its color	Intentional	Color (purple *vs.* orange)	2	Intrinsic	5	Recog	Direct	24	n/a	Proportion of correct source attributions	High = Low	Only YA
†Dutton & Carroll, 2001	1 (central details)	Watch the video and read the passage	Incidental	Modality (seen *vs.* read)	2	Intrinsic	10080	Recog	Indirect (after CR for item memory)	0	n/a	Proportion of correct source attributions	Medium > High	-
	1 (peripheral details)	Watch the video and read the passage	Incidental	Modality (seen *vs.* read)	2	Intrinsic	10080	Recog	Indirect (after CR for item memory)	0	n/a	Proportion of correct source attributions	Medium > Low	-
Kensinger & Schacter, 2005a	1 (word-picture)	Decide if the object is bigger or smaller than a shoebox	Incidental	Modality (seen *vs.* imagined)	2	Extrinsic	1140-2800	Recog	Direct	150	n/a	n/a	High > Low	-

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Kensinger & Schacter, 2005a	1 (word-only)	Decide if the object is bigger or smaller than a shoebox	Incidental	Modality (seen *vs.* imagined)	2	Extrinsic	1140-2800	Recog	Direct	150	n/a	n/a	High = Low (misattributions)	-
†Kensinger & Schacter, 2005b	1 (word-picture)	Decide if the object is bigger or smaller than a shoebox	Incidental	Modality (seen *vs.* imagined)	2	Extrinsic	1140-2800	Recog	Direct	150	n/a	n/a	High > Low	-
	1 (word-only)	Decide if the object is bigger or smaller than a shoebox	Incidental	Modality (seen *vs.* imagined)	2	Extrinsic	1140-2800	Recog	Direct	150	n/a	n/a	High > Low	-
Mather & Knight, 2008	5	Remember face and location	Intentional	Location (3x3 grid)	8	Intrinsic	0	Recog	Direct	0	n/a	n/a	n/a	-
Mather & Nesmith, 2008	1 (no-lag)	Watch pictures	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Direct	0	Self-paced	Proportion of correct picture-location conjunctions	High & Medium > Low	No lag/no-lag effect
	1 (lag)	Watch pictures	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Direct	0	Self-paced	Proportion of correct picture-location conjunctions	High & Medium > Low	No lag/no-lag effect
	2 (alternating picture and dot)	Watch pictures	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Direct	0	Self-paced	Proportion of correct picture-location conjunctions	High & Medium > Low	No alternating/blocked effect
	2 (blocked dots)	Watch pictures	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Direct	0	Self-paced	Proportion of correct picture-location conjunctions	High & Medium > Low	No alternating/blocked effect
	3 (seen with arousing picture)	Watch pictures	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Direct	0	Self-paced	Proportion of correct picture-location conjunctions	High & Medium > Low	No seen with arousing/ nonarousing picture effect
	3 (seen with nonarousing picture)	Watch pictures	Incidental	Location (3x3 grid)	8	Intrinsic	0	Recog	Direct	0	Self-paced	Proportion of correct picture-location conjunctions	High & Medium > Low	No seen with arousing/ nonarousing picture effect
	4	Respond if picture appears on the top/bottom half of the window	Incidental	Location	4	Intrinsic	0	Recog	Indirect (after old/new)	12	Self-paced	Proportion of correct picture-location conjunctions	High & Medium > Low	-
Nashiro & Mather, 2011a	1	Different encoding tasks: color of the dot; if picture is higher/lower than shape	Incidental	Location (3x3 grid)	8	4	Intrinsic	0	Indirect (after FR)	0	n/a	Proportion of correct source attributions	High > Low	Only YA

Study	Experiment (condition)	Instruction(s) study	Intentional *vs.* Incidental	Type of SM feature	*N* tested features	Intrinsic *vs.* extrinsic (Mather, 2007)	Retention interval (min)	Type of SM test	Direct *vs.* indirect test	*N* new items (total)	Time test items (s)	SM score calculation	Main finding(s)	Note(s)
Nashiro & Mather, 2011b	1	Different encoding tasks: color of the dot; if picture is higher/lower than shape	Incidental	Location (3x3 grid)	8	4	Intrinsic	0	Indirect (after FR)	0	n/a	Proportion of correct source attributions	High > Low	Only YA
	2	Different encoding tasks: color of the dot; if picture is higher/lower than shape	Incidental	Location (3x3 grid)	8	4	Intrinsic	0	Indirect (after old/new)	16	n/a	Proportion of correct source attributions	High > Low	Only YA

Note. A = Arousal; ANEW = Affective Norms for English Words; CAPS = Chinese Affective Picture System; CAWS = Chinese Affective Word System; CG = Control Group; CR = Cued Recall; edu = Education; emo = Emotional; Ext = External; FA = False Alarms; FR = Free Recall; HA = High Arousal; IAPS = International Affective Picture System; Int = Internal; LA = Low Arousal; MA = Medium Arousal; Mod. = Modality; N = No; n/a = Not Applicable/Available; P = Pictures; pos = Positive; neg = Negative; neu = Neutral; Recog = Recognition; RM = Reality-Monitoring; SM = Source Memory; V = Valence; W = Words; Y = Yes; YA = Young Adults; ↑ = Increased; ∩ = Inverted "U" Relationship; † = Experiments initially eligible for the meta-analysis and for which information was extracted, but that ended up being excluded from the statistical analysis due to unavailable data required for the computation of effect sizes.

Table S9

Methodological Quality Assessment of the Studies Included in the Meta-analysis Organized by Alphabetic Order

Authors	Experiment	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	Score (%)
Aupée, 2007		1	1	1	1	1	2	1	1	1	1	1	0	83.33
Blix & Brennen, 2011		1	1	1	0	2	2	1	1	1	1	1	0	66.67
Boywitt, 2005	E1	1	1	0	0	1	0	1	1	1	1	1	0	66.67
	E2	1	1	0	0	1	0	1	1	1	1	1	0	66.67
†Cook et al., 2007	E1	1	1	0	0	1	0	1	1	1	1	1	1	n/a
	E2	1	1	0	0	1	0	1	1	1	1	1	1	n/a
	E3	1	1	0	0	1	0	1	1	1	1	1	1	n/a
	E4	1	1	0	0	1	0	1	1	1	1	1	0	n/a
†D'Argembeau & Van der Linden, 2004	E1	1	1	0	0	1	2	1	1	1	1	1	0	n/a
	E2	1	1	0	0	1	2	1	1	1	0	1	0	n/a
	E3	1	1	0	0	1	2	1	1	1	1	1	0	n/a
	E4	1	1	0	0	1	2	1	1	1	1	1	0	n/a
D'Argembeau & Van der Linden, 2005		1	1	0	0	1	1	1	1	1	1	1	0	75.00
Davidson et al., 2006	E1	1	1	1	1	1	1	1	0	1	1	1	0	83.33
	E2	1	1	1	1	1	1	1	0	1	1	1	0	83.33
†Diaz-Asper et al., 2008		1	1	1	1	1	2	0	1	1	1	0	0	n/a
†Doerksen & Shimamura, 2001	E1	1	1	1	0	1	2	1	1	1	1	1	0	n/a
	E2	1	1	1	0	1	2	1	1	1	1	1	0	n/a
Dougal et al., 2007		1	1	0	1	1	1	1	1	1	1	1	1	91.67
Durbin et al., 2017	E1	1	1	1	1	1	1	1	1	1	1	1	0	91.67
	E2	1	1	1	1	1	1	1	1	1	1	1	0	91.67
†Dutton & Carroll, 2001		1	1	0	0	2	0	1	1	1	1	1	0	n/a
Ferré et al., 2019	E1	1	1	0	0	1	2	1	1	1	1	1	0	66.67
	E2	1	1	0	0	1	2	1	1	1	1	1	0	66.67
	E3	1	1	0	0	1	2	1	1	1	1	1	0	66.67
Gallant et al., 2019	E1	1	1	1	1	1	1	1	1	1	1	1	0	91.67
	E2	1	1	1	1	1	1	1	1	1	1	1	0	91.67

Authors	Experiment	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	Score (%)
Gallant & Yang, 2014		1	1	1	1	1	1	1	1	1	1	1	0	91.67
Gallo et al., 2011		1	1	0	0	1	2	1	0	1	1	1	0	58.33
Jia et al., 2018		1	1	0	0	1	1	1	1	1	1	1	0	75.00
Kaida et al., 2015		1	1	1	1	1	1	1	1	1	1	0	0	83.33
†Kensinger et al., 2011		1	1	1	1	1	1	1	1	1	1	0	1	n/a
Kensinger & Corkin, 2003	E2	1	1	1	1	1	2	1	1	1	1	1	0	83.33
	E4	1	1	1	1	1	2	1	1	1	1	1	1	91.67
Kensinger & Schacter, 2005a		1	1	1	1	1	1	1	1	1	1	0	0	83.33
†Kensinger & Schacter, 2005b		1	1	1	1	1	1	1	1	1	1	0	0	n/a
Kensinger et al., 2007	E1	1	1	1	1	1	1	1	1	1	1	1	0	91.67
	E2	1	1	1	1	1	1	1	1	1	1	1	0	91.67
Kensinger & Schacter, 2006a		1	1	0	1	1	2	1	1	1	1	1	0	75.00
Kensinger & Schacter, 2006b	E1	1	1	1	1	1	2	1	0	1	1	1	0	75.00
Kensinger & Schacter, 2006b	E2	1	1	1	1	1	2	1	0	1	1	1	0	75.00
Kensinger & Schacter, 2006b	E3A/3B	1	1	1	1	1	2	1	0	1	1	1	0	75.00
Koenig & Mecklinger, 2008		1	1	0	0	1	1	1	1	1	1	1	0	75.00
Le Bigot et al., 2018	E1	1	1	0	0	1	1	1	1	1	1	1	0	75.00
	E2	1	1	0	0	1	1	1	1	1	1	1	0	75.00
Leshikar et al., 2015		1	1	1	1	1	1	1	1	0	1	1	0	83.33
MacKenzie et al., 2015	E1	1	1	0	0	1	2	1	1	1	1	1	0	66.67
	E2	1	1	0	0	1	2	1	1	1	1	1	0	66.67
Maddock & Frein, 2009	E1A/1B	1	1	0	0	1	1	1	0	1	1	1	0	66.67
	E3	1	1	0	0	1	1	1	0	1	1	1	0	66.67
	E4	1	1	0	0	1	1	1	0	1	1	1	0	66.67
Mao et al., 2017	E1	1	1	0	0	1	2	1	1	1	1	1	1	75.00
	E2	1	1	1	0	1	2	1	1	1	1	1	0	75.00
Mao et al., 2015		1	1	1	1	1	1	1	1	1	1	1	0	91.67
Mather & Knight, 2008	E5	1	1	0	0	1	2	0	1	1	1	0	0	50.00
Mather & Nesmith, 2008	E1	1	1	0	0	1	2	1	1	1	1	1	0	66.67
	E2	1	1	0	0	1	2	1	1	1	1	1	0	66.67

Authors	Experiment	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	Score (%)
Mather & Nesmith, 2008	E3	1	1	0	0	1	2	1	1	1	1	1	0	66.67
	E4	1	1	0	0	1	1	1	1	1	1	1	0	75.00
Mathews et al., 2013	E1	1	1	0	0	1	2	0	1	0	1	1	0	50.00
	E2	1	1	0	0	1	2	0	1	0	1	1	0	50.00
	E3	1	1	0	0	1	2	0	1	0	1	1	0	50.00
McKague et al., 2012		1	1	0	0	1	2	1	1	0	1	1	1	66.67
Minor & Herzmann, 2019		1	1	0	0	1	1	1	1	1	1	1	0	75.00
Morrison & Haddock, 1997		1	1	0	1	1	1	1	1	1	1	0	0	75.00
Nashiro & Mather, 2011a		1	1	1	1	1	1	1	1	1	1	1	0	91.67
Nashiro & Mather, 2011b	E1	1	1	1	1	1	1	1	1	1	1	1	0	91.67
	E2	1	1	1	1	1	1	1	1	1	1	1	0	91.67
Newsome et al., 2012		1	1	1	1	1	1	1	1	1	1	1	0	91.67
Novak & Mather, 2009	E1	1	1	0	0	1	2	1	1	1	1	1	0	66.67
Otani, Libkuman et al., 2012		1	1	0	0	1	2	1	1	1	1	1	0	66.67
Otani, Jaffa et al., 2012		1	1	0	0	1	2	1	1	1	1	1	0	66.67
Pereira et al., 2019	E1	1	1	1	1	1	1	1	1	1	1	1	0	91.67
	E2	1	1	1	1	1	1	1	1	1	1	1	0	91.67
Rimmele et al., 2011	E1	1	1	0	0	1	2	1	1	1	1	1	0	66.67
Rimmele et al., 2012	E1	1	1	0	0	1	2	1	1	1	1	1	0	66.67
	E2	1	1	0	0	1	2	1	1	1	1	1	0	66.67
	E3	1	1	0	0	1	2	1	1	1	1	1	0	66.67
Ritchey et al., 2019		1	1	1	1	1	1	1	1	1	1	1	0	91.67
Robin et al., 2018		1	1	1	1	1	1	1	1	1	1	1	0	91.67
Schmidt et al., 2011	E1	1	1	1	1	1	2	1	1	1	1	1	0	83.33
	E2	1	1	1	1	1	2	1	1	1	1	1	0	83.33
Sharot & Yonelinas, 2008		1	1	0	0	1	2	1	1	1	1	1	0	66.67
Wang, 2012		1	1	1	0	1	2	1	1	1	1	1	0	75.00
Wang & Fu, 2011		1	1	1	0	1	1	1	1	1	1	1	0	83.33
Yick et al., 2015		1	1	1	1	1	1	1	0	1	1	1	1	91.67
Zhang et al., 2018	E1	1	1	1	1	1	2	1	1	1	1	1	0	83.33

Authors	Experiment	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	Score (%)
Zhang et al., 2018	E2	1	1	1	1	1	2	1	1	1	1	1	0	83.33

Note. 0 = Not Verified; 1 = Verified; 2 = Unclear (there was not enough information to rate the item); E = Experiment; n/a = Not Applicable; Item 1 (I1) – The goals of the study were stated clearly; Item 2 (I2) – The outcome measures of the study were stated clearly; Item 3 (I3) – Relevant sociodemographic (e.g., mean age; sex; formal education) and health-related characteristics (e.g., no psychiatric or neurological conditions) were described. Of note, "1" was only attributed to studies that mentioned both sociodemographic and health-related information, otherwise "0" was attributed; Item 4 (I4) - There was a description of the inclusion and exclusion criteria; Item 5 (I5) - Considering between-subjects designs, the participants were randomly allocated to one of the groups. Considering within-subjects designs, the order of the stimuli was (pseudo-) randomized and/or counterbalanced across participants; Item 6 (I6) - Information regarding the participants who did not complete the study was provided; Item 7 (I7) - There was a detailed description of the experimental manipulation regarding valence/arousal and the source memory task; Item 8 (I8) - The participants were not overtly instructed to the emotional manipulation of the stimuli; Item 9 (I9) - There was a comparison/control condition; Item 10 (I10) - The data for the defined outcome measures was reported; Item 11 (I11) - The outcome data was non-selectively reported; Item 12 (I12) - Other sources of bias were not detected. In the case of the current meta-analysis, the possible sources of bias detected were: incomplete information regarding the selection of materials; incomplete information regarding the affective ratings of the materials used; unbalanced number of stimuli across valence categories (e.g., more negative than positive); merging negative and positive stimuli to create the emotional condition; the arousal properties were not controlled across emotional conditions, even though this is a common approach in different studies. In such cases, it is difficult to dissociate the contributions of valence and/arousal to memory performance; no control regarding other relevant properties of the stimuli such as frequency, imagery and concreteness in the case of words, or visual complexity, brightness, contrast and number of represented categories (e.g., people; animals; food; objects, landscapes) in the case of pictures; the lists of stimuli were not counterbalanced across participants (same old and new items for all participants); even though participants were not overtly instructed about the emotion manipulation, they were hinted about the appearance of emotional stimuli; given that age influences memory performance, the experiment included a mixed sample of distinct age groups, such as middle-aged adults and elderly, without addressing the possible impact of the variable age on the results; unbalanced number of old/new items during the test phase. † = Experiments initially eligible for the meta-analysis and for which information was extracted, but that ended up being excluded from the statistical analysis due to unavailable data required for the computation of effect sizes.

Figure S1. Funnel plot graphs after implementing the trim-and-fill method. Of note, whereas in the negative *vs.* positive and high *vs.* low arousal analyses no missing study was imputed, in the negative *vs.* neutral and positive *vs.* neutral analyses a total of 12 and 11 missing studies were imputed on the left side, respectively (see both funnel plots on the top).

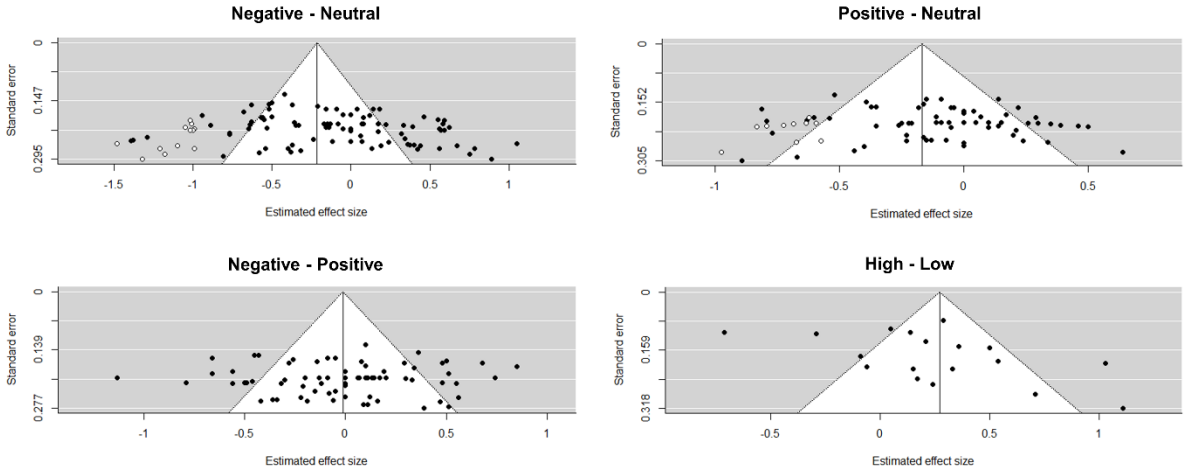
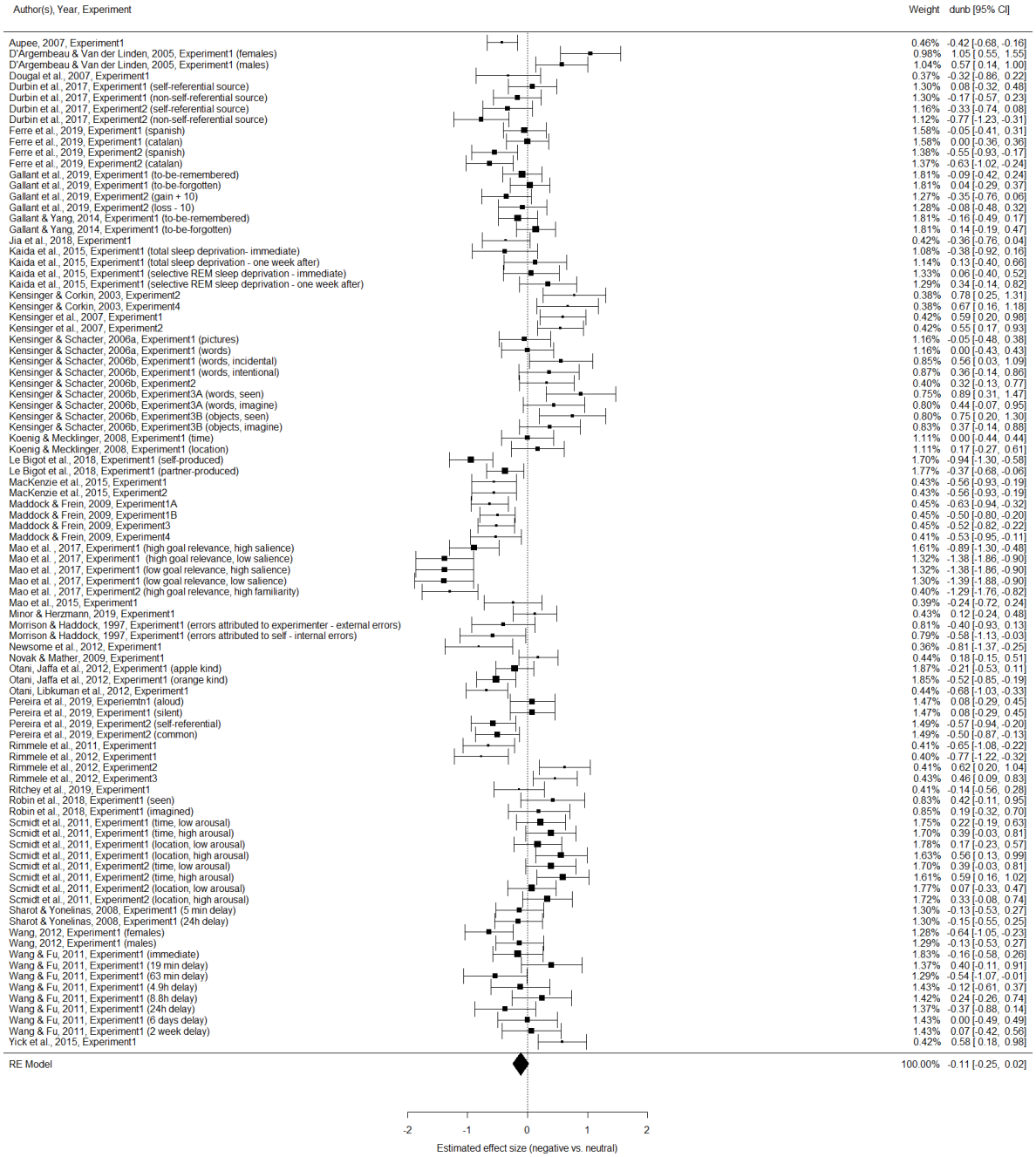
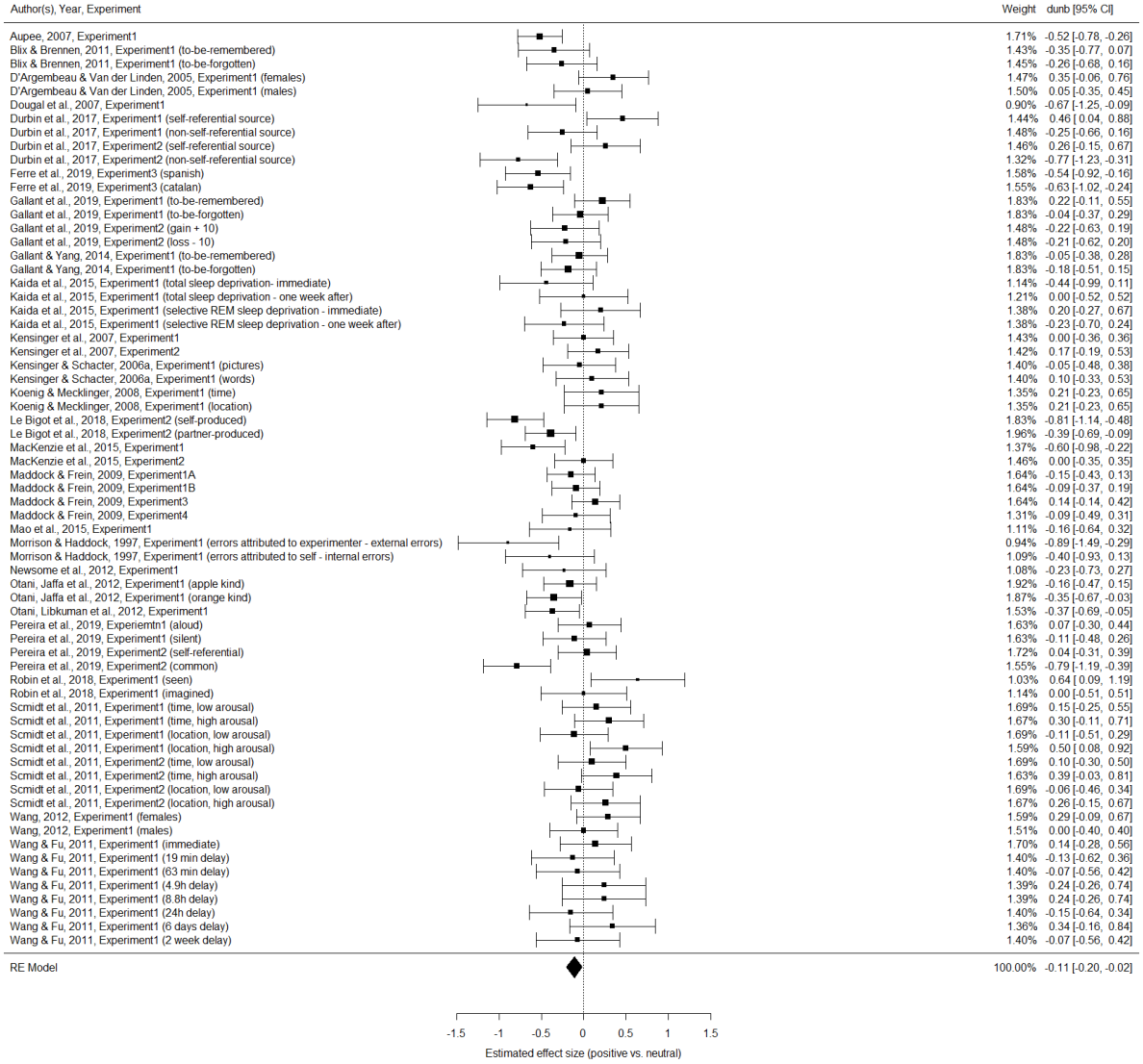


Figure S2. Forest plots for the six meta-analyses: negative *vs.* neutral (panel A); positive *vs.* neutral (panel B); negative *vs.* positive (panel C); high *vs.* low (panel D); medium *vs.* low (panel E); high *vs.* medium (panel F).

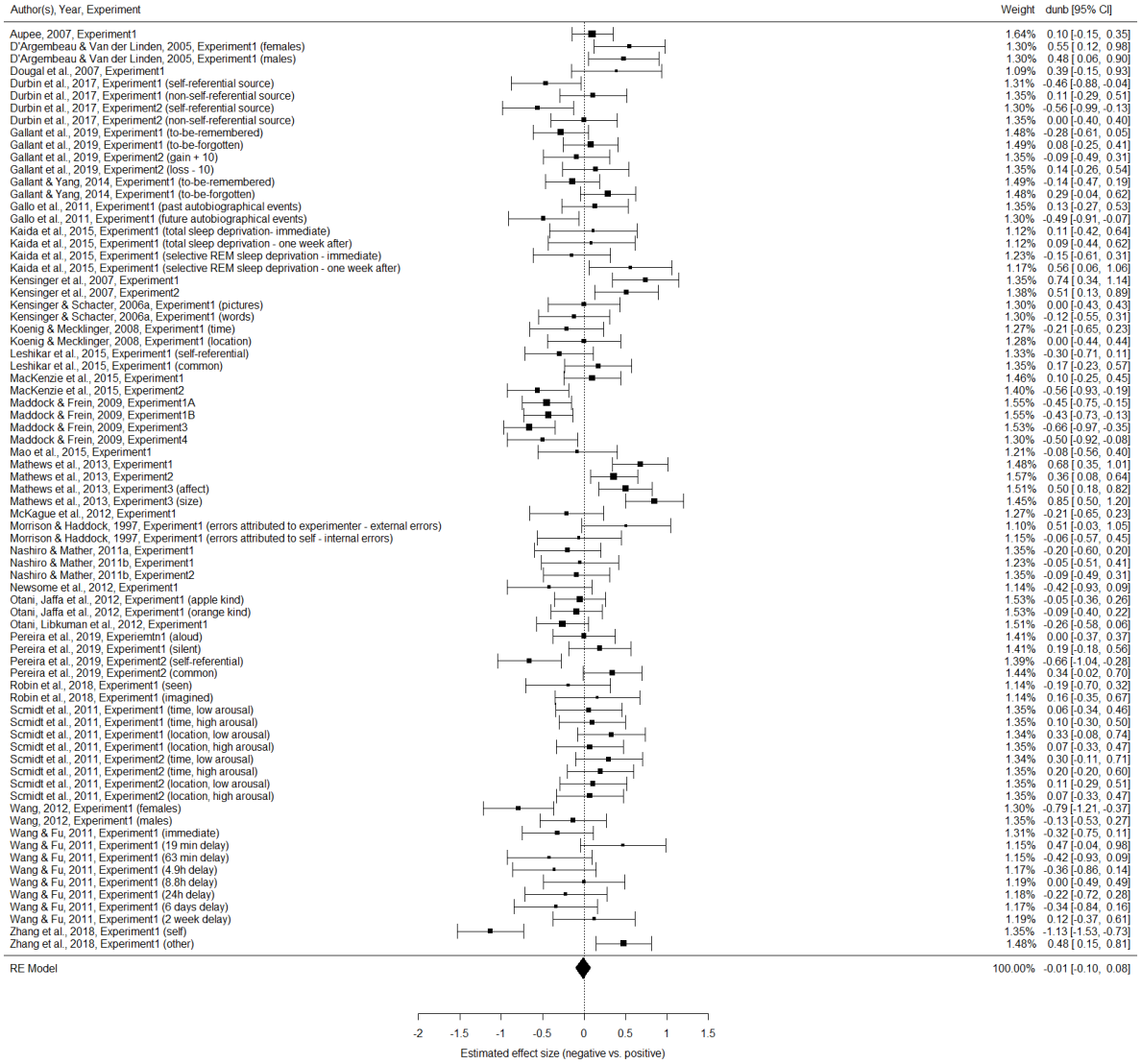
(A)



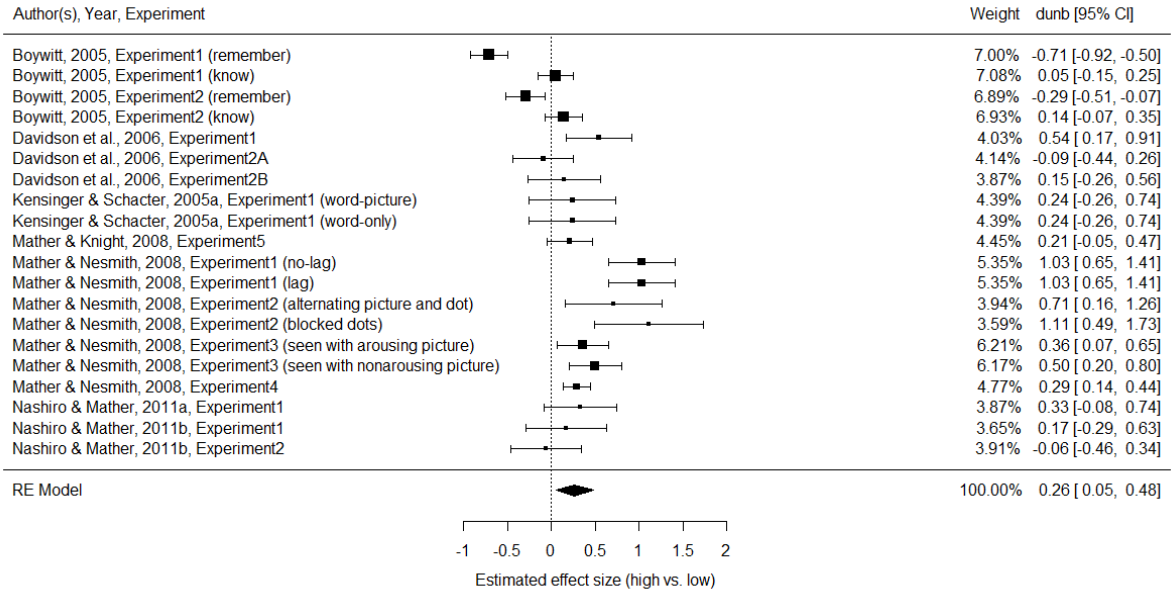
(B)


Author(s), Year, Experiment
Weight
dunb [95% CI]
Aupee, 2007, Experiment1 1.71% -0.52 [-0.78, -0.26]
Blix & Brennen, 2011, Experiment1 (to-be-remembered) 1.43% -0.35 [-0.77, 0.07]
Blix & Brennen, 2011, Experiment1 (to-be-forgotten) 1.45% -0.26 [-0.68, 0.16]
D'Argembeau & Van der Linden, 2005, Experiment1 (females) 1.47% 0.35 [-0.06, 0.76]
D'Argembeau & Van der Linden, 2005, Experiment1 (males) 1.50% 0.05 [-0.35, 0.45]
Dougal et al., 2007, Experiment1 0.90% -0.67 [-1.25, -0.09]
Durbin et al., 2017, Experiment1 (self-referential source) 1.44% 0.46 [0.04, 0.88]
Durbin et al., 2017, Experiment1 (non-self-referential source) 1.48% -0.25 [-0.66, 0.16]
Durbin et al., 2017, Experiment2 (self-referential source) 1.46% 0.26 [-0.15, 0.67]
Durbin et al., 2017, Experiment2 (non-self-referential source) 1.32% -0.77 [-1.23, -0.31]
Ferre et al., 2019, Experiment3 (spanish) 1.58% -0.54 [-0.92, -0.16]
Ferre et al., 2019, Experiment3 (catalan) 1.55% -0.63 [-1.02, -0.24]
Gallant et al., 2019, Experiment1 (to-be-remembered) 1.83% 0.22 [-0.11, 0.55]
Gallant et al., 2019, Experiment1 (to-be-forgotten) 1.83% -0.04 [-0.37, 0.29]
Gallant et al., 2019, Experiment2 (gain + 10) 1.48% -0.22 [-0.63, 0.19]
Gallant et al., 2019, Experiment2 (loss - 10) 1.48% -0.21 [-0.62, 0.20]
Gallant & Yang, 2014, Experiment1 (to-be-remembered) 1.83% -0.05 [-0.38, 0.28]
Gallant & Yang, 2014, Experiment1 (to-be-forgotten) 1.83% -0.18 [-0.51, 0.15]
Kaida et al., 2015, Experiment1 (total sleep deprivation- immediate) 1.14% -0.44 [-0.99, 0.11]
Kaida et al., 2015, Experiment1 (total sleep deprivation - one week after) 1.21% 0.00 [-0.52, 0.52]
Kaida et al., 2015, Experiment1 (selective REM sleep deprivation - immediate) 1.38% 0.20 [-0.27, 0.67]
Kaida et al., 2015, Experiment1 (selective REM sleep deprivation - one week after) 1.38% -0.23 [-0.70, 0.24]
Kensinger et al., 2007, Experiment1 1.43% 0.00 [-0.36, 0.36]
Kensinger et al., 2007, Experiment2 1.42% 0.17 [-0.19, 0.53]
Kensinger & Schacter, 2006a, Experiment1 (pictures) 1.40% -0.05 [-0.48, 0.38]
Kensinger & Schacter, 2006a, Experiment1 (words) 1.40% 0.10 [-0.33, 0.53]
Koenig & Mecklinger, 2008, Experiment1 (time) 1.35% 0.21 [-0.23, 0.65]
Koenig & Mecklinger, 2008, Experiment1 (location) 1.35% 0.21 [-0.23, 0.65]
Le Bigot et al., 2018, Experiment2 (self-produced) 1.83% -0.81 [-1.14, -0.48]
Le Bigot et al., 2018, Experiment2 (partner-produced) 1.96% -0.39 [-0.69, -0.09]
MacKenzie et al., 2015, Experiment1 1.37% -0.60 [-0.98, -0.22]
MacKenzie et al., 2015, Experiment2 1.46% 0.00 [-0.35, 0.35]
Maddock & Frein, 2009, Experiment1A 1.64% -0.15 [-0.43, 0.13]
Maddock & Frein, 2009, Experiment1B 1.64% -0.09 [-0.37, 0.19]
Maddock & Frein, 2009, Experiment3 1.64% 0.14 [-0.14, 0.42]
Maddock & Frein, 2009, Experiment4 1.31% -0.09 [-0.49, 0.31]
Mao et al., 2015, Experiment1 1.11% -0.16 [-0.64, 0.32]
Morrison & Haddock, 1997, Experiment1 (errors attributed to experimenter - external errors) 0.94% -0.89 [-1.49, -0.29]
Morrison & Haddock, 1997, Experiment1 (errors attributed to self - internal errors) 1.09% -0.40 [-0.93, 0.13]
Newsome et al., 2012, Experiment1 1.08% -0.23 [-0.73, 0.27]
Otani, Jaffa et al., 2012, Experiment1 (apple kind) 1.92% -0.16 [-0.47, 0.15]
Otani, Jaffa et al., 2012, Experiment1 (orange kind) 1.87% -0.35 [-0.67, -0.03]
Otani, Libkuman et al., 2012, Experiment1 1.53% -0.37 [-0.69, -0.05]
Pereira et al., 2019, Experiemtn1 (aloud) 1.63% 0.07 [-0.30, 0.44]
Pereira et al., 2019, Experiment1 (silent) 1.63% -0.11 [-0.48, 0.26]
Pereira et al., 2019, Experiment2 (self-referential) 1.72% 0.04 [-0.31, 0.39]
Pereira et al., 2019, Experiment2 (common) 1.55% -0.79 [-1.19, -0.39]
Robin et al., 2018, Experiment1 (seen) 1.03% 0.64 [0.09, 1.19]
Robin et al., 2018, Experiment1 (imagined) 1.14% 0.00 [-0.51, 0.51]
Scmidt et al., 2011, Experiment1 (time, low arousal) 1.69% 0.15 [-0.25, 0.55]
Scmidt et al., 2011, Experiment1 (time, high arousal) 1.67% 0.30 [-0.11, 0.71]
Scmidt et al., 2011, Experiment1 (location, low arousal) 1.69% -0.11 [-0.51, 0.29]
Scmidt et al., 2011, Experiment1 (location, high arousal) 1.59% 0.50 [0.08, 0.92]
Scmidt et al., 2011, Experiment2 (time, low arousal) 1.69% 0.10 [-0.30, 0.50]
Scmidt et al., 2011, Experiment2 (time, high arousal) 1.63% 0.39 [-0.03, 0.81]
Scmidt et al., 2011, Experiment2 (location, low arousal) 1.69% -0.06 [-0.46, 0.34]
Scmidt et al., 2011, Experiment2 (location, high arousal) 1.67% 0.26 [-0.15, 0.67]
Wang, 2012, Experiment1 (females) 1.59% 0.29 [-0.09, 0.67]
Wang, 2012, Experiment1 (males) 1.51% 0.00 [-0.40, 0.40]
Wang & Fu, 2011, Experiment1 (immediate) 1.70% 0.14 [-0.28, 0.56]
Wang & Fu, 2011, Experiment1 (19 min delay) 1.40% -0.13 [-0.62, 0.36]
Wang & Fu, 2011, Experiment1 (63 min delay) 1.40% -0.07 [-0.56, 0.42]
Wang & Fu, 2011, Experiment1 (4.9h delay) 1.39% 0.24 [-0.26, 0.74]
Wang & Fu, 2011, Experiment1 (8.8h delay) 1.39% 0.24 [-0.26, 0.74]
Wang & Fu, 2011, Experiment1 (24h delay) 1.40% -0.15 [-0.64, 0.34]
Wang & Fu, 2011, Experiment1 (6 days delay) 1.36% 0.34 [-0.16, 0.84]
Wang & Fu, 2011, Experiment1 (2 week delay) 1.40% -0.07 [-0.56, 0.42]
RE Model 100.00% -0.11 [-0.20, -0.02]
-1.5
-1
-0.5
0
0.5
1
1.5
Estimated effect size (positive vs. neutral)

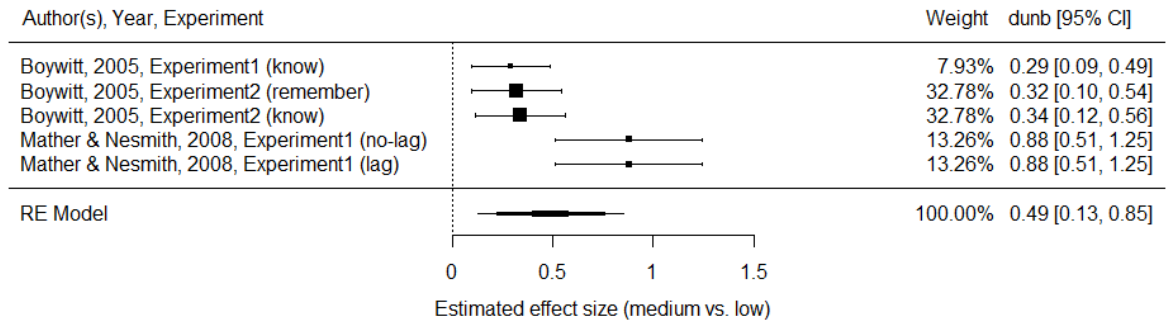
(C)


Author(s), Year, Experiment
Weight
dunb [95% CI]
Aupee, 2007, Experiment1 1.64% 0.10 [-0.15, 0.35]
D'Argembeau & Van der Linden, 2005, Experiment1 (females) 1.30% 0.55 [0.12, 0.98]
D'Argembeau & Van der Linden, 2005, Experiment1 (males) 1.30% 0.48 [0.06, 0.90]
Dougal et al., 2007, Experiment1 1.09% 0.39 [-0.15, 0.93]
Durbin et al., 2017, Experiment1 (self-referential source) 1.31% -0.46 [-0.88, -0.04]
Durbin et al., 2017, Experiment1 (non-self-referential source) 1.35% 0.11 [-0.29, 0.51]
Durbin et al., 2017, Experiment2 (self-referential source) 1.30% -0.56 [-0.99, -0.13]
Durbin et al., 2017, Experiment2 (non-self-referential source) 1.35% 0.00 [-0.40, 0.40]
Gallant et al., 2019, Experiment1 (to-be-remembered) 1.48% -0.28 [-0.61, 0.05]
Gallant et al., 2019, Experiment1 (to-be-forgotten) 1.49% 0.08 [-0.25, 0.41]
Gallant et al., 2019, Experiment2 (gain + 10) 1.35% -0.09 [-0.49, 0.31]
Gallant et al., 2019, Experiment2 (loss - 10) 1.35% 0.14 [-0.26, 0.54]
Gallant & Yang, 2014, Experiment1 (to-be-remembered) 1.49% -0.14 [-0.47, 0.19]
Gallant & Yang, 2014, Experiment1 (to-be-forgotten) 1.48% 0.29 [-0.04, 0.62]
Gallo et al., 2011, Experiment1 (past autobiographical events) 1.35% 0.13 [-0.27, 0.53]
Gallo et al., 2011, Experiment1 (future autobiographical events) 1.30% -0.49 [-0.91, -0.07]
Kaida et al., 2015, Experiment1 (total sleep deprivation- immediate) 1.12% 0.11 [-0.42, 0.64]
Kaida et al., 2015, Experiment1 (total sleep deprivation - one week after) 1.12% 0.09 [-0.44, 0.62]
Kaida et al., 2015, Experiment1 (selective REM sleep deprivation - immediate) 1.23% -0.15 [-0.61, 0.31]
Kaida et al., 2015, Experiment1 (selective REM sleep deprivation - one week after) 1.17% 0.56 [0.06, 1.06]
Kensinger et al., 2007, Experiment1 1.35% 0.74 [0.34, 1.14]
Kensinger et al., 2007, Experiment2 1.38% 0.51 [0.13, 0.89]
Kensinger & Schacter, 2006a, Experiment1 (pictures) 1.30% 0.00 [-0.43, 0.43]
Kensinger & Schacter, 2006a, Experiment1 (words) 1.30% -0.12 [-0.55, 0.31]
Koenig & Mecklinger, 2008, Experiment1 (time) 1.27% -0.21 [-0.65, 0.23]
Koenig & Mecklinger, 2008, Experiment1 (location) 1.28% 0.00 [-0.44, 0.44]
Leshikar et al., 2015, Experiment1 (self-referential) 1.33% -0.30 [-0.71, 0.11]
Leshikar et al., 2015, Experiment1 (common) 1.35% 0.17 [-0.23, 0.57]
MacKenzie et al., 2015, Experiment1 1.46% 0.10 [-0.25, 0.45]
MacKenzie et al., 2015, Experiment2 1.40% -0.56 [-0.93, -0.19]
Maddock & Frein, 2009, Experiment1A 1.55% -0.45 [-0.75, -0.15]
Maddock & Frein, 2009, Experiment1B 1.55% -0.43 [-0.73, -0.13]
Maddock & Frein, 2009, Experiment3 1.53% -0.66 [-0.97, -0.35]
Maddock & Frein, 2009, Experiment4 1.30% -0.50 [-0.92, -0.08]
Mao et al., 2015, Experiment1 1.21% -0.08 [-0.56, 0.40]
Mathews et al., 2013, Experiment1 1.48% 0.68 [0.35, 1.01]
Mathews et al., 2013, Experiment2 1.57% 0.36 [0.08, 0.64]
Mathews et al., 2013, Experiment3 (affect) 1.51% 0.50 [0.18, 0.82]
Mathews et al., 2013, Experiment3 (size) 1.45% 0.85 [0.50, 1.20]
McKague et al., 2012, Experiment1 1.27% -0.21 [-0.65, 0.23]
Morrison & Haddock, 1997, Experiment1 (errors attributed to experimenter - external errors) 1.10% 0.51 [-0.03, 1.05]
Morrison & Haddock, 1997, Experiment1 (errors attributed to self - internal errors) 1.15% -0.06 [-0.57, 0.45]
Nashiro & Mather, 2011a, Experiment1 1.35% -0.20 [-0.60, 0.20]
Nashiro & Mather, 2011b, Experiment1 1.23% -0.05 [-0.51, 0.41]
Nashiro & Mather, 2011b, Experiment2 1.35% -0.09 [-0.49, 0.31]
Newsome et al., 2012, Experiment1 1.14% -0.42 [-0.93, 0.09]
Otani, Jaffa et al., 2012, Experiment1 (apple kind) 1.53% -0.05 [-0.36, 0.26]
Otani, Jaffa et al., 2012, Experiment1 (orange kind) 1.53% -0.09 [-0.40, 0.22]
Otani, Libkuman et al., 2012, Experiment1 1.51% -0.26 [-0.58, 0.06]
Pereira et al., 2019, Experiemtn1 (aloud) 1.41% 0.00 [-0.37, 0.37]
Pereira et al., 2019, Experiment1 (silent) 1.41% 0.19 [-0.18, 0.56]
Pereira et al., 2019, Experiment2 (self-referential) 1.39% -0.66 [-1.04, -0.28]
Pereira et al., 2019, Experiment2 (common) 1.44% 0.34 [-0.02, 0.70]
Robin et al., 2018, Experiment1 (seen) 1.14% -0.19 [-0.70, 0.32]
Robin et al., 2018, Experiment1 (imagined) 1.14% 0.16 [-0.35, 0.67]
Scmidt et al., 2011, Experiment1 (time, low arousal) 1.35% 0.06 [-0.34, 0.46]
Scmidt et al., 2011, Experiment1 (time, high arousal) 1.35% 0.10 [-0.30, 0.50]
Scmidt et al., 2011, Experiment1 (location, low arousal) 1.34% 0.33 [-0.08, 0.74]
Scmidt et al., 2011, Experiment1 (location, high arousal) 1.35% 0.07 [-0.33, 0.47]
Scmidt et al., 2011, Experiment2 (time, low arousal) 1.34% 0.30 [-0.11, 0.71]
Scmidt et al., 2011, Experiment2 (time, high arousal) 1.35% 0.20 [-0.20, 0.60]
Scmidt et al., 2011, Experiment2 (location, low arousal) 1.35% 0.11 [-0.29, 0.51]
Scmidt et al., 2011, Experiment2 (location, high arousal) 1.35% 0.07 [-0.33, 0.47]
Wang, 2012, Experiment1 (females) 1.30% -0.79 [-1.21, -0.37]
Wang, 2012, Experiment1 (males) 1.35% -0.13 [-0.53, 0.27]
Wang & Fu, 2011, Experiment1 (immediate) 1.31% -0.32 [-0.75, 0.11]
Wang & Fu, 2011, Experiment1 (19 min delay) 1.15% 0.47 [-0.04, 0.98]
Wang & Fu, 2011, Experiment1 (63 min delay) 1.15% -0.42 [-0.93, 0.09]
Wang & Fu, 2011, Experiment1 (4.9h delay) 1.17% -0.36 [-0.86, 0.14]
Wang & Fu, 2011, Experiment1 (8.8h delay) 1.19% 0.00 [-0.49, 0.49]
Wang & Fu, 2011, Experiment1 (24h delay) 1.18% -0.22 [-0.72, 0.28]
Wang & Fu, 2011, Experiment1 (6 days delay) 1.17% -0.34 [-0.84, 0.16]
Wang & Fu, 2011, Experiment1 (2 week delay) 1.19% 0.12 [-0.37, 0.61]
Zhang et al., 2018, Experiment1 (self) 1.35% -1.13 [-1.53, -0.73]
Zhang et al., 2018, Experiment1 (other) 1.48% 0.48 [0.15, 0.81]
RE Model 100.00% -0.01 [-0.10, 0.08]
-2
-1.5
-1
-0.5
0
0.5
1
1.5
Estimated effect size (negative vs. positive)

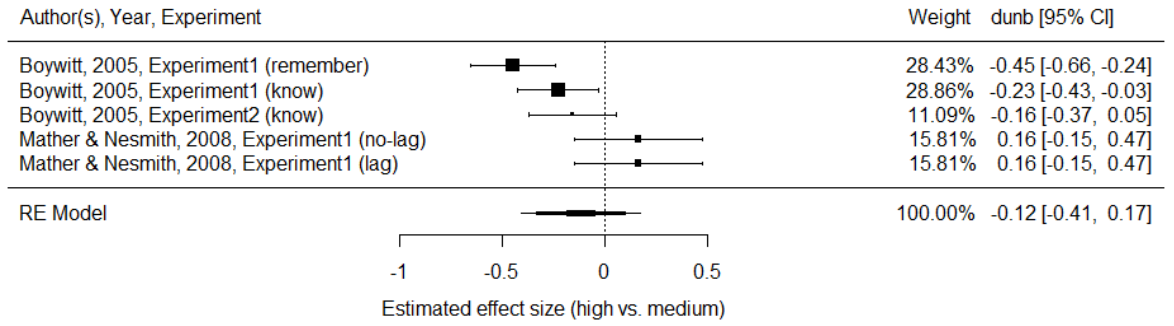
(D)


Author(s), Year, Experiment
Weight
dunb [95% CI]
Boywitt, 2005, Experiment1 (remember) 7.00% -0.71 [-0.92, -0.50]
Boywitt, 2005, Experiment1 (know) 7.08% 0.05 [-0.15, 0.25]
Boywitt, 2005, Experiment2 (remember) 6.89% -0.29 [-0.51, -0.07]
Boywitt, 2005, Experiment2 (know) 6.93% 0.14 [-0.07, 0.35]
Davidson et al., 2006, Experiment1 4.03% 0.54 [0.17, 0.91]
Davidson et al., 2006, Experiment2A 4.14% -0.09 [-0.44, 0.26]
Davidson et al., 2006, Experiment2B 3.87% 0.15 [-0.26, 0.56]
Kensinger & Schacter, 2005a, Experiment1 (word-picture) 4.39% 0.24 [-0.26, 0.74]
Kensinger & Schacter, 2005a, Experiment1 (word-only) 4.39% 0.24 [-0.26, 0.74]
Mather & Knight, 2008, Experiment5 4.45% 0.21 [-0.05, 0.47]
Mather & Nesmith, 2008, Experiment1 (no-lag) 5.35% 1.03 [0.65, 1.41]
Mather & Nesmith, 2008, Experiment1 (lag) 5.35% 1.03 [0.65, 1.41]
Mather & Nesmith, 2008, Experiment2 (alternating picture and dot) 3.94% 0.71 [0.16, 1.26]
Mather & Nesmith, 2008, Experiment2 (blocked dots) 3.59% 1.11 [0.49, 1.73]
Mather & Nesmith, 2008, Experiment3 (seen with arousing picture) 6.21% 0.36 [0.07, 0.65]
Mather & Nesmith, 2008, Experiment3 (seen with nonarousing picture) 6.17% 0.50 [0.20, 0.80]
Mather & Nesmith, 2008, Experiment4 4.77% 0.29 [0.14, 0.44]
Nashiro & Mather, 2011a, Experiment1 3.87% 0.33 [-0.08, 0.74]
Nashiro & Mather, 2011b, Experiment1 3.65% 0.17 [-0.29, 0.63]
Nashiro & Mather, 2011b, Experiment2 3.91% -0.06 [-0.46, 0.34]
RE Model 100.00% 0.26 [0.05, 0.48]
-1
-0.5
0
0.5
1
1.5
2
Estimated effect size (high vs. low)

(E)


Author(s), Year, Experiment
Weight
dunb [95% CI]
Boywitt, 2005, Experiment1 (know) 7.93% 0.29 [0.09, 0.49]
Boywitt, 2005, Experiment2 (remember) 32.78% 0.32 [0.10, 0.54]
Boywitt, 2005, Experiment2 (know) 32.78% 0.34 [0.12, 0.56]
Mather & Nesmith, 2008, Experiment1 (no-lag) 13.26% 0.88 [0.51, 1.25]
Mather & Nesmith, 2008, Experiment1 (lag) 13.26% 0.88 [0.51, 1.25]
RE Model 100.00% 0.49 [0.13, 0.85]
0
0.5
1
1.5
Estimated effect size (medium vs. low)

(F)


Author(s), Year, Experiment
Weight
dunb [95% CI]
Boywitt, 2005, Experiment1 (remember) 28.43% -0.45 [-0.66, -0.24]
Boywitt, 2005, Experiment1 (know) 28.86% -0.23 [-0.43, -0.03]
Boywitt, 2005, Experiment2 (know) 11.09% -0.16 [-0.37, 0.05]
Mather & Nesmith, 2008, Experiment1 (no-lag) 15.81% 0.16 [-0.15, 0.47]
Mather & Nesmith, 2008, Experiment1 (lag) 15.81% 0.16 [-0.15, 0.47]
RE Model 100.00% -0.12 [-0.41, 0.17]
-1
-0.5
0
0.5
Estimated effect size (high vs. medium)

References[1]

Borenstein, M., Hedged, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, USA: Routledge.

Cook, G. I., Hicks, J. L., & Marsh, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition*, *35*(2), 222–230. https://doi.org/10.3758/BF03193443

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, USA: Routledge.

D'Argembeau, A., & Van Der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4*(2), 173–188. https://doi.org/10.1037/1528-3542.4.2.173

Diaz-Asper, C., Malley, J., Genderson, M., Apud, J., & Elvevåg, B. (2008). Context binding in schizophrenia: Effects of clinical symptomatology and item content. *Psychiatry Research*, *159*(3), 259–270. https://doi.org/10.1016/j.psychres.2007.02.018

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1*(1), 5–11. https://doi.org/10.1037/1528-3542.1.1.5

Dutton, A., & Carroll, M. (2001). Eyewitness testimony: Effects of source. *Australian Journal of Psychology*, *53*(2), 83–91. https://doi.org/10.1080/00049530108255128

Kensinger, E. A., Addis, D. R., & Atapattu, R. K. (2011). Amygdala activity at encoding corresponds with memory vividness and with memory for select episodic details. *Neuropsychologia*, *49*(4), 663–673. https://doi.org/10.1016/j.neuropsychologia.2011.01.017

Kensinger, E. A., & Schacter, D. L. (2005). Retrieving accurate and distorted memories: Neuroimaging evidence for effects of emotion. *NeuroImage*, *27*(1), 167–177. https://doi.org/10.1016/j.neuroimage.2005.03.038

[1] To avoid repetition, the complete list of references regarding studies included in the current meta-analysis can be consulted in the main manuscript. They are preceded by an asterisk. Here, references are only included for the formulas used to calculate effect sizes, as well as for the studies that were initially eligible for the meta-analysis, but that ended up being excluded from the statistical analysis due to unavailable data required for the computation of effect sizes.

APPENDIX B

SUPPLEMENTARY MATERIALS - CHAPTER III

Table S1

Affective and Psycholinguistic Properties of the Stimuli Selected for Both Experiments

		Valence category			*F, p*
		Negative *M* (*SD*)	Neutral *M* (*SD*)	Positive *M* (*SD*)	
Experiment 1	Valence	2.73 (0.42)	5.07 (0.27)	7.16 (0.43)	2013.09, < .001
	Arousal	5.43 (0.53)	4.01 (0.65)	5.14 (0.94)	63.62, < .001*
	Frequency	18.67 (23.99)	18.90 (16.80)	23.43 (15.81)	1.17, .31
	N of letters	7.12 (1.83)	6.83 (1.72)	7.30 (2.01)	0.96, .39
	N of syllables	3.07 (0.90)	2.92 (0.79)	3.22 (0.94)	1.75, .18
Experiment 2	Valence	2.88 (0.40)	5.06 (0.66)	7.37 (0.38)	978.29, < .001
	Arousal	4.83 (0.61)	4.45 (0.78)	4.73 (1.06)	2.55, .82
	Frequency	14.54 (25.64)	20.41 (24.11)	16.40 (18.28)	0.825, .44
	N of letters	7.52 (2.04)	7.38 (2.10)	7.88 (2.19)	0.71, .49
	N of syllables	3.40 (0.89)	3.23 (0.93)	3.58 (0.99)	1.72, .18

Note. * Post-hoc comparisons revealed that the arousal ratings for the neutral stimuli were lower than both negative and positive stimuli in Experiment 1.

Table S2

Response Frequencies for Each Valence Condition in Experiment 1

	Negative				Neutral				Positive			
	Aloud	Silent	DNK	New	Aloud	Silent	DNK	New	Aloud	Silent	DNK	New
Aloud	255	60	77	28	223	48	93	46	243	48	90	39
Silent	21	249	99	51	19	239	90	72	23	230	94	73
New	11	17	37	775	1	15	44	778	2	24	63	751

Note. DNK = Do Not Know.

Table S3

Response Frequencies for each Valence Condition in Experiment 2

	Negative				Neutral				Positive			
	SR	Common	DNK	New	SR	Common	DNK	New	SR	Common	DNK	New
SR	244	29	52	59	292	18	32	42	302	24	32	26
Common	48	164	74	97	33	206	66	79	69	152	81	81
New	16	16	59	672	10	11	47	699	16	19	76	657

Note. DNK = Do Not Know; SR = Self-Reference.

Table S4

Multinomial Model Parameters for Item and Source Memory for each Experimental Condition (Source x Valence) in Experiments 1 and 2

	Experiment 1 – *M* (*SD*)						Experiment 2 – *M* (*SD*)					
	Aloud			Silent			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
D	.928	.882	.896	.868	.815	.806	.826	.880	.921	.712	.774	.753
d	.602	.590	.604	.508	.566	.523	.681	.822	.790	.471	.599	.355
a/g	.129	.094	.100	.346	.292	.301	.258	.221	.289	.225	.220	.248

	Negative	Neutral	Positive	Negative	Neutral	Positive
a/g DK	.525	.614	.615	.518	.560	.507
b	.077	.072	.106	.119	.089	.145

Note. DNK = Do Not Know.

Table S5

Descriptive Statistics of Confidence Ratings and Response Time (JOSs Ratings, Source Correct Judgments, Correct Rejections) in each Experimental Condition (Task x Emotion) of Experiments 1 and 2

	Experiment 1 – *M* (*SD*)						Experiment 2 – *M* (*SD*)					
	Aloud			Silent			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
Confidence source correct	5.38 (0.56)	5.41 (0.39)	5.22 (0.53)	5.09 (0.57)	5.03 (0.74)	5.08 (0.68)	5.27 (0.77)	5.42 (0.76)	5.30 (0.85)	4.91 (0.96)	5.16 (0.85)	4.80 (1.11)
RT JOSs ratings	1715 (868)	1650 (827)	1646 (886)	1876 (1151)	1768 (932)	1763 (1185)	1876 (718)	1902 (762)	1592 (574)	1782 (766)	1812 (935)	1668 (679)
RT source correct	3230 (1076)	3007 (835)	3535 (1598)	3558 (990)	3455 (958)	3909 (1291)	3245 (1282)	2827 (829)	2888 (746)	3600 (1150)	3471 (1062)	3488 (1187)

	Negative	Neutral	Positive	Negative	Neutral	Positive
Confidence CR	5.06 (0.75)	5.03 (0.76)	4.75 (0.85)	4.80 (0.86)	4.96 (0.81)	4.63 (0.92)
RT CR	2092 (354)	2039 (472)	2319 (608)	2278 (621)	2131 (533)	2373 (659)

Note. CR = Correct Rejections; JOSs = Judgments of Source; RT = Response Time.

Table S6

Results from the Pair-by-Pair Wilcoxon Tests Applied to the Proportion of Source Incorrect Responses, Do Not Know Responses, and Misses for Experiment 1 and Experiment 2

	Experiment 1					Experiment 2				
	n	*T*	*z*	*r*	*p*	*n*	*T*	*z*	*r*	*p*
			Aloud					Self		
Source incorrect										
Negative - Neutral	28	61	- 1.10	- .15	.272	32	44	- 1.86	- .23	.063
Negative - Positive	28	65	- 1.21	- .16	.225	32	71	- 0.64	- .08	.524
Neutral - Positive	28	78	- 0.35	- .05	.725	32	40	- 1.17	- .15	.243
Do not know										
Negative - Neutral	28	-	-	-	-	32	49	- 2.56	- .32	.011
Negative - Positive	28	-	-	-	-	32	40	- 2.45	- .31	.014
Neutral - Positive	28	-	-	-	-	32	76	- 0.02	- .003	.981
Misses										
Negative - Neutral	28	50	- 2.30	- .31	.022	32	75	- 1.93	- .24	.054
Negative - Positive	28	71	- 1.56	- .21	.119	32	59	- 2.83	- .35	.005*
Neutral - Positive	28	77	- 1.06	- .14	.289	32	51	- 2.08	- .26	.038
			Silent					Common		
Source incorrect										
Negative - Neutral	28	50	- 0.16	- .02	.871	32	40	- 0.28	- .29	.022
Negative - Positive	28	32	- 0.96	- .13	.339	32	65	- 2.02	- .25	.043
Neutral - Positive	28	31	- 1.05	- .14	.292	32	26	- 3.72	- .47	< .001*
Do not know										
Negative - Neutral	28	-	-	-	-	32	74	- 0.50	- .06	.616
Negative - Positive	28	-	-	-	-	32	127	- 0.68	- .09	.499
Neutral - Positive	28	-	-	-	-	32	51	- 1.79	- .22	.074
Misses										
Negative - Neutral	28	64	- 2.07	- .28	.038	32	123	- 1.61	- .20	.107
Negative - Positive	28	57	- 2.06	- .28	.039	32	148	- 0.98	- .12	.329
Neutral - Positive	28	153	- 0.27	- .04	.787	32	176	- 0.44	- .06	.744

	Experiment 1					Experiment 2				
	n	*T*	*z*	*r*	*p*	*n*	*T*	*z*	*r*	*p*
						Negative				
Source incorrect										
Aloud-Silent/ Self - Common	28	25	- 3.18	- .42	.001*	32	135	- 0.11	- .01	.914
Do not know										
Aloud-Silent/ Self - Common	28	-	-	-	-	32	88	- 2.04	- .26	.042
Misses										
Aloud-Silent/ Self - Common	28	24	- 2.88	- .42	.004*	32	47	- 3.44	- .43	.001*
						Neutral				
Source incorrect										
Aloud-Silent/ Self - Common	28	56	- 2.10	- .28	.036	32	64	- 0.22	- .03	.830
Do not know										
Aloud-Silent/ Self - Common	28	-	-	-	-	32	26	- 3.13	- .39	.002*
Misses										
Aloud-Silent/ Self - Common	28	80	- 1.77	- .24	.077	32	36	- 3.28	- .41	.001*
						Positive				
Source incorrect										
Aloud-Silent/ Self - Common	28	65	- 1.50	- .20	.133	32	73	- 2.23	- .28	.026
Do not know										
Aloud-Silent/ Self - Common	28	-	-	-	-	32	4	- 4.00	- .50	<.001*
Misses										
Aloud-Silent/ Self - Common	28	56	- 2.90	- .39	.004*	32	41	- 3.44	- .43	.001*

Note. * $p < .006$

Table S7

Output of the Model-based Bayesian Analysis (BF_{10}) with JASP in Experiment 1

Models	P(M)	P(M | data)	BF_{M}	BF_{10}	BF_{01}†	error %
			Item recognition			
Null model (incl. subject)	0.200	0.001	0.005	1.000	1.000	
Source	0.200	0.080	0.350	66.568	0.015	1.002
Valence	0.200	0.008	0.031	6.416	0.156	1.020
Source + Valence	0.200	0.811	17.196	671.247	0.001	2.209
Source + Valence + Source ✻ Valence	0.200	0.099	0.441	82.160	0.012	3.368
			Item Br (response bias)			
Null model (incl. subject)	0.200	0.016	0.063	1.000	1.000	
Source	0.200	0.145	0.681	9.314	0.107	1.189
Valence	0.200	0.056	0.236	3.576	0.280	0.709
Source + Valence	0.200	0.659	7.725	42.206	0.024	1.496
Source + Valence + Source ✻ Valence	0.200	0.124	0.568	7.965	0.126	2.145
			Source recognition			
Null model (incl. subject)	0.200	0.458	3.379	1.000	1.000	
Source	0.200	0.471	3.554	1.027	0.973	4.276
Valence	0.200	0.033	0.136	0.072	13.893	0.859
Source + Valence	0.200	0.034	0.142	0.075	13.334	4.479
Source + Valence + Source ✻ Valence	0.200	0.004	0.017	0.009	107.686	1.745
			JOSs ratings			
Null model (incl. subject)	0.200	5.501e -24	2.200e -23	1.000	1.000	
Source	0.200	0.001	0.005	2.046e +20	4.793e -21	1.211
Valence	0.200	2.752e -23	1.101e -22	5.003	0.202	0.962
Source + Valence	0.200	0.891	32.843	1.620e +23	5.922e -24	1.253
Source + Valence + Source ✻ Valence	0.200	0.107	0.482	1.953e +22	5.077e -23	2.279
			Gamma correlations			
Null model (incl. subject)	0.200	0.720	10.288	1.000	1.000	
Source	0.200	0.213	1.084	0.296	1.591e -13	1.183
Valence	0.200	0.050	0.209	0.069	0.088	0.863
Source + Valence	0.200	0.015	0.060	0.021	9.545e -16	1.354
Source + Valence + Source ✻ Valence	0.200	0.002	0.009	0.003	7.627e -22	2.856

Note. JOSs = Judgments of Source; All models include subject; † BF_{01} was added to observe how the null model is being favored in relation to the remaining models.

Table S8

Output of the Model-based Bayesian Analysis (BF_{10}) with JASP in Experiment 2

Models	P(M)	P(M | data)	BF_M	BF_{10}	BF_{01}†	error %
			Item recognition			
Null model (incl. subject)	0.200	9.313e -9	3.725e -8	1.000	1.000	
Source	0.200	0.042	0.174	4.470e +6	2.237e -7	2.179
Valence	0.200	5.631e -8	2.252e -7	6.046	0.165	2.280
Source + Valence	0.200	0.811	17.166	8.708e +7	1.148e -8	1.183
Source + Valence + Source ✻ Valence	0.200	0.147	0.691	1.582e +7	6.321e -8	4.872
			Item Br (response bias)			
Null model (incl. subject)	0.200	2.327e -9	9.309e -9	1.000	1.000	
Source	0.200	9.368e -4	0.004	402550.432	2.484e -6	2.121
Valence	0.200	3.105e -7	1.242e -6	133.435	0.007	0.952
Source + Valence	0.200	0.720	10.266	3.092e +8	3.234e -9	2.064
Source + Valence + Source ✻ Valence	0.200	0.279	1.551	1.201e +8	8.328e -9	1.817
			Source recognition			
Null model (incl. subject)	0.200	1.073e -15	4.292e -15	1.000	1.000	
Source	0.200	7.083e -5	2.834e -4	6.601e +10	1.515e -11	1.400
Valence	0.200	1.331e -14	5.323e -14	12.400	0.081	1.462
Source + Valence	0.200	0.011	0.044	1.007e +13	9.931e -14	14.643
Source + Valence + Source ✻ Valence	0.200	0.989	363.778	9.218e +14	1.085e -15	6.687
			JOSs ratings			
Null model (incl. subject)	0.200	2.462e -4	9.851e -4	1.000	1.000	
Source	0.200	0.304	1.747	1234.836	8.098e -4	1.029
Valence	0.200	2.533e -4	0.001	1.029	0.972	0.733
Source + Valence	0.200	0.465	3.477	1888.860	5.294e -4	1.902
Source + Valence + Source ✻ Valence	0.200	0.230	1.198	935.928	0.001	2.137
			Gamma correlations			
Null model (incl. subject)	0.200	0.734	11.031	1.000	1.000	
Source	0.200	0.210	1.066	0.287	3.489	0.718
Valence	0.200	0.042	0.178	0.058	17.270	1.063
Source + Valence	0.200	0.012	0.049	0.016	61.245	1.419
Source + Valence + Source ✻ Valence	0.200	0.001	0.005	0.002	571.741	3.106

Note. JOSs = Judgments of Source; All models include subject; † BF_{01} was added to observe how the null model is being favored in relation to the remaining models.

Method

Data analysis

Multinomial models.

The first step was to bring together all the participants' responses in a 3 x 4 table, where rows correspond to the source testing responses ("read in silence"; "read aloud"; "new") and the columns correspond to participants' responses ("read in silence"; "read aloud"; "read, but do not know if silently/aloud"; "new"). This table was computed for each valence condition (see Table S2 and S3). The model adopted here followed Leshikar and colleagues (2015), which relied on the proposal of Batchelder and Riefer (1990). Accordingly, the following parameters were considered: 'D' as the probability of correctly recognizing studied stimuli irrespective of the encoding source; 'd' as the probability of correctly recognizing the stimulus source given that it was accurately identified as a studied item; 'b' as the probability of correctly guessing whether a previously studied item was studied or the probability of erroneously considering a new stimulus as old; 'g' as the probability of guessing the stimulus source given that the item was already assessed as old; 'a' as the probability of guessing the source given that the stimulus was correctly detected as study items. By imposing the constraint 'a' = 'g' (Batchelder & Riefer, 1990; Dodson, Holland, & Shimamura, 1998; Leshikar et al., 2015), eight parameters were estimated for each valence condition – two 'D', two 'd', three 'a', one 'b' – giving rise to a 24-parameter full model. Of note, two additional restrictions were imposed: all the parameter values could only vary between 0.00000001 and 0.999999999 (Dodson, Prinzmetal, & Shimamura, 1998). As we had three 'a'/'g' parameters to estimate, their sum was constrained to one. Both parameters' estimation and model fit were computed using the excel solver function following Dodson, Prinzmetal et al. (1998), which employs the maximum likelihood ratio and the likelihood statistic (G^2). Additionally, goodness of fit for different models was assessed by comparing G^2 statistics with a chi-square distribution (alpha = .05). After an initial parameter estimation for the full 24 parameters' model, the goodness of fit of the general model was tested by changing the parameters until a satisfactory solution was found. Then, we compared the goodness of fit between different nested models in a two by two fashion, contemplating item/source memory accuracy and item/source memory response bias. As stated by Dodson, Holland et al. (1998), the idea is to compare models in which the parameters can vary without restrictions with models in which specific parameters are constrained to be equal. If the model fit does not differ significantly between models, it might be the case that the parameters are not different; if the model with free parameters reveals a better fit than the restricted one, it may suggest that the parameters are different.

Goodman-Kruskal gamma correlation.

To compute gamma, the following elements were considered: (a) the number of correct "remember" predictions by adding the cases that received a rating between 4 and 6; (b) the number of incorrect "remember" predictions by adding the cases with ratings between 4 and 6 that were later forgotten; (c) the number of incorrect "forget" predictions by adding the cases with ratings between 1 and 3 that were later remembered; (d) the number of correct "forget" predictions by adding the cases with ratings between 1 and 3 that were actually forgotten. These frequencies were then inserted into the following formula: $G = (ad - bc)/(ad - bc)$. However, it was not always possible to calculate the formula as some of the elements were equal to zero. To overcome this issue, an adjustment was employed as recommended by Snodgrass and Corwin (1988), and as adopted by previous studies (e.g., Bastin et al., 2012; Grainger, Williams, & Lind, 2016). More specifically, the value of 0.5 was added to each prediction frequency (a, b, c, d), and the result was then divided by the total number of judgments plus one ($N + 1$).

Gamma correlations provide a measure of association between the predictions about which words will be later remembered and forgotten and the actual performance of the participant. Large and positive gamma values are indicative of a good metamnemonic resolution, whereas values equal or below zero do not support an accurate relation between prediction and performance.

Results

The multinomial model results and specific statistical analyses performed on the response times and confidence ratings are presented for both Experiment 1 and Experiment 2. The descriptive statistics are shown in Table S5. Additionally, for both experiments, the proportion of incorrect source responses, do not know responses, misses, correct rejections, and (corrected) false alarms were subjected to a 3 (valence: negative *vs*. neutral *vs*. positive) x 2 (source: aloud *vs*. silent) Friedman's ANOVA. Wilcoxon tests with Bonferroni corrections were used as follow-up tests in the case a statistically significant result was obtained in the context of the Friedman's ANOVA. Non-parametric tests were used given that most of the experimental conditions revealed a non-normal distribution according to the Shapiro-Wilk test and the absolute values of skewness and kurtosis.

Of note, a one sample *t*-test (test value = 0) was computed for each recognition measure from Experiment 1 to ensure participants could discriminate between old and new items, as well as between the two sources [$t(27) \geq 8.65$, $p < .001$ for all experimental conditions]. The same was done in the

case of Experiment 2, and a similar result was observed [$t(31) \geq 5.62$, $p < .001$ for all experimental conditions].

Experiment 1

Multinomial model results.

After running the solver function on the 24-parameter model, it was possible to verify that the solution was not a good fit for the data, because the obtained G^2 value of 25.35 was above the critical chi-square value of 12.59 (considering six degrees of freedom). To obtain a G^2 value below 12.59, we tried to keep most of the parameters yielded by the initial solution and changed three 'a'/'g' parameters of only one valence condition. This alteration resulted in a G^2 of 12.42 which is below the critical chi-square value. All the parameter values are presented in Table S4. With this model, the ANOVA results were tested.

In the case of item memory, when both positive and negative parameters were set to be equal, the G^2 was 20.06, which suggests that the parameters might be different after all, $G^2(2) = 7.64$, $p < .05$. When both aloud *vs.* silence conditions were equated, the G^2 was 38.63, which supports the ANOVA result that words read aloud are better recognized than words read silently, $G^2(3) = 26.21$, $p < .05$. In the case of source memory and considering no statistically significant differences emerged from the repeated-measures ANOVA, we set all the source 'd' to be equal regardless of the experimental condition, and the G^2 value obtained was 17.80. This result indicates that the parameters are not different, $G^2(5) = 5.38$, $p > .05$, in good agreement with the ANOVA outcome. Overall, the multinomial-based results are consistent with the ANOVA results.

Incorrect source responses.

The 3 x 2 Friedman's ANOVA yielded a statistically significant result, $X^2(5) = 17.25$, $p = .004$. Specifically, the Wilcoxon tests with Bonferroni corrections ($p < .006$) showed that the proportion of incorrect responses was higher for negative words read aloud than for negative words read silently. No other comparisons reached the statistically significant threshold (see Table S6).

Do not know responses.

The Friedman's ANOVA was not statistically significant, $X^2(5) = 5.15$, $p = .398$.

Misses.

The Friedman's ANOVA was statistically significant, $X^2(5) = 24.11$, $p < .001$. Only two comparisons survived the Wilcoxon tests with Bonferroni corrections ($p < .006$): the proportion of misses was higher for both negative and positive words read silently during the study phase when compared to negative and positive words read aloud, respectively (see Table S6).

Correct rejections and corrected false alarm rates.

Regarding the proportion of correct rejections, no statistically significant difference emerged, $X^2(2) = 5.93$, $p = .052$. Considering the proportion of corrected false alarm rates, the The Friedman's ANOVA showed a statistically significant effect, $X^2(2) = 6.17$, $p = .046$. Nonetheless, when applying the Wilcoxon tests with Bonferroni corrections ($p < .017$), none of the comparisons survived this correction. Taken together, no significant differences were observed between the responses to negative/neutral/positive new words during the recognition test.

Response time.

JOSs ratings. A 3 (valence: negative, neutral, positive) x 2 (source: silent, aloud) repeated-measures ANOVA yielded no statistically significant results (valence: $F(2, 54) = 1.37$, $p = .263$, $\eta_p^2 = .05$; source : $F(1, 27) = 2.90$, $p = .100$, $\eta_p^2 = .10$; valence x source: $F(2, 54) = 0.12$, $p = .884$, $\eta_p^2 = .01$). Thus, it seems that participants took the same time to judge if they would later recall information differing in valence and production mode (see Table S5).

Source correct judgments. By applying the same 3 x 2 repeated-measures ANOVA on correct source judgments, a main effect of source, $F(1, 25) = 6.55$, $p = .017$, $\eta_p^2 = .21$, and a main effect of valence, $F(2, 50) = 4.64$, $p = .021$, $\eta_p^2 = .16$, $\varepsilon = .82$, were observed. However, there was no interaction effect, $F(2, 50) = 0.10$, $p = .869$, $\eta_p^2 = .004$, $\varepsilon = .80$. Specifically, the response time was slower in the case of positive words ($M = 3722$, $SE = 256$) when compared with neutral words ($M = 3231$, $SE = 152$, $p = .022$). The response time was also slower for words read silently ($M = 3641$, $SE = 186$) than words read aloud ($M = 3257$, $SE = 189$, $p = .017$).

Correct rejections. In the case of the response time for correct rejections of new words, a repeated measures ANOVA with the factor valence showed a statistically significant effect, $F(2, 54) = 4.80$, $p = .012$, $\eta_p^2 = .16$, revealing once again that the time to respond to new positive words ($M = 2319$, $SE = 115$) was slower in comparison with new neutral words ($M = 2039$, $SE = 89$, $p = .015$).

Confidence ratings.

Source correct judgments. When analyzing the confidence ratings after accurate source memory judgments, the 3 x 2 repeated-measures ANOVA revealed only a statistically significant main effect of source, $F(1, 26) = 6.99$, $p = .014$, $\eta_p^2 = .21$, (valence: $F(2, 54) = 1.37$, $p = .263$, $\eta_p^2 = .05$; valence x source: $F(2, 54) = 0.12$, $p = .884$, $\eta_p^2 = .01$): confidence ratings for words read aloud ($M = 5.34$, $SE = 0.08$) were higher than confidence ratings for words read silently ($M = 5.07$, $SE = 0.12$, $p = 0.14$).

Correct rejections. In the case of confidence ratings for accurately rejected new words, a repeated-measures ANOVA run on the factor valence yielded a significant effect, $F(2, 54) = 12.57$, $p < .001$, $\eta_p^2 = .32$, $\varepsilon = .81$, which showed that participants were more confident when rejecting both new negative ($M = 5.07$, $SE = 0.14$, $p = .001$) and new neutral words ($M = 5.03$, $SE = 0.14$, $p = .004$) than new positive words ($M = 4.75$, $SE = 0.16$, $p = 0.14$).

Experiment 2

Mean proportion of "yes" responses.

The mean proportion of "yes" responses in the self-referential condition was .09 ($SD = .09$), .31 ($SD = .18$), and .69 ($SD = .17$) for negative, neutral, and positive words, respectively. In the case of the common condition, this proportion was .64 ($SD = .16$), .78 (.16), .90 (.12) for negative, neutral, and positive words, respectively. A 3 (valence: negative *vs.* neutral *vs.* positive) x 2 (source: self-reference *vs.* common) repeated-measures ANOVA revealed a main effect of valence, $F(2, 62) = 231.46$, $p < .001$, $\eta_p^2 = .88$, a main effect of source, $F(1, 31) = 263.91$, $p < .001$, $\eta_p^2 = .90$, and an interaction effect, $F(2, 62) = 20.84$, $p < .001$, $\eta_p^2 = .40$. In general, the mean proportion of "yes" responses was higher in the case of common judgments compared to self-referential judgments (all $p < .001$), and it also differed according to word valence: negative < neutral < positive (all $p < .01$).

Multinomial model results.

When modelling the results of this experiment based on the 24-parameter model, we came across the same problem reported in Experiment 1, that is, the solution found with solver was not the best fit for the data as the G^2 value of 44.39 was above the critical chi-square value of 12.59 (considering six degrees of freedom). In these circumstances, we applied the same strategy reported in Experiment 1 to achieve a better solution. We ended up obtaining a G^2 of 11.15 which is below the critical chi-square value. All the parameter values are presented in Table S4. In the case of item

memory, we started by testing if there was a difference between the parameters of neutral and negative stimuli. For this, negative and neutral parameters were set to be equal regardless of the source condition. The obtained G^2 value was 43.46 which supports the difference between neutral and negative words, $G^2(3) = 32.31$, $p < .05$. To confirm the main effect of source (self-reference *vs.* common), the item parameters were set to be equal for each valence condition, which revealed a G^2 value of 71.23. So, the model that posits the source conditions as different is a better fit for the data, $G^2(3) = 60.08$, $p < .05$, which is in accordance with the ANOVA results. Concerning the source memory accuracy results, specifically in the context of the self-referential condition, the 'd' parameter was set to be equal for both neutral and positive stimuli. The G^2 value was 11.85, which suggests that these parameters might not differ, $G^2(1) = 0.70$, $p > .05$. Once again, this result supports the ANOVA results showing that the source of both neutral and positive stimuli is better recognized than the source of negative stimuli. In the case of the common condition, the ANOVA results showed that the source of emotional words was less accurately recognized when compared with neutral stimuli. Thus, we tested a model wherein both negative and positive 'd' parameters of the common condition were equal while keeping the 'd' neutral parameter varying freely. The obtained G^2 value was 14.02, which shows that the former parameters do not differ, $G^2(1) = 2.87$, $p > .05$. Again, the multinomial-based results are consistent with the ANOVA results.

Incorrect source responses.

The 3 x 2 Friedman's ANOVA revealed a statistically significant effect, $X^2(5) = 26.80$, $p < .001$. Nonetheless, only one comparison survived the statistical significance threshold imposed by the Wilcoxon tests with Bonferroni corrections ($p < .006$). Specifically, the proportion of incorrect source responses for positive words encoded in the common condition was higher in comparison with neutral words in the common condition (see Table S6).

Do not know responses.

The 3 x 2 Friedman's ANOVA was statistically significant, $X^2(5) = 36.98$, $p < .001$. The Wilcoxon tests with Bonferroni corrections ($p < .006$) showed that for both neutral and positive words studied in the common condition received more "do not know" responses in contrast to neutral and positive words in the self-referential condition, respectively (see Table S6).

Misses.

A statistically significant result was obtained with the Friedman's ANOVA, $X^2(5) = 45.08$, $p < .001$. In the case of words studied self-referentially, the proportion of misses was higher in the case of negative words in comparison with positive words. Moreover, irrespective of valence, words encoded in the common condition presented higher proportion of misses in contrast to words studied in the self-referential condition (see Table S6).

Correct rejections and corrected false alarm rates.

In the case of correct rejections, the Friedman's ANOVA yielded a statistically significant result, $X^2(2) = 9.60$, $p = .008$. During the test phase, participants correctly identified more new neutral words as "new" when compared to emotionally laden words (negative *vs*. neutral: $T = 44$, $z = - 2.89$, $p = .004$, r = - .36; positive *vs*. neutral: T = 57, $z = - 3.05$, $p = .002$, r = - .38). Concerning the false alarms, a similar result was obtained, $X^2(2) = 9.19$, $p = .010$., as new emotional words demonstrated higher false alarm rates than new neutral words (negative *vs*. neutral: $T = 54$, $z = - 2.87$, $p = .004$, $r = - .36$; positive *vs*. neutral: T = 58, $z = - 3.04$, $p = .002$, $r = - .38$).

Response time.

JOSs ratings. A 3 (valence: negative, neutral, positive) x 2 (source: silent, aloud) repeated-measures ANOVA revealed only a main effect of valence, $F(2, 62) = 5.74$, $p = .009$, $\eta_p^2 = .16$, $\varepsilon = .83$; the main effect of source, $F(1, 31) = 0.24$, $p = .627$, $\eta_p^2 = .01$, and the interaction between both factors, $F(2, 62) = 0.85$, $p = .431$, $\eta_p^2 = .03$, were not statistically significant. This effect revealed that participants were faster in the metamemory evaluation of positive stimuli ($M = 1630$, $SE = 102$) when compared to both negative ($M = 1829$, $SE = 115$, $p = .007$) and neutral stimuli ($M = 1857$, $SE = 138$, $p = .043$).

Source correct judgments. The 3 x 2 repeated-measures ANOVA showed a main effect of source, $F(1, 31) = 30.12$, $p < .001$, $\eta_p^2 = .49$, and a main effect of valence, $F(2, 62) = 4.23$, $p = .019$, $\eta_p^2 = .12$, but no interaction effect, $F(2, 62) = 1.03$, $p = .362$, $\eta_p^2 = .03$. Accurate source judgments were faster in the context of neutral ($M = 3149$, $SE = 156$) compared to negative words ($M = 3422$, $SE = 197$, $p = .016$). Furthermore, accurate source responses were also faster for words encoded in the self-referential condition ($M = 2987$, $SE = 154$) when compared to words encoded in the common condition ($M = 3520$, $SE = 178$, $p < .001$).

Correct rejections. Considering the response time for correct rejections, the repeated-measures ANOVA yielded a statistically significant effect of valence, $F(2, 62) = 4.84$, $p = .011$, $\eta_p^2 = .14$: participants were faster in the rejection of neutral ($M = 2131$, $SE = 94$) compared to positive words ($M = 2373$, $SE = 117$, $p = .037$).

Confidence ratings.

Source correct judgments. The analysis of the confidence ratings of accurately identified sources showed a main effect of source, $F(1, 31) = 16.32$, $p < .001$, $\eta_p^2 = .35$, and a main effect of valence, $F(2, 62) = 6.22$, $p = .003$, $\eta_p^2 = .18$, but no statistically significant interaction effect, $F(2, 62) = 1.17$, $p = .317$, $\eta_p^2 = .04$. The confidence ratings of both emotional words (negative: $M = 5.10$, $SE = 0.14$, $p = .034$; positive: $M = 5.10$, $SE = 0.15$) was significantly lower than the confidence ratings of neutral words ($M = 5.29$, $SE = 0.14$, $p = .011$).

Correct rejections. In the case of confidence ratings after correct rejections, a main effect of valence was observed, $F(2, 62) = 19.78$, $p < .001$, $\eta_p^2 = .39$, showing statistically significant differences between all valence conditions ($p < .01$): neutral ($M = 4.96$, $SE = 0.14$) > negative ($M = 4.80$, $SE = 0.15$) > positive ($M = 4.63$, $SE = 0.16$).

References

Bastin, C., Feyers, D., Souchay, C., Guillaume, B., Pepin, J. L., Lemaire, C., ... Salmon, E. (2012). Frontal and posterior cingulate metabolic impairment in the behavioral variant of frontotemporal dementia with impaired autonoetic consciousness. *Human Brain Mapping*, *33*(6), 1268–1278. https://doi.org/10.1002/hbm.21282

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*(4), 548-564. https://doi.org/10.1037/0033-295X.97.4.548

Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1121-1136. https://doi.org/10.1037//0278-7393.24.5.1121

Dodson, C. S., Prinzmetal, W., & Shimamura, A. P. (1998). Using excel to estimate parameters from observed data: An example from source memory data. *Behavior Research Methods, Instruments, & Computers, 30*(3), 517-526. https://doi.org/10.3758/BF03200685

Grainger, C., Williams, D. M., & Lind, S. E. (2016). Judgment of learning accuracy in high-functioning adolescents and adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(11), 3570–3582. https://doi.org/10.1007/s10803-016-2895-1

JASP Team. (2018). JASP (Version 0.9.0.1) [Computer software].

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

APPENDIX C

SUPPLEMENTARY MATERIALS - CHAPTER IV

Table S1

Main Characteristics of the Pool of Adjectives and Images Selected for Experiment 1 and Experiment 2

		Valence category			*F, p*
		Negative *M* (*SD*)	Neutral *M* (*SD*)	Positive *M* (*SD*)	
Images	Valence	2.61 (0.29)	4.79 (0.65)	7.58 (0.38)	849.00, < .001*
	Arousal	5.33 (0.23)	5.33 (0.29)	5.44 (0.34)	1.56, .22
	Valence – E1†	2.91 (0.52)	4.67 (0.77)	7.00 (0.61)	303.95, < .001*
	Valence – E2†	2.99 (0.58)	4.69 (0.84)	7.18 (0.59)	286.53, < .001*
		M (SD)			
Adjectives	Valence	5.03 (0.61)			
	Arousal	5.10 (0.81)			
	Frequency	8.56 (15.19)			
	N of letters	9.04 (2.10)			
	N of syllables	4.10 (0.98)			

Note. E1 = Experiment 1; E2 = Experiment 2; † Rated by the participants; *Bonferroni post-hoc comparisons revealed statistically significant differences between all valence categories.

Table S2

Response Frequencies for Each Valence and Source Condition in Experiment 1

	Negative				Neutral				Positive			
	SR	Common	DNK	New	SR	Common	DNK	New	SR	Common	DNK	New
SR	255	37	62	96	268	27	64	90	272	29	64	85
Common	62	155	92	141	63	186	72	129	69	155	74	152
New	57	57	111	2475	57	57	111	2475	57	57	111	2475

Note. DK = Do Not Know; SR = Self-Reference.

Table S3

Response Frequencies for Each Valence and Source Condition in Experiment 2

	Negative				Neutral				Positive			
	SR	Common	DNK	New	SR	Common	DNK	New	SR	Common	DNK	New
SR	248	26	83	63	261	30	69	60	258	37	60	65
Common	53	149	114	104	49	157	112	102	42	168	89	121
New	26	28	53	733	18	20	59	743	24	30	58	728

Note. DNK = Do Not Know; SR = Self-Reference.

Table S4

Multinomial Model Parameters for Item and Source Memory for each Experimental Condition (Source x Valence) for Experiment 1 and Experiment 2

	Experiment 1 – *M* (*SD*)						Experiment 2 – *M* (*SD*)					
	Self			Common			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
D	.767	.781	.794	.658	.687	.632	.828	.839	.823	.716	.726	.668
d	.630	.654	.647	.341	.466	.387	.611	.665	.669	.346	.368	.414
a/g	.275	.299	.307	.260	.237	.240	.247	.220	.225	.214	.222	.282

	Negative	Neutral	Positive	Negative	Neutral	Positive
a/g DNK	.465	.465	.455	.538	.557	.493
b	.083	.083	.083	.128	.116	.133

Note. DNK = Do Not Know.

Table S5

Results from the Pair by Pair Wilcoxon Tests Applied to the Response Time of Source Correct Responses in Experiment 2

	n	*T*	*z*	*r*	*p*
			Self		
Negative - Neutral	28	182	- 0.48	- .06	.633
Negative - Positive	28	158	- 1.03	- .14	.305
Neutral - Positive	28	200	- 0.07	- .01	.946
			Common		
Negative - Neutral	26	141	- 0.88	- .12	.381
Negative - Positive	26	175	- 0.01	- .001	.990
Neutral - Positive	27	184	- 0.12	- .02	.904
			Negative		
Self - Common	27	60	-3.10	- .42	.002*
			Neutral		
Self - Common	27	99	-2.16	- .29	.031
			Positive		
Self - Common	27	34	-3.72	- .51	< .001*

Note. * $p < .006$

Table S6

Results from the Pair by Pair Wilcoxon Tests Applied to the Proportion of Source Incorrect Responses, Do Not Know Responses, and Misses for Experiment 1 and Experiment 2

	Experiment 1					Experiment 2				
	n	*T*	*z*	*r*	*p*	*n*	*T*	*z*	*r*	*p*
					Self					
Source incorrect										
Negative - Neutral	30	45	- 1.81	- .23	.070	28	85	- 0.05	- .007	.964
Negative - Positive	30	69	- 1.08	- .14	.280	28	54	- 1.09	- .15	.275
Neutral - Positive	30	45	- 0.48	- .06	.630	28	50	- 0.60	- .08	.546
Do not know										
Negative - Neutral	30	111	- 0.16	- .02	.874	28	52	- 1.49	- .20	.136
Negative - Positive	30	81	- 0.20	- .03	.844	28	54	- 1.91	- .26	.056
Neutral - Positive	30	113	- 0.11	- .01	.915	28	72	- 0.59	- .08	.555
Misses										
Negative - Neutral	-	-	-	-	-	28	100	- 0.21	- .03	.836
Negative - Positive	-	-	-	-	-	28	89	- 0.26	- .03	.792
Neutral - Positive	-	-	-	-	-	28	141	- 0.59	- .08	.559
Item Br measure										
Negative - Neutral	30	115	- 0.72	- .09	.475	-	-	-	-	-
Negative - Positive	30	121	- 1.12	- .14	.264	-	-	-	-	-
Neutral - Positive	30	113	- 0.46	- .06	.649	-	-	-	-	-
					Common					
Source incorrect										
Negative - Neutral	30	115	- 0.04	- .005	.972	28	99	- 0.60	- .08	.551
Negative - Positive	30	114	- 0.05	- .006	.958	28	104	- 1.35	- .18	.177
Neutral - Positive	30	94	- 0.04	- .005	.968	28	81	- 0.90	- .12	.368
Do not know										
Negative - Neutral	30	49	- 1.86	- .24	.063	28	142	- 0.23	- .03	.818
Negative - Positive	30	88	- 1.55	- .20	.122	28	82	- 1.47	- .19	.143
Neutral - Positive	30	128	- 0.31	- .04	.758	28	52	- 2.22	- .30	.027

	Experiment 1					Experiment 2				
	n	*T*	*z*	*r*	*p*	*n*	*T*	*z*	*r*	*p*
Misses										
Negative - Neutral	-	-	-	-	-	28	162	- 0.01	- .001	.989
Negative - Positive	-	-	-	-	-	28	84	- 1.41	- .19	.157
Neutral - Positive	-	-	-	-	-	28	68	- 1.66	- .22	.097
Item Br measure										
Negative - Neutral	30	153	- 0.27	- .03	.788	-	-	-	-	-
Negative - Positive	30	169	- 1.05	- .14	.294	-	-	-	-	-
Neutral - Positive	30	113	- 1.06	- .14	.290	-	-	-	-	-
					Negative					
Source incorrect										
Self - Common	30	88	- 1.55	- .20	.121	28	20	- 3.05	- .41	.002*
Do not know										
Self - Common	30	35	- 2.45	- .32	.014	28	22	- 2.41	- .32	.016
Misses										
Self - Common	-	-	-	-	-	28	46	- 3.16	- .42	.002*
Item Br measure										
Self - Common	30	99	- 1.94	- .25	.052	-	-	-	-	-
					Neutral					
Source incorrect										
Self - Common	30	51	- 3.03	- .39	.002*	28	40	- 2.00	- .27	.046
Do not know										
Self - Common	30	69	- 0.72	- .09	.471	28	47	- 2.95	- .39	.003*
Misses										
Self - Common	-	-	-	-	-	28	75	- 2.75	- .37	.006
Item Br measure										
Self - Common	30	102	- 1.88	- .24	.060	-	-	-	-	-

	Experiment 1					Experiment 2				
	n	*T*	*z*	*r*	*p*	*n*	*T*	*z*	*r*	*p*
					Positive					
Source incorrect										
Self - Common	30	49	- 2.93	- .38	.003*	28	54	- 0.37	- .05	.710
Do not know										
Self - Common	30	101	- 0.84	- .11	.404	28	67	- 1.96	- .26	.050
Misses										
Self - Common	-	-	-	-	-	28	28	- 3.63	- .49	.000*
Item Br measure										
Self - Common	30	66	- 3.12	- .40	.002*	-	-	-	-	-

Note. * $p < .006$

Table S7

Means (and SDs) for the PANAS Scores Obtained in Experiment 2

	Baseline	Post-Neg block	Post-Neu block	Post-Pos block	Diff (neg block – baseline)	Diff (neu block – baseline)	Diff (pos block – baseline)
Positive affect	27.29 (5.59)	25.18 (7.16)	24.54 (6.60)	25.46 (7.05)	- 2.11 (4.87)	- 2.75 (4.59)	- 1.82 (5.30)
Negative affect	11.14 (1.41)	11.32 (2.51)	11.07 (1.96)	10.71 (1.78)	0.18 (1.91)	- 0.07 (1.49)	- 0.43 (0.92)

Note. Diff = Difference; Neg = Negative; Neu = Neutral; Pos = Positive.

Table S8

Output of the Model-based Bayesian Analysis (BF_{10}) Obtained with JASP for Experiment 1

Models	P(M)	P(M | data)	BF_M	BF_{10}	BF_{01}†	error %
			Item recognition			
Null model (incl. subject)	0.200	1.331e -7	5.326e -7	1.000	1.000	
Valence	0.200	1.081e -8	4.324e -8	0.081	12.316	0.803
Source	0.200	0.898	35.151	6.743e +6	1.483e -7	0.931
Valence + Source	0.200	0.081	0.352	606788.848	1.648e -6	2.084
Valence + Source + Valence ✻ Source	0.200	0.021	0.087	160557.386	6.228e -6	1.503
			Source recognition			
Null model (incl. subject)	0.200	7.426e -19	2.970e -18	1.000	1.000	
Valence	0.200	1.368e -19	5.470e -19	0.184	5.430	0.739
Source	0.200	0.646	7.298	8.699e +17	1.150e -18	4.598
Valence + Source	0.200	0.301	1.721	4.051e +17	2.469e -18	1.350
Valence + Source + Valence ✻ Source	0.200	0.053	0.225	7.171e +16	1.394e -17	2.214
			RT (study phase)			
Null model (incl. subject)	0.200	3.180e -5	1.272e -4	1.000	1.000	
Valence	0.200	4.512e -6	1.805e -5	0.142	7.049	0.812
Source	0.200	0.839	20.875	26388.919	3.789e -5	0.891
Valence + Source	0.200	0.146	0.684	4590.445	2.178e -4	3.475
Valence + Source + Valence ✻ Source	0.200	0.015	0.060	465.014	0.002	2.633
			RT (source correct responses)			
Null model (incl. subject)	0.200	1.339e -7	5.358e -7	1.000	1.000	
Valence	0.200	8.593e -9	3.437e -8	0.064	15.587	0.860
Source	0.200	0.927	50.913	6.922e +6	1.445e -7	1.193
Valence + Source	0.200	0.062	0.262	459163.247	2.178e -6	2.979
Valence + Source + Valence ✻ Source	0.200	0.011	0.046	84658.689	1.181e -5	1.492

Note. RT = Response Time; All models include subject; † BF_{01} was added to observe how the null model is being favored in relation to the remainder models.

Table S9

Output of the Model-based Bayesian Analysis (BF_{10}) Obtained with JASP for Experiment 2

Models	P(M)	P(M | data)	BF_M	BF_{10}	BF_{01}†	error %
			Item recognition			
Null model (incl. subject)	0.200	7.832e -7	3.133e -6	1.000	1.000	
Valence	0.200	1.082e -7	4.330e -7	0.138	7.236	2.344
Source	0.200	0.835	20.215	1.066e +6	9.382e -7	1.326
Valence + Source	0.200	0.143	0.670	183115.444	5.461e -6	1.610
Valence + Source + Valence ✻ Source	0.200	0.022	0.089	27791.013	3.598e -5	4.479
			Source recognition			
Null model (incl. subject)	0.200	1.232e -16	4.929e -16	1.000	1.000	
Valence	0.200	9.910e -18	3.964e -17	0.080	12.435	1.074
Source	0.200	0.892	33.099	7.240e +15	1.381e -16	1.177
Valence + Source	0.200	0.091	0.401	7.386e +14	1.354e -15	2.399
Valence + Source + Valence ✻Source	0.200	0.017	0.068	1.364e +14	7.334e -15	1.502
			RT (study phase)			
Null model (incl. subject)	0.200	0.007	0.027	1.000	1.000	
Valence	0.200	0.022	0.090	3.259	0.307	1.257
Source	0.200	0.164	0.784	24.385	0.041	1.050
Valence + Source	0.200	0.719	10.238	107.049	0.009	1.492
Valence + Source + Valence ✻ Source	0.200	0.089	0.388	13.179	0.076	1.282
			RT (source correct responses)			
Null model (incl. subject)	0.200	6.926e -5	2.771e -4	1.000	1.000	
Valence	0.200	6.169e -6	2.467e -5	0.089	11.228	0.869
Source	0.200	0.901	36.288	13004.239	7.690e -5	1.834
Valence + Source	0.200	0.084	0.365	1207.947	8.279e -4	1.284
Valence + Source + Valence ✻ Source	0.200	0.016	0.063	224.426	0.004	2.592

Note. RT = Response Time; All models include subject; † BF_{01} was added to observe how the null model is being favored in relation to the remainder models.

Table S10

Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Congruency x Source x Valence) for Experiment 1

	Yes – *M* (*SD*)						No – *M* (*SD*)					
	Self			Common			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
Item Pr measure (*n* = 27)	0.70 (0.22)	0.72 (0.20)	0.74 (0.24)	0.61 (0.17)	0.65 (0.15)	0.56 (0.20)	0.71 (0.18)	0.72 (0.21)	0.73 (0.16)	0.62 (0.30)	0.68 (0.22)	0.60 (0.30)
Item Br measure (*n* = 26)	0.45 (0.41)	0.48 (0.40)	0.58 (0.42)	0.26 (0.23)	0.29 (0.28)	0.28 (0.31)	0.40 (0.33)	0.48 (0.38)	0.46 (0.37)	0.50 (0.42)	0.50 (0.41)	0.43 (0.40)
SM measure (*n* = 27)	0.63 (0.31)	0.66 (0.32)	0.66 (0.32)	0.26 (0.27)	0.30 (0.34)	0.28 (0.29)	0.44 (0.30)	0.50 (0.29)	0.55 (0.24)	0.24 (0.37)	0.34 (0.40)	0.23 (0.44)

Note. SM = Source Memory.

Table S11

Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Congruency x Source x Valence) for Experiment 2

	Yes – *M* (*SD*)						No – *M* (*SD*)					
	Self			Common			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
Item Pr measure (*n* = 24)	0.76 (0.21)	0.76 (0.16)	0.72 (0.23)	0.61 (0.20)	0.69 (0.18)	0.56 (0.18)	0.70 (0.17)	0.76 (0.17)	0.70 (0.23)	0.66 (0.25)	0.64 (0.22)	0.66 (0.28)
Item Br measure (*n* = 16)	0.76 (0.34)	0.64 (0.35)	0.67 (0.34)	0.43 (0.22)	0.58 (0.35)	0.39 (0.21)	0.69 (0.32)	0.70 (0.38)	0.54 (0.32)	0.69 (0.38)	0.53 (0.39)	0.53 (0.36)
SM measure (*n* = 24)	0.61 (0.38)	0.67 (0.30)	0.62 (0.28)	0.23 (0.22)	0.25 (0.33)	0.22 (0.30)	0.48 (0.20)	0.53 (0.31)	0.45 (0.39)	0.31 (0.42)	0.28 (0.30)	0.46 (0.32)

Note. SM = Source Memory.

Table S12

Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Source x Valence) for Experiment 1 and Experiment 2 Wherein Valence was Recoded According to the Participants' Ratings

	Experiment 1 – *M* (*SD*)						Experiment 2 – *M* (*SD*)					
	Self			Common			Self			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
Item Pr measure (*n* = 29 / 28)	0.71 (0.16)	0.69 (0.14)	0.73 (0.14)	0.60 (0.17)	0.60 (0.16)	0.61 (0.21)	0.70 (0.20)	0.74 (0.14)	0.73 (0.17)	0.60 (0.19)	0.65 (0.19)	0.61 (0.19)
Item Br measure (*n* = 29 / 25)	0.35 (0.29)	0.31 (0.23)	0.44 (0.37)	0.27 (0.26)	0.24 (0.20)	0.31 (0.32)	0.48 (0.35)	0.53 (0.37)	0.48 (0.37)	0.33 (0.26)	0.45 (0.34)	0.35 (0.31)
SM measure (*n* = 29 / 28)	0.53 (0.28)	0.50 (0.22)	0.48 (0.28)	0.22 (0.24)	0.31 (0.24)	0.24 (0.28)	0.53 (0.28)	0.52 (0.23)	0.59 (0.27)	0.20 (0.23)	0.28 (0.25)	0.33 (0.24)
Extreme ratings for negative and positive pictures												
Item Pr measure (*n* = 25 / 25)	0.72 (0.17)	0.69 (0.14)	0.73 (0.17)	0.60 (0.20)	0.60 (0.17)	0.58 (0.24)	0.72 (0.26)	0.76 (0.14)	0.71 (0.29)	0.60 (0.25)	0.68 (0.17)	0.59 (0.22)
Item Br measure (*n* = 25/ 20)	0.46 (0.37)	0.32 (0.24)	0.50 (0.40)	0.29 (0.27)	0.26 (0.20)	0.35 (0.36)	0.52 (0.36)	0.57 (0.36)	0.56 (0.41)	0.38 (0.33)	0.47 (0.33)	0.41 (0.35)
SM measure (*n* = 25 / 25)	0.56 (0.30)	0.53 (0.22)	0.54 (0.36)	0.21 (0.34)	0.34 (0.23)	0.29 (0.35)	0.52 (0.27)	0.55 (0.22)	0.55 (0.39)	0.19 (0.28)	0.32 (0.22)	0.29 (0.32)

Note. SM = Source Memory.

Figure S1. Outputs of the model-based Bayesian analysis (BF_{10}) obtained with JASP for the comparison between experiments 1 and 2. (A) item memory recognition; (B) source memory recognition; (C) Response time during encoding; (D) Response time for correct source responses.

(A)

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Null model (incl. subject)	0.053	8.088e-15	1.456e-13	1.000	
Valence	0.053	7.940e-16	1.429e-14	0.098	0.722
Source	0.053	0.639	31.913	7.905e+13	1.558
Valence + Source	0.053	0.084	1.652	1.039e+13	1.229
Valence + Source + Valence ✻ Source	0.053	0.021	0.379	2.551e+12	7.839
Experimental_Group	0.053	2.129e-15	3.832e-14	0.263	1.158
Valence + Experimental_Group	0.053	2.097e-16	3.774e-15	0.026	1.195
Source + Experimental_Group	0.053	0.187	4.134	2.309e+13	1.744
Valence + Source + Experimental_Group	0.053	0.026	0.475	3.176e+12	2.275
Valence + Source + Valence ✻ Source + Experimental_Group	0.053	0.006	0.100	6.854e+11	3.027
Valence + Experimental_Group + Valence ✻ Experimental_Group	0.053	1.330e-17	2.395e-16	0.002	1.236
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	0.002	0.038	2.580e+11	16.738
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	3.725e-4	0.007	4.605e+10	2.094
Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.030	0.561	3.737e+12	2.367
Valence + Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.004	0.072	4.923e+11	1.813
Valence + Source + Valence ✻ Source + Experimental_Group + Source ✻ Experimental_Group	0.053	9.266e-4	0.017	1.146e+11	3.926
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	2.732e-4	0.005	3.378e+10	3.814
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	6.155e-5	0.001	7.611e+9	2.968
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group + Valence ✻ Source ✻ Experimental_Group	0.053	8.927e-6	1.607e-4	1.104e+9	24.192

Note. All models include subject.

(B)

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Null model (incl. subject)	0.053	5.979e-36	1.076e-34	1.000	
Valence	0.053	6.327e-37	1.139e-35	0.106	0.838
Source	0.053	0.585	25.338	9.779e+34	2.183
Valence + Source	0.053	0.159	3.402	2.659e+34	1.293
Valence + Source + Valence ✻ Source	0.053	0.011	0.196	1.803e+33	6.333
Experimental_Group	0.053	1.315e-36	2.366e-35	0.220	2.098
Valence + Experimental_Group	0.053	1.364e-37	2.455e-36	0.023	1.785
Source + Experimental_Group	0.053	0.158	3.369	2.637e+34	1.771
Valence + Source + Experimental_Group	0.053	0.043	0.813	7.225e+33	3.009
Valence + Source + Valence ✻ Source + Experimental_Group	0.053	0.003	0.047	4.346e+32	1.941
Valence + Experimental_Group + Valence ✻ Experimental_Group	0.053	1.073e-38	1.932e-37	0.002	5.989
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	0.004	0.074	6.829e+32	2.141
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	2.562e-4	0.005	4.285e+31	3.241
Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.027	0.501	4.532e+33	1.763
Valence + Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.008	0.142	1.310e+33	2.965
Valence + Source + Valence ✻ Source + Experimental_Group + Source ✻ Experimental_Group	0.053	5.116e-4	0.009	8.558e+31	4.167
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	0.002	0.041	3.837e+32	68.528
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	4.711e-5	8.480e-4	7.880e+30	4.094
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group + Valence ✻ Source ✻ Experimental_Group	0.053	1.428e-5	2.570e-4	2.388e+30	4.213

Note. All models include subject.

(C)

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Null model (incl. subject)	0.053	1.065e-7	1.916e-6	1.000	
Valence	0.053	2.552e-7	4.594e-6	2.398	6.590
Source	0.053	0.106	2.139	997571.456	3.487
Valence + Source	0.053	0.420	13.039	3.946e+6	7.197
Valence + Source + Valence ✻ Source	0.053	0.028	0.519	263140.215	2.985
Experimental_Group	0.053	5.978e-8	1.076e-6	0.562	0.602
Valence + Experimental_Group	0.053	1.378e-7	2.481e-6	1.295	1.883
Source + Experimental_Group	0.053	0.059	1.122	551218.104	1.365
Valence + Source + Experimental_Group	0.053	0.222	5.137	2.086e+6	2.201
Valence + Source + Valence ✻ Source + Experimental_Group	0.053	0.017	0.309	158367.424	5.422
Valence + Experimental_Group + Valence ✻ Experimental_Group	0.053	3.838e-8	6.909e-7	0.361	3.536
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	0.073	1.419	686584.665	2.129
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	0.006	0.102	53026.544	6.318
Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.011	0.193	99745.007	1.566
Valence + Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.041	0.771	385973.553	2.696
Valence + Source + Valence ✻ Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.003	0.054	28283.552	2.983
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	0.013	0.241	124246.941	2.981
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	0.001	0.025	12992.867	30.373
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group + Valence ✻ Source ✻ Experimental_Group	0.053	1.016e-4	0.002	954.653	4.950

Note. All models include subject.

(D)

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
Null model (incl. subject)	0.053	1.631e-13	2.936e-12	1.000	
Valence	0.053	6.178e-15	1.112e-13	0.038	1.384
Source	0.053	0.064	1.236	3.940e+11	1.437
Valence + Source	0.053	0.002	0.044	1.508e+10	1.635
Valence + Source + Valence ✻ Source	0.053	2.468e-4	0.004	1.514e+9	1.666
Experimental_Group	0.053	1.903e-12	3.425e-11	11.668	1.016
Valence + Experimental_Group	0.053	7.188e-14	1.294e-12	0.441	1.663
Source + Experimental_Group	0.053	0.765	58.606	4.691e+12	2.765
Valence + Source + Experimental_Group	0.053	0.030	0.552	1.826e+11	3.544
Valence + Source + Valence ✻ Source + Experimental_Group	0.053	0.003	0.053	1.810e+10	3.338
Valence + Experimental_Group + Valence ✻ Experimental_Group	0.053	6.363e-15	1.145e-13	0.039	2.130
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	0.003	0.056	1.885e+10	8.597
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group	0.053	2.862e-4	0.005	1.755e+9	3.705
Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.126	2.586	7.702e+11	1.951
Valence + Source + Experimental_Group + Source ✻ Experimental_Group	0.053	0.005	0.096	3.260e+10	12.706
Valence + Source + Valence ✻ Source + Experimental_Group + Source ✻ Experimental_Group	0.053	4.963e-4	0.009	3.043e+9	7.128
Valence + Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	4.560e-4	0.008	2.796e+9	2.227
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group	0.053	4.629e-5	8.333e-4	2.839e+8	2.711
Valence + Source + Valence ✻ Source + Experimental_Group + Valence ✻ Experimental_Group + Source ✻ Experimental_Group + Valence ✻ Source ✻ Experimental_Group	0.053	9.616e-6	1.731e-4	5.896e+7	3.450

Note. All models include subject.

Method

Materials: The case of a new pool of 645 European Portuguese adjectives

The European Portuguese (EP) adaptation of the Affective Norms for English Words (ANEW; Bradley & Lang, 1999; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012) has been an important step towards the development of well-controlled experimental tasks using specific affective properties in the Portuguese context. Even so, it is not always possible to obtain a satisfactory pool of stimuli which comply with some experimental needs. Particularly, for both experiments described in the main manuscript and other ongoing works in our laboratory, we could not obtain a suffice number of adjectives to conduct the experiments. Thus, we selected a new pool of 645 adjectives from *Procura-PALavras* (P-PAL), which is a recently created EP lexical database available online (Soares et al., 2010; Soares et al., 2014; Soares et al., 2018). These adjectives were selected using specific filters of frequency by million (0.43 – 125.78), number of letters (2.00 – 14.00), and number syllables (1.00 – 6.00), following similar characteristics already reported for the adjectives available in the EP ANEW.

We then conducted a pilot study with a group of 125 EP native speakers (85F; M_{age} = 31.31, *SD* = 7.46; $M_{education}$ = 17.15, *SD* = 2.71) with the purpose of collecting valence (varying from negative/unpleasant to positive/pleasant) and arousal (ranging from low/calm state to high/excited state) ratings for each adjective. The procedure was similar to the one described in the EP adaptation of the ANEW (Soares et al., 2012), thus we used the same instructions, the 9-point Self-Assessment Manikin (SAM) scale (Bradley & Lang, 1999), and both web-based and paper-and-pencil approaches. Additionally, 35 adjectives from the EP of the ANEW were used as control words. The 645 words plus the 35 control words were then randomly distributed across of five lists of 134 stimuli each for the paper-and-pencil version or 10 lists of 67 stimuli each for the web-based version. Each participant completed only one list. Each adjective was evaluated 14.40 times on average. Of note, the dominance dimension, which represents the sense of control over the environment during an emotional event (ranging from in control to out of control), was not included here as it has been less consensual in the literature (e.g., Bradley & Lang, 1994) and less relevant for the goals of our experiments.

A preliminary analysis of the results showed a positive correlation between valence and arousal when considering pleasant adjectives (r = .21, p < .001), meaning that highly pleasant words were associated with higher levels of arousal, and also a negative correlation between valence and arousal for unpleasant stimuli (r = - .30, p < .001), showing that the adjectives with lower valence values were judged as more arousing. In fact, the unpleasant adjectives revealed higher arousal ratings than pleasant adjectives [t(633) = 7.02, p < .001]. Moreover, no statistically significant differences were

found between the ratings of the 35 words common to the ANEW and the new pool. Strong correlations between the valence ($r = .96$, $p < .001$) and the arousal ratings ($r = .86$, $p < .001$) of these words were also found. Overall, although the starting point was different since we used material already available in EP instead of an adaptation from a foreign language, the results yielded similar patterns of valence and arousal as in the EP version of the ANEW. A more robust data collection and analysis for these 645 words is needed to append them to the ANEW. Nonetheless, given the preliminary results described here, we were confident to use this new pool of adjectives together with the ANEW in our experimental studies.

Data analysis

Multinomial models.

For a full and detailed description of the procedure used to conduct the multinomial approach, see the supplementary material section of Pereira, Sampaio, and Pinheiro (2019)[1]. In summary, for each experiment, a 24-parameter model was estimated based on the following parameters for each valence condition: two 'D' (probability of item memory hits); two 'd' (probability of source memory hits given the stimuli were already identified as "old"); one 'b' (probability of guessing an item as "old" or committing false alarms); three 'g' (probability of correctly guessing the source given the item was already guessed as 'old'); three 'a' (probability of correctly guessing the source given the item was correctly regarded as "old"); of note, 'a' and 'g' were constraint to be equal (Batchelder & Riefer, 1990; Dodson, Holland, & Shimamura, 1998; Leshikar, Dulas, & Duarte, 2015), giving rise to eight parameters to be estimated for each valence condition (see Tables S2 to S4). The Dodson, Prinzmetal et al. (1998) excel solver function was used to estimate the parameters and gauge model fit, using the maximum likelihood ratio and the likelihood statistic (G^2). Specifically, the G^2 statistics with a chi-square distribution (alpha = .05) was employed to compare the goodness of fit of different models.

Bayesian analysis.

Whenever possible, Bayesian analysis with JASP (JASP Team, 2018, Version 0.9.0.1) were also computed to test the results obtained with the frequentist analysis (Wagenmakers, Love et al., 2018; Wagenmakers, Marsman et al., 2018; see Supplementary Tables S8-S9 and Figure S1). Bayes factors (BF_{10}) with values above three were interpreted as being in favor of the alternative hypothesis,

[1] https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01326/full#supplementary-material

whilst values below 0.3 were deemed in favor of the null hypothesis (Jarosz & Wiley, 2014; Wagenmakers, Love et al., 2018).

Results

One sample *t*-tests (test value = 0) were computed for each item and source memory recognition measure to ascertain the performance of the participants was above chance level in Experiment 1 [all $t(29) \geq 5.47$, $p < .001$] and in Experiment 2 [all $t(27) \geq 5.81$, $p < .001$].

The multinomial model results are described separately for each experiment (see also Tables S2 to S4). In the case of Experiment 1, a 3 (valence: negative *vs*. neutral *vs*. positive) x 2 (source: self *vs*. common) Friedman's ANOVA was also run on the proportion of incorrect source responses and do not know responses, and a 3 x 2 repeated-measures ANOVA was performed on the proportion of misses. For Experiment 2, a 3 x 2 Friedman's ANOVA was run on the proportion of incorrect source responses, do not know responses, misses, correct rejections, and false alarms. If statistically significant results were found in the Friedman's ANOVA, Wilcoxon tests with Bonferroni corrections were used as follow-up tests. A non-parametric approach was used when data did not follow a normal distribution according to the skewness and kurtosis absolute values and the Shapiro-Wilk test.

As additional manipulation checks, a repeated-measures ANOVA was applied on the proportion of "yes" responses during encoding to ascertain the lists were balanced across valence conditions (for both experiments). Considering that responding "yes" or "no" during encoding may also influence recognition as stimuli that are congruent with a given category, such as self-descriptiveness, are more likely to be remembered than non-congruent stimuli (e.g., Rogers, Kuiper, & Kirker, 1977; Sanquist, Rohrbaugh, Syndulko, & Lindsley, 1980), a repeated-measures 2 (congruency: "yes" *vs*. "no" response during encoding) x 3 (valence) x 2 (source) ANOVA including the factor congruency was computed. Also, for Experiment 1, a 3 x 2 Friedman's ANOVA was computed for the number of matches between the original valence rating of the picture and the valence rating of the participants, in the case of old hits.

Experiment 1

Multinomial model results.

A G^2 value of 14.96 was obtained after running for the first time the solver function on the 24-parameter model data, which indicates not a good fit for the data as the value is above the critical chi-square of 12.59 (six degrees of freedom). To achieve a better fit to the data, we slightly adjusted the

'a'/'g' parameters of one of the valence conditions, resulting in an acceptable G^2 value of 12.56 (see Table S4 for a complete list of the parameters obtained with this solution).

In this experiment, the ANOVA results demonstrated a main effect of source, with the item memory recognition performance being superior for words encoded in a self-referential manner. To test this finding, both self-reference and common 'D' parameters were constraint to be equal for each valence condition separately, resulting in a G^2 value of 59.65, which is in accordance with the ANOVA results, given the model fit of the constraint model was significantly worse than the initial 24-parameter model, $G^2(3) = 47.09$, $p < .05$.

The analysis of the source memory recognition measures yielded a similar main effect source, but also a main effect of valence in which the source of items studied in a negative background was worst recognized than items encoded in a neutral picture background. To test the first finding, a similar approach as described for item memory was used, but instead of constraining the 'D' parameters, the 'd' parameters were the ones set to be equal. The G^2 value was 64.88, which once again indicates a worst fit than the initial model, $G^2(3) = 52.33$, $p < .05$. So, the source recognition of self-referent and common items appears to be different. Concerning the second finding, we began by constraining the 'd' parameters to be equal between negative and neutral conditions for both self-referential and common conditions separately. The result was a G^2 value of 17.27, giving rise to a $G^2(2) = 4.71$, $p > .05$. Considering that the pairwise comparisons results also yielded no significant differences between positive encoding contexts and both negative/neutral encoding contexts, we additionally equated all these parameters. The G^2 value obtained was 17.40, resulting in $G^2(3) = 4.84$, $p > .05$. These findings contrast with the ANOVA results as the parameters might not differ after all. This also supports the need to analyze other response indicators (e.g., incorrect source responses; do not know responses; misses) to clarify the source memory impairment observed in the case of negative encoding conditions. Thus, the next subsections are devoted to such analyses.

Bayesian results.

The repeated-measures ANOVA run on the item recognition measure showed a main effect of source that was characterized by a better recognition of words encoded in a self-referential manner when compared to words encoded in the common condition. The Bayesian analysis supported the main effect of source as the single effect of source model received the most positive evidence against the null model ($BF_{10} = 6.743x10^6 \pm 0.93\%$; valence + source: $BF_{10} = 606788.848 \pm 2.08\%$; valence + source + interaction: $BF_{10} = 160557.386 \pm 1.50\%$; see Supplementary Table S8). Specifically, data were

11 times ($6.743x10^{6}$/606788.848) more likely under the single model of source than the two main effects model.

In the case of the source recognition measure, while the repeated-measures ANOVA showed both a main effect of source and a main effect of valence, Bayes factors revealed that the single source model received the most evidence against the null model (BF_{10} = $8.699x10^{17}$±4.60%), followed by the two main effects model (BF_{10} = $4.051x10^{17}$±1.35%), and then by the model with the interaction term (BF_{10} = $7.171x10^{16}$±2.21%). Although data were 5.65 times ($4.051x10^{17}$/$7.171x10^{16}$) more likely under the two main effects model than the model with interaction, the single source model appeared to be 2.15 times ($8.699x10^{17}$/$4.051x10^{17}$) more probable than the two main effects model. Moreover, post-hoc comparisons using Bayesian repeated-measures ANOVA yielded unreliable evidence in favor of a difference between words studied in negative encoding conditions in comparison with words studied in neutral encoding conditions ($BF_{10(uncorrected)}$ = 1.902±0.001%). Thus, even though pairwise comparisons revealed that the source of words studied with neutral pictures was better recognized than words encoded with negative pictures, the valence effect was inconclusive when following a Bayesian approach.

Concerning the analysis of response time, the single effect of source found for mean response times during encoding was supported by the Bayesian analysis (BF_{10} = 26388.919±0.89%; valence + source: BF_{10} = 4590.445±3.48%; valence + source + interaction: BF_{10} = 465.014±2.63%). During recognition and specifically considering accurate source responses, Bayes factors were once again in favor of the single source model (BF_{10} = 6.922x106±1.19%; valence + source: BF_{10} = 459163.247±2.98%; valence + source + interaction: BF_{10} = 84658.689±1.49%).

Incorrect source responses.

The 3 x 2 Friedman's ANOVA was statistically significant, $X^2(5) = 20.60$, $p = .001$. By applying the Bonferroni correction for multiple comparisons ($p < .006$), two comparisons survived (see Table S6): the proportion of incorrect source responses was higher for words studied in the common condition in comparison with the self-referential condition, namely when the words appeared together with a neutral ($p = .002$) or a positive background ($p = .003$).

Do not know responses.

No statistically significant effects emerged in the case of proportion of do not know responses, $X^2(5) = 7.80$, $p = .168$ (see Table S6), which suggests that participants used the do not know option in a similar way regardless of the encoding task and of the emotional background.

Misses.

The repeated-measures ANOVA revealed only a main effect of source, $F(1, 29) = 29.27$, $p < .001$, $\eta_p^2 = .50$, wherein words studied in the common condition were more frequently regarded as new ($M = .31$, $SE = 0.03$) when compared to words encoded in the self-referential condition ($M = .20$, $SE = 0.02$; $p < .001$).

Additional manipulation check.

When analyzing the proportion of "yes" responses provided during encoding, the repeated-measures ANOVA revealed a main effect of source, $F(1, 29) = 47.09$, $p < .001$, $\eta_p^2 = .62$, but no main effect of valence, $F(2, 58) = 0.24$, $p = .788$, $\eta_p^2 = .01$, nor source x valence interaction, $F(2, 58) = 1.14$, $p = .326$, $\eta_p^2 = .04$. The Bonferroni-corrected post-hoc test showed that the proportion of "yes" responses was higher when participants judged words in the common condition ($M = 0.67$, $SE = 0.03$) than in the self-referential condition ($M = 0.42$, $SE = 0.02$; $p < .001$). The absence of valence effects supports the notion that the word lists were properly balanced for common and self-referential responses across valence conditions.

Considering the 2 x 3 x 2 repeated-measures ANOVA including the congruency factor (see Supplementary Table S10), in the case of item memory only a statistically significant main effect of source was found, $F(1, 26) = 20.38$, $p < .001$, $\eta_p^2 = .44$, supporting once again the advantage of self-referential encoding conditions on recognition memory. In the case of source memory, together with a main effect a source, $F(1, 26) = 10.13$, $p = .004$, $\eta_p^2 = .28$, a main effect of congruency, $F(1, 26) = 86.58$, $p < .001$, $\eta_p^2 = .77$, and congruency x source interaction effect also emerged, $F(1, 26) = 5.82$, $p = .023$, $\eta_p^2 = .18$. Bonferroni-corrected post-hoc tests showed that source memory performance for words studied in the self-referential condition was better when participants gave a congruent "yes" response ($M = 0.65$, $SE = 0.04$) in comparison when they gave a "no" response ($M = 0.50$, $SE = 0.04$; $p < .001$, 95% CI [0.07, 0.23]). In the case of words studied in the common condition no such difference was observed ($p = .738$). Moreover, regardless if the word was accompanied by a "yes" or

"no" response during encoding, words studied in the self-referential condition (yes: M = 0.65, SE = 0.04; no: M = 0.50, SE = 0.04) were associated with better source memory than words studied in the common condition (yes: M = 0.28, SE = 0.05; p < .001, 95% CI [0.30, 0.44]; no: M = 0.27, SE = 0.05; p < .001, 95% CI [0.13, 0.33]).

The 3 x 2 Friedman's ANOVA suggested no statistically significant differences in the recognition of the valence of the background pictures, $X^2(5) = 10.56$, $p = .061$. Thus, during recognition, participants did not particularly remember the emotional content of pictures from a specific source and/or valence condition. As in previous investigations (Bowen & Kensinger, 2017; Bridge, Chiao, & Paller, 2010), the results suggest that the encoding context was not readily available during the recognition of neutral stimuli. Also, in the context of the experiment, this information is relevant as it confirms that participants complied with the experimental instructions, since they were asked to passively observe the pictures and to respond only to the words, not to the pair or the association between stimuli.

Experiment 2

Multinomial model results.

An initial G^2 value of 4.71 was obtained after running the solver function for the 24-parameter model of Experiment 2, which suggests a good fit for the data as the G^2 value is below the critical chi-square value of 12.59 (six degrees of freedom). The estimated parameters are presented in Table S4.

The main effect emerging from the repeated-measures ANOVA for both item and source memory recognition was the self-referential benefit in comparison with the common condition. Thus, in different iterations, we restricted both 'D' and 'd' parameters to be equal for each valence condition. In the case of item memory, we obtained a G^2 value of 52.89, which supports the notion that words studied in the self-referential condition are better recognized than words studied in the non-self-referential condition as the original 24-parameter model presented a better fit than the restricted 21-parameter model, $G^2(3) = 48.18$, $p < .05$. When a similar analysis was applied to the 'd' parameters concerning source memory recognition, the G^2 value was 67.35, which agrees once more with the self-referential benefit found in the ANOVA as the original model showed a better fit to the data, $G^2(3) = 62.64$, $p < .05$.

Bayesian results.

In the case of item memory, the repeated-measures ANOVA revealed a main effect of source supporting the memory advantage of words studied in the self-referential condition over words studied

in the common condition, and no other main or interaction effect was detected. Bayesian analysis supported the former result as the single source model gathered the strongest evidence against the null model (BF_{10} = 1.066x106±1.33%; valence + source: BF_{10} = 183115.444±1.61%; valence + source + interaction: BF_{10} = 27791.013±4.48%; see Supplementary Table S9). The Bayes factors also supported the ANOVA analysis run on the response bias (single source: BF_{10} = 249.358±1.08%; valence + source: BF_{10} = 53.091±1.03%; valence + source + interaction: BF_{10} = 6.487±2.64%).

Regarding source memory recognition, the repeated-measures ANOVA results were in the same direction as item memory, showing only a main effect of source. In the same fashion, the single source model was the one that gathered the most positive evidence in contrast with the null model (BF_{10} = $7.240x10^{15}$±1.18%; valence + source: BF_{10} = $7.386x10^{14}$±2.40%; valence + source + interaction: BF_{10} = $1.364x10^{14}$±1.50%), in good agreement with the participant-based results. With respect to the additional analyses considering the individual valence ratings, the repeated-measures ANOVA yielded two main effects: source and valence. Bayesian analysis supported this result as the two main effects model had the most positive evidence against the null model (BF_{10} = $7.624x10^{11}$±7.78%; single valence: BF_{10} = 0.392±0.73%; single source: BF_{10} = $6.409x10^{11}$±0.90%; valence + source + interaction: BF_{10} = $1.232x10^{11}$±3.51%). This model was also 1.19 times ($7.624x10^{11}$/$6.409x10^{11}$) more likely than the single source model. More specifically, post-hoc comparisons using Bayesian repeated-measures ANOVA also supported the difference between words studied with positive background pictures and words studied with negative background pictures ($BF_{10(uncorrected)}$ = 12.372±$1.340x10^{-7}$%).

For the analysis computed on response time during encoding, two main effects emerged: source and valence. In fact, the two main effects model received the most positive evidence in comparison with the null model (BF_{10} = 107.049±1.49%; single source: BF_{10} = 24.385±1.05%; valence + source + interaction: BF_{10} = 13.179±1.28%). When testing the model with an interaction, the two main effects model remained as the most probable model (107.049/13.179: around 8.12 times more likely than the interaction model). In the case of the response time analysis of correct source responses, the Bayesian analysis (single source: BF_{10} = 13004.239±1.83%; valence + source: BF_{10}= 1207.947±1.28%; valence + source + interaction: BF_{10} = 224.426±2.59%) was also in line with the non-parametric analysis described in the main text.

Incorrect source responses.

The 3 x 2 Friedman's ANOVA was statistically significant, $\chi^2(5) = 14.85$, $p = .011$. The proportion of incorrect source responses was higher for words encoded in the common condition than

words encoded in the self-referential condition, specifically when negative pictures were presented in the background (p = .002; see Table S6).

Do not know responses.

Contrary to Experiment 1, Experiment 2 yielded a statistically significant 3 x 2 Friedman's ANOVA, $X^2(5) = 21.74$, $p = .001$. The proportion of do not know responses was higher for words studied with the common judgments when compared to the self-referential judgments particularly in the block with neutral pictures (p = .003; see Table S6).

Misses.

The 3 x 2 Friedman's ANOVA was statistically significant, $X^2(5) = 32.92$, $p < .001$. As in Experiment 1, the Wilcoxon tests suggested that words studied in the common condition received a higher proportion of "new" attributions in comparison with words judged self-referentially regardless of the emotional valence of the background picture (see Table S6; of note, the effect detected for the neutral block was marginally significant, p = .006).

Correct rejections and false alarms.

No statistically significant differences emerged between negative *vs*. neutral *vs*. positive blocks, $X^2(2) = 0.09$, $p = .956$. Thus, the proportion of correct rejections and false alarms was similar across blocks.

Additional manipulation check.

The results obtained for the proportion of "yes" responses were in the same direction as in Experiment 1. Thus, a main effect of source was observed, $F(1, 27) = 34.77$, $p < .001$, $\eta_p^2 = .56$, with no significant valence effect, $F(2, 54) = 0.24$, $p = .785$, $\eta_p^2 = .01$, nor interaction effect, $F(2, 54) = 1.28$, $p = .286$, $\eta_p^2 = .05$. Once more, the pairwise comparisons revealed that the proportion of "yes" responses was increased in the common condition ($M = 0.71$, $SE = 0.03$) relative to the self-referential condition ($M = 0.45$, $SE = 0.03$; $p < .001$).

Regarding the analysis including the congruency factor (see Supplementary Table S11) and considering item memory recognition, the results were in the same direction as Experiment 1, thus only the main effect of source was statistically significant, $F(1, 23) = 18.57$, $p < .001$, $\eta_p^2 = .45$. In the case of source memory recognition, the repeated-measures ANOVA revealed a main effect of source,

$F(1, 23) = 38.94$, $p < .001$, $\eta_p^2 = .63$, but also a congruency x source interaction, $F(1, 23) = 18.74$, $p < .001$, $\eta_p^2 = .45$. Follow-up pair-by-pair comparisons indicated that for words self-referentially encoded, source memory performance was better when the words matched the study category, i.e., when participants responded "yes" during encoding (yes: $M = 0.64$, $SE = 0.05$; no: $M = 0.49$, $SE = 0.05$; $p = .004$, 95% CI [0.05, 0.24]). The reverse pattern was observed in the case of words studied in the common condition, that is source memory was worst when there was a match with the common category (yes: $M = 0.23$, $SE = 0.04$; no: $M = 0.35$, $SE = 0.05$; $p = .027$, 95% CI [0.02, 0.22]. More importantly, regardless if the participants responded "yes" or "no" during encoding, the source of words studied in the self-referential condition (yes: $M = 0.64$, $SE = 0.05$; no: $M = 0.49$, $SE = 0.05$) was better recognized than words studied in the common condition (yes: $M = 0.23$, $SE = 0.04$; $p < .001$, 95% CI [0.32, 0.49]; no: $M = 0.35$, $SE = 0.05$; $p = .037$, 95% CI [0.01, 0.26]).

Comparison across experiments

The results of Experiments 1 and 2 were included in the same cross-sectional analysis to test whether participants were approaching the experiments in a similar way, and whether differences in the experimental design could emerge. Considering that the participants' sociodemographic characteristics, the stimuli, and the general procedures shared commonalities across Experiment 1 and Experiment 2, we conducted a mixed-design ANOVA (within-subjects factors: source and valence; between-subject factor: mixed *vs.* blocked design) on the corrected item and source recognition measures, and on the response times during encoding and correct source memory recognition. In the case of both item and source memory corrected measures, only a main effect of source was observed [item memory: $F(1, 56) = 57.40$, $p < .001$, $\eta_p^2 = .51$, $BF_{10} = 7.905x10^{13}\pm1.56\%$; source memory: $F(1, 56) = 149.41$, $p < .001$, $\eta_p^2 = .73$, $BF_{10} = 9.779x10^{34}\pm2.18\%$], replicating the memory advantage in the case of words encoded self-referentially (item memory: $M = 0.72$, $SE = 0.02$; source memory: $M = 0.53$, $SE = 0.03$) in comparison with the common judgment condition (item memory: $M = 0.61$, $SE = 0.02$; $p < .001$, $d = 1.00$, 95% CI [0.08, 0.14]; source memory: $M = 0.26$, $SE = 0.02$; $p < .001$, $d = 1.62$, 95% CI [0.22, 0.31]). The slower response time for words that were self-referentially encoded ($M = 2003$, $SE = 30.31$) in contrast to words judged in the common condition ($M = 1904$, $SE = 30.10$; $p < .001$, $d = 0.71$, 95% CI [61.98, 135.63]) was again observed, $F(1, 56) = 28.89$, $p < .001$, $\eta_p^2 = .34$. Additionally, a main effect of valence emerged, $F(2, 112) = 4.41$, $p = .014$, $\eta_p^2 = .07$, showing that, in general, the response time for words studied in the negative condition ($M = 1990$, $SE = 32.42$) was significantly slower than words studied in the neutral condition ($M = 1925$, $SE = 32.28$; $p = .006$, $d =$

0.41, 95% CI [15.59, 114.54]; valence + source: BF_{10} = 3.946x10^{6}±7.18%). Although this effect was not detected in Experiment 1, it was present in Experiment 2. When analyzing the response times for source recognition hits, a main effect of source was observed, $F(1, 51) = 47.51$, $p < .001$, $\eta_p^2 = .48$, BF_{10} = 3.946x10^{6}±7.18%, showing that the source recognition of words encoded self-referentially was faster (M = 2256, SE = 77.20) than words encoded in the common judgment task (M = 2739, SE = 118.21; $p < .001$, d = 0.96, 95% CI [342.25, 623.54]). However, a main effect of the experimental design was also observed, $F(1, 51) = 9.49$, $p = .003$, $\eta_p^2 = .16$, showing that in general participants were slower in the mixed design test phase (M = 2785, SE = 130.95) than in the blocked design test phase (M = 2209, SE = 133.45; $p = .003$, d = 0.42, 95% CI [200.63, 951.34]; source + experimental design: BF_{10} = 4.691x10^{12}±2.77%). The difference in response time for source hits might be accounted by the fact that participants had to recall not only the source of the item but also the valence of the picture presented together with it. This was an experimental control feature that was only present in the mixed design experiment. Overall, as no interaction effect emerged between the type of experimental design and source or valence, this suggests that using a mixed or a blocked design did not significantly change how participants approached the tasks.

References

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*(4), 548-564. https://doi.org/10.1037/0033-295X.97.4.548

Bowen, H. J., & Kensinger, E. A. (2017). Recapitulation of emotional source context during memory retrieval. *Cortex, 91*, 142–156. https://doi.org/10.1016/j.cortex.2016.11.004

Bradley, M., M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. https://doi.org/10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Bridge, D. J., Chiao, J. Y., & Paller, K. A. Emotional context at learning systematically biases memory for facial information. *Memory & Cognition, 38*(2), 125-133. https://doi.org/10.3758/MC.38.2.125

Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1121-1136. https://doi.org/10.1037//0278-7393.24.5.1121

Dodson, C. S., Prinzmetal, W., & Shimamura, A. P. (1998). Using excel to estimate parameters from observed data: An example from source memory data. *Behavior Research Methods, Instruments, & Computers, 30*(3), 517-526.

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting bayes factors. *Journal of Problem Solving, 7*(1), 2-9. https://doi.org/10.7771/1932-6246.1167

JASP Team. (2018). JASP (Version 0.9.0.1) [Computer software].

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150.

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy in internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology, 10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1980). Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology, 17*(6), 568-576. https://doi.org/10.1111/j.1469-8986.1980.tb02299.x

Soares, A. P., Comesaña, M., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., ...Machado, J. (2010). P-PAL: Uma base lexical com índices psicolinguísticos do Português Europeu [P-PAL: A European Portuguese lexical database]. *Linguamática, 2*(3), 67-72.

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Soares, A. P., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., França, P., ...Comesaña, M. (2014). Procura-PALavras (P-PAL): A new measure of word frequency for contemporary European

Portuguese. *Psicologia: Reflexão e Crítica, 27*(1), 110-123. https://doi.org/10.1590/S0102-79722014000100013

Soares, A. P., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., Machado, J., ...Perea, M. (2018). Procura-PALavras (P-PAL): A web-based interface for a new European Portuguese lexical database. *Behavior Research Methods, 50*(4), 1461-1481. https://doi.org/10.3758/s13428-018-1058-z

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Wagenmakers, E-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ...Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*(1), 58-78. https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ...Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25*(1), 35-57. https://doi.org/10.3758/s13423-017-1343-3

APPENDIX D

SUPPLEMENTARY MATERIALS - CHAPTER V

Table S1

Main Characteristics of the Selected Lists of Adjectives

	Valence category			*F, p*
	Negative - *M* (*SD*) *n* = 120	Neutral - *M* (*SD*) *n* = 120	Positive - *M* (*SD*) *n* = 120	
Valence	2.87 (0.43)	5.00 (0.55)	7.19 (0.44)	2435.95, < .001*
Arousal	5.15 (0.73)	4.92 (0.97)	4.95 (0.34)	2.12, .12
Frequency	4.99 (12.47)	7.76 (12.27)	5.99 (11.34)	1.62, .20
N of letters	8.44 (2.18)	8.44 (2.21)	8.59 (2.22)	0.19, .83
N of syllables	3.81 (0.99)	3.78 (1.06)	3.91 (1.03)	0.50, .61

Note. Adjectives were selected from two different sources: the European Portuguese (EP) version of the Affective Norms for English Words (ANEW; Bradley & Lang, 1999; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), and from an ongoing study that aims to extend the EP version of the ANEW. Specifically, the new list of 645 adjectives was extracted from the EP lexical database *Procura-PALavras* (P-PAL; Soares et al., 2010; Soares et al., 2014; Soares et al., 2018). Following the procedures of the adaptation of the EP version of the ANEW (Soares et al., 2012), these adjectives were then rated by 125 EP native speakers using the 9-point Self-Assessment Manikin (SAM) scale (Bradley & Lang, 1999). In the case of valence, 1 corresponds to the most negative ratings, while 9 corresponds to the most positive ratings. In the case of arousal, ratings closer to 1 indicate the lowest arousal levels, whereas ratings closer to 9 indicate the highest arousal levels. On average, each adjective was rated 14.40 times. *Bonferroni post-hoc comparisons revealed statistically significant differences between all valence categories.

Table S2

Mean Number of Accepted Trials in the ERP Analysis During Study and Test Phases as a Function of Encoding Task and Valence Conditions

			Valence category								
			Negative			Neutral			Positive		
			M	*SD*	*Range*	*M*	*SD*	*Range*	*M*	*SD*	*Range*
Study phase	Word encoding JOSs ratings	Self	37.50	1.95	34-40	37.23	2.14	33-40	38.00	2.12	32-40
		Common	37.55	2.56	32-40	37.55	2.54	32-40	37.73	2.10	33-40
		Self	35.14	3.44	27-40	34.32	3.87	25-40	34.18	4.85	23-40
		Common	35.05	3.26	28-40	35.09	3.56	27-40	36.09	3.24	28-40
Test phase	Self *vs.* Common *vs.* New	Self	21.95	5.51	12-32	27.36	4.75	17-34	27.23	5.46	16-36
		Common	20.27	5.54	12-35	21.68	5.93	12-33	19.36	4.28	12-28
		New	28.82	4.53	20-37	31.82	3.33	22-37	29.59	4.82	19-37

Note. ERP = Event-related Potential; JOSs = Judgments of Source.

Table S3

Response Frequencies for Each Encoding Task and Valence Conditions in the Test Phase

	Negative				Neutral				Positive			
	SR	Common	DNK	New	SR	Common	DNK	New	SR	Common	DNK	New
SR	530	109	93	148	651	88	51	90	659	72	53	96
Common	78	472	110	180	54	518	88	220	106	465	108	201
New	35	58	75	711	19	45	42	773	58	39	65	718

Note. DNK = Do Not Know; SR = Self-Reference.

Table S4

Multinomial Model Parameters for Item and Source Memory as a Function of Encoding Task and Valence Conditions

	Self-reference			Common		
	Negative	Neutral	Positive	Negative	Neutral	Positive
D	0.792	0.884	0.866	0.735	0.716	0.720
d	0.669	0.794	0.777	0.576	0.616	0.576
a/g	0.239	0.193	0.347	0.385	0.478	0.317
	Negative		Neutral		Positive	
a/g DNK	0.376		0.329		0.340	
b	0.191		0.121		0.184	

Note. DNK = Do Not Know.

Table S5

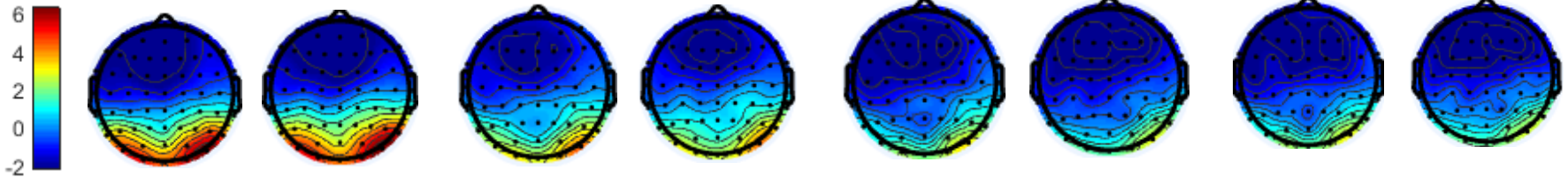
Results from the Pair by Pair Wilcoxon Tests Applied to the Proportion of Incorrect Source Responses

	n	*T*	*z*	*r*	*p*
			Self		
Negative - Neutral	22	65	- 1.53	- .23	.127
Negative - Positive	22	44	- 2.91	- .44	.022
Neutral - Positive	22	61	- 1.40	- .21	.163
			Common		
Negative - Neutral	22	41	- 2.22	- .33	.027
Negative - Positive	22	68	- 1.09	- .16	.276
Neutral - Positive	22	25	- 3.00	- .45	.003*
			Negative		
Self - Common	22	53	- 1.44	- .22	.149
			Neutral		
Self - Common	22	40	- 2.26	- .34	.024
			Positive		
Self - Common	22	42	- 1.66	- .25	.096

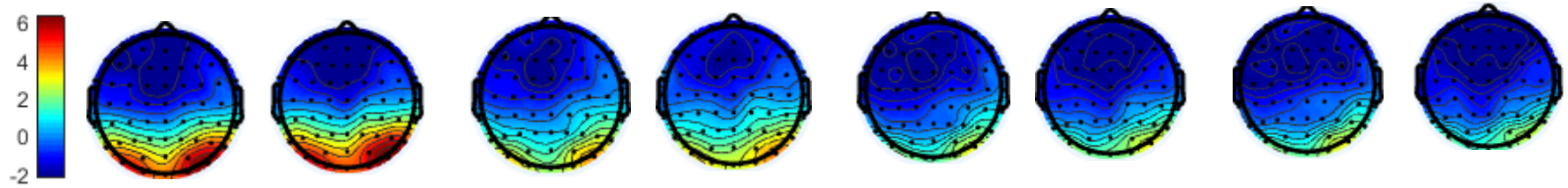
Note. * $p < .006$

Table S6

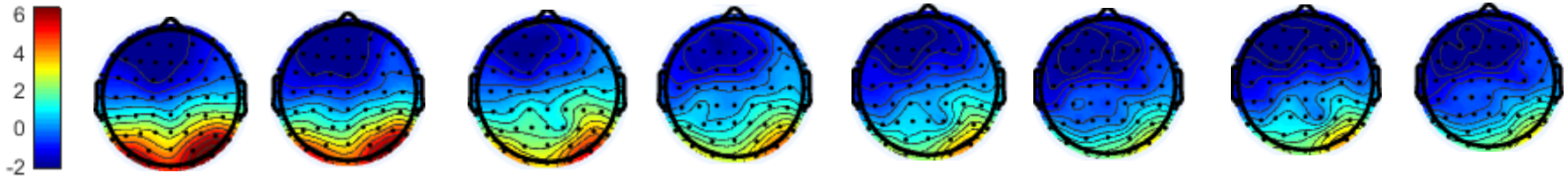
Mean Amplitude (μV) for the JOS Ratings in the Four Time Windows Analyzed for Middle Frontal (F1/Fz/F2) and Middle Parietal (P1/Pz/P2) ROIs

	300 – 500 ms - *M* (*SD*)				500 – 700 ms - *M* (*SD*)				700 – 1000 ms - *M* (*SD*)				1000 – 1300 ms - *M* (*SD*)			
	Self		Common		Self		Common		Self		Common		Self		Common	
Valence	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P
Negative	-2.32 (4.42)	1.72 (3.80)	-2.75 (4.18)	2.16 (4.23)	-1.84 (4.49)	0.55 (3.51)	-2.35 (3.91)	0.84 (3.65)	-2.14 (4.27)	-0.23 (3.50)	-2.73 (4.17)	-0.39 (3.97)	-2.05 (4.22)	-0.08 (3.20)	-2.50 (3.99)	0.34 (3.76)
Neutral	-2.07 (4.53)	1.86 (4.79)	-2.25 (3.62)	2.44 (3.39)	-1.83 (4.39)	0.60 (4.14)	-1.85 (3.81)	1.06 (3.54)	-2.09 (4.01)	-0.31 (4.43)	-2.36 (3.94)	-0.09 (3.81)	-2.16 (4.26)	-0.32 (4.27)	-2.44 (3.71)	-0.45 (4.42)
Positive	-1.99 (3.74)	2.95 (3.93)	-2.07 (4.00)	2.00 (3.58)	-1.38 (3.95)	1.57 (3.65)	-1.58 (4.50)	1.04 (3.81)	-1.89 (3.73)	0.52 (3.38)	-2.05 (4.22)	-0.05 (3.20)	-2.13 (3.52)	-0.34 (3.76)	-2.05 (4.29)	-0.02 (3.50)
Effects from RMANOVAs	ROI: $F(4, 84) = 54.93$, $p < .001$, $\eta_p^2 = .72$, $\varepsilon = .29$ Laterality: $F(2, 42) = 9.60$, $p < .001$, $\eta_p^2 = .31$ ROI x laterality: $F(8, 168) = 4.14$, $p = .009$, $\eta_p^2 = .17$, $\varepsilon = .39$ Source x valence x ROI: $F(8, 168) = 2.83$, $p = .046$, $\eta_p^2 = .12$, $\varepsilon = .37$				ROI: $F(4, 84) = 18.93$, $p < .001$, $\eta_p^2 = .47$, $\varepsilon = .28$ Laterality: $F(2, 42) = 8.08$, $p = .001$, $\eta_p^2 = .28$ ROI x laterality: $F(8, 168) = 2.61$, $p = .048$, $\eta_p^2 = .11$, $\varepsilon = .45$				ROI: $F(4, 84) = 11.95$, $p = .002$, $\eta_p^2 = .36$, , $\varepsilon = .28$ Laterality: $F(2, 42) = 9.62$, $p < .001$, $\eta_p^2 = .31$ ROI x laterality: $F(8, 168) = 3.19$, $p = .020$, $\eta_p^2 = .13$, $\varepsilon = .46$				ROI: $F(4, 84) = 13.72$, $p = .001$, $\eta_p^2 = .40$, , $\varepsilon = .29$ Laterality: $F(2, 42) = 7.26$, $p = .002$, $\eta_p^2 = .26$ Valence x laterality†: $F(4, 84) = 3.00$, $p = .023$, $\eta_p^2 = .13$ ROI x laterality: $F(8, 168) = 3.48$, $p = .010$, $\eta_p^2 = .14$, $\varepsilon = .51$			
Topographical maps - Negative Words																

Topographical maps - Neutral Words



Topographical maps - Positive Words



Note. F = Frontal; P = Parietal; RMANOVAs = Repeated-Measures Analysis of Variance; ROI = Region of Interest. † Follow-up post-hoc tests yielded no differences between valence categories, yet showed that for both neutral and positive words the mean amplitudes were less negative at right electrode locations (neutral: M = - 0.60, SE = 0.65; positive: M = - 0.11, SE = 0.56) in contrast to left (neutral: M = - 1.34, SE = 0.63; p = .038; positive: M = - 1.28, SE = 0.56; p < .001) and middle electrode locations (neutral: M = - 1.40, SE = 0.63; p = .019; positive: M = - 0.83, SE = 0.68; p = .025).

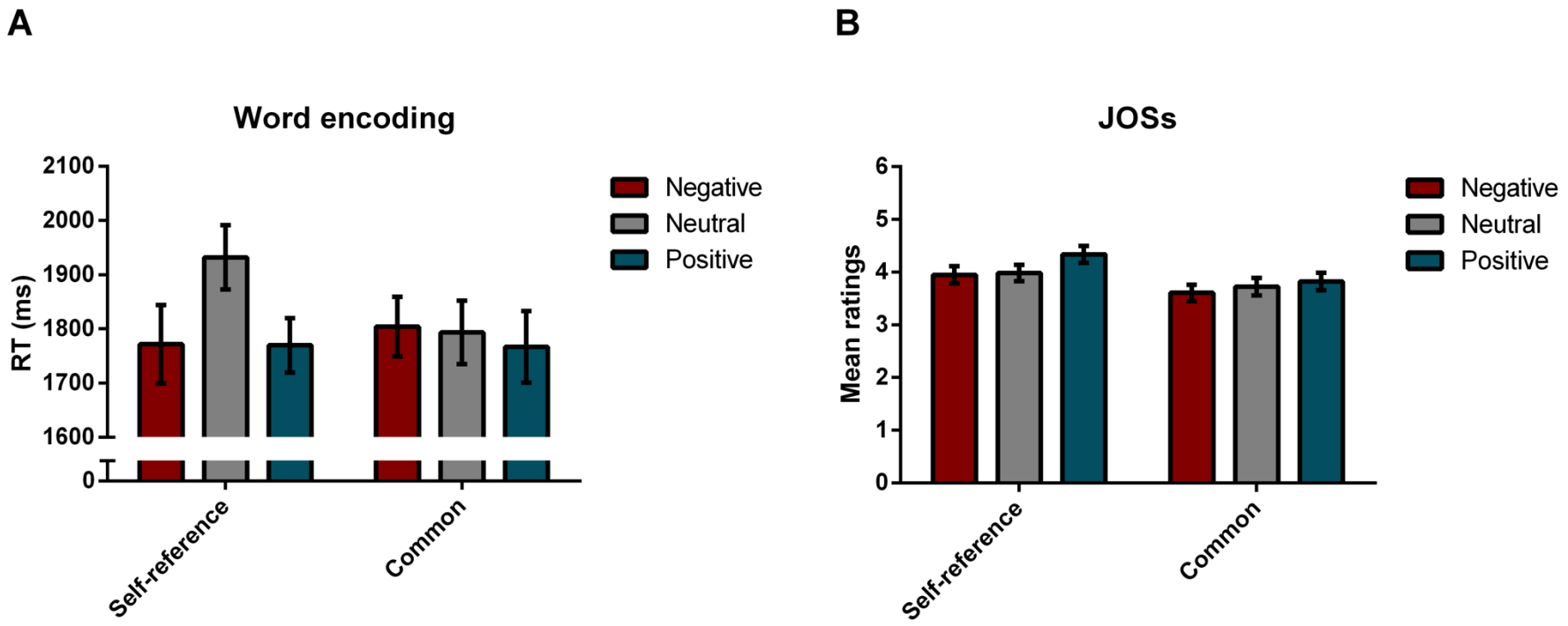


Figure S1. In panel A, the mean response time (RT) obtained during word encoding is plotted on the y-axis as a function of encoding task (self-reference *vs.* common) and valence (negative *vs.* neutral *vs.* positive). In panel B, a plot of the mean judgments of source (JOSs) ratings obtained during the study phase is also presented. The error bars represent the standard errors of the means.

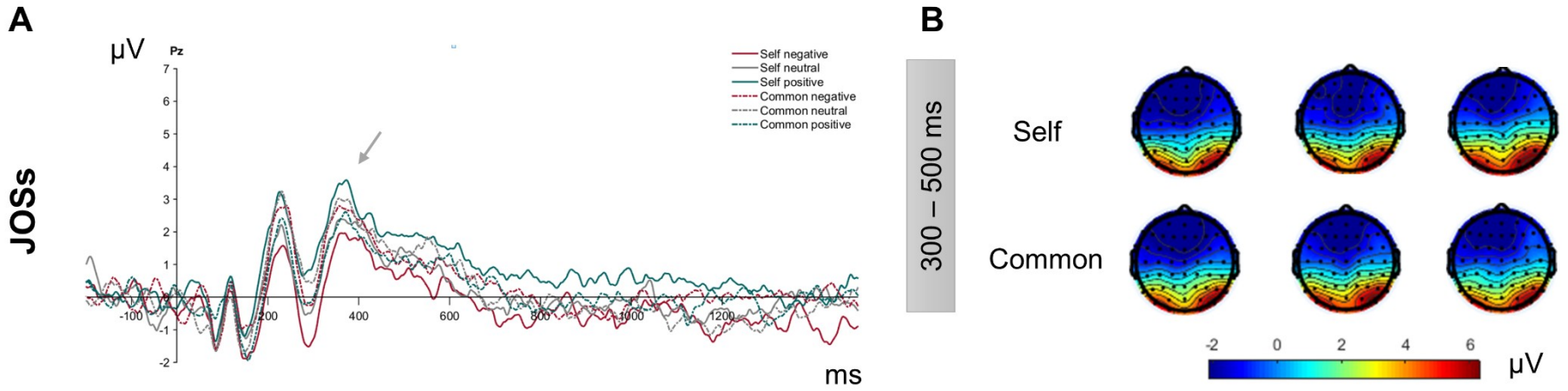


Figure S2. Illustration of grand average waveforms obtained during judgments of source (JOSs; panel A) together with the topographic maps (panels B) considering each encoding task (self *vs*. common) x valence (negative *vs*. neutral *vs*. positive) condition. It is worth noting that most ERP components underlying JOSs ratings in this study were not significantly modulated by valence and/or self-relevance. Even so, considering the first latency window (300 – 500 ms) and parietal electrodes (see the gray arrow), positive words studied in the self-referential condition were associated with a larger mean amplitude than positive words studied in the common condition.

Method

Data analysis

Multinomial model.

For each valence condition (negative *vs*. neutral *vs*. positive), a 24-parameter model was estimated given the following parameters: two 'D' parameters corresponding to the probability of correct item memory responses; two 'd' parameters corresponding to the probability of correct source responses if the items were accurately judged as 'old'; one 'b' parameter corresponding to the probability of saying an item was 'old' by guessing, or to the probability of saying that a new item was studied before (i.e., false alarms); three 'g' parameters corresponding to the probability of correctly guessing both item and source memory responses; three 'a' parameters corresponding to the probability of correctly guessing source memory responses if items were already judged as 'old'. In particular, eight parameters were estimated for each valence condition as parameters 'a' and 'g' were constrained to be equal following previous studies (Batchelder & Riefer, 1990; Dodson, Holland, & Shimamura, 1998; Leshikar, Dulas, & Duarte, 2015). This estimation was conducted using an excel solver function (Dodson, Prinzmetal et al., 1998). Also, the same function was used to compare the goodness of fit of different models by applying the maximum likelihood ratio and the likelihood statistic (G^2) with a chi-square distribution (alpha = .05). Of note, a similar multinomial approach was described elsewhere (Pereira, Sampaio, & Pinheiro, 2019)[1].

Goodman-Kruskal gamma correlation.

Gamma correlations vary between – 1 and 1 and inform about the degree of association between the participants' memory performance and their prediction regarding which items will be later remembered or forgotten. A good metamnemonic resolution, that is a strong relation between objective memory performance and subjective memory predictions, is shown by large and positive gamma correlation values. On the contrary, gamma correlation values equal, close or below zero stand against the relationship between memory performance and prediction.

Given that the judgments of source (JOSs) ratings were provided on a 6-point scale, ranging from 1 ("Sure I will not remember the judgment") to 6 ("Sure I will remember the judgment"), four different frequency values – *a, b, c, d* - were extracted from each experimental condition and from each participant: (*a*) the number of correct source memory predictions for items that received JOSs ratings between 4 and 6 during the study phase, and whose source was later remembered; (*b*) the number of

[1] https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01326/full#supplementary-material

incorrect source memory predictions for items that also received JOSs ratings between 4 and 6, but whose source was later forgotten; (*c*) the number of incorrect source memory predictions for items rated between 1 and 3 whose source was later remembered; (*d*) the number of correct source predictions for items also rated between 1 and 3, but whose source was later forgotten. Then, the former values were used in the gamma correlation formula: $G = (ad - bc)/(ad - bc)$. It is worth noting that when some of the elements of the formula were equal to zero, making impossible to obtain a gamma estimate, we added 0.5 to each frequency (*a, b, c, d*) and then divided the result by the total number of judgments plus one ($N + 1$), following the recommendation of Snodgrass and Corwin (1988) (see Bastin et al., 2012; Castel, Rhodes, & Friedman, 2013; Grainger, Williams, & Lind, 2016 for gamma correlations examples).

Results

In a first step, we computed one sample *t*-tests (test value = 0) for each item and source memory measure to evaluate if the participants' performance was above chance level. In all cases, the above chance level performance was verified, $t(21) \geq 10.14$, $p < .001$. Regarding the results from the multinomial model approach, the response frequencies, and the estimated parameters as a function of encoding task and valence conditions are presented in Tables S3 and S4, respectively. For additional analysis of the proportions of "do not know" responses, misses, correct rejections, false alarms, and "yes" responses during the study phase, a 2 (self-reference: self *vs*. common) x 3 (valence: negative *vs*. neutral *vs*. positive) repeated-measures ANOVA was used. Also, mean response times were extracted considering the period of word encoding and JOSs ratings (trials with response times below 200 ms were not included; Leshikar et al., 2015). Data were also submitted to a 2 x 3 repeated-measures ANOVA. As an exception, a 2 x 3 Friedman's ANOVA was run on the proportion of incorrect source memory responses as data were not normally distributed following the Shapiro-Wilk test and the skewness/kurtosis absolute values. In this context, when statistically significant results were found, Wilcoxon tests with Bonferroni corrections were used as follow-up tests. Moreover, similar 2 x 3 repeated-measures ANOVAs were run on the mean JOSs ratings and Goodman-Kruskal gamma correlations. Regarding JOSs, we also present an analysis of the electrophysiological response including four consecutive time windows (following a previous study by Müller et al., 2016): 300 – 500; 500 – 700; 700 – 1000; 1000 – 1300 ms.

Multinomial model results

Initially, a G^2 value of 17.58 was obtained when using the solver function on the 24-parameter model. As this value was above the critical chi-square of 12.59 (six degrees of freedom), this was indicative of poor fit to the data. In this case, the 'a'/'g' parameters of one of the valence conditions were slightly adjusted until a value below 12.59 was obtained (G^2 = 12.58). Table S4 presents the estimated parameters obtained with this solution.

The participant-based results using repeated-measures ANOVA showed that words studied under the self-referential condition resulted in better item memory performance than words studied in the common condition. Moreover, while no differences between emotional categories were observed for words encoded in the common condition, in the self-referential condition there were differences for all valence categories: neutral > positive > negative. To test the first result using the model-based data, we started by equating the 'D' parameters of the self-referential and common condition for each valence category separately. A G^2 value of 283.50 was obtained, which yielded a worse model fit when compared to G^2 value of the initial 24-parameter model, $G^2(3) = 270.92$, $p < .05$. Thus, the model- and participant-based results were in the same direction. The second result was tested by constraining the 'D' parameters of the common condition to be equal across valence categories. The G^2 value was 14.73, which resulted in a $G^2(2) = 2.15$, $p > .05$, suggesting no significant differences between the constraint model and the 24-parameter model, thus supporting the initial ANOVA result. To test the neutral > positive > negative result for the self-referential condition, we then decided to constraint the 'D' parameters of this condition to be equal across valence categories. A G^2 value of 83.21 was obtained, which indicates a worse model fit when compared to the G^2 value of the initial 24-parameter model, $G^2(2) = 70.63$, $p < .05$; this outcome also supports the ANOVA result as the parameters may differ after all. Taken together, model-based results supported the participant-based results in the case of item memory recognition.

Considering source memory performance, the participant-based results demonstrated a behavioral dissociation: the source memory performance for neutral and positive words encoded in the self-referential condition was improved compared to the performance for negative words, whilst in the common judgment task positive words were associated with a worse source memory performance when compared to neutral words. In the model-based approach, we started by constraining the neutral and positive 'd' parameters referring to the self-referential condition to be equal. As an outcome, the G^2 value was 13.28, which supported the behavioral dissociation for the self-referential condition as the constraint model also showed a good fit to the data and did not statistically differ from the 24-

parameter model, $G^2(1) = 0.70$, $p > .05$. In the case of the common condition, we rather forced the neutral and positive 'd' parameters to be equal. A G^2 value of 14.59 was obtained, thus suggesting that the parameters may not differ, $G^2(1) = 2.01$, $p > .05$. Following this result, we constrained all 'd' parameters of the common condition to be equal across valence categories, and a G^2 value of 14.09 was obtained, supporting the absence of differences between negative and both neutral and positive words, $G^2(2) = 1.51$, $p > .05$. In this context, the model-based findings partially supported the participant-based findings.

Response time during word encoding

The analysis of the mean response time during word encoding yielded a main effect of valence, $F(2, 42) = 13.81$, $p < .001$, $\eta^2_p = .40$, and a self-reference x valence effect, $F(2, 42) = 7.59$, $p = .004$, $\eta^2_p = .27$, $\varepsilon = .78$ [source: $F(1, 21) = 1.89$, $p = .184$, $\eta^2_p = .08$]. Considering neutral items, participants took longer to respond in the self-referential condition ($M = 1932$, $SE = 59.12$) when compared to the common condition ($M = 1793$, $SE = 58.69$; $p < .001$, 95% CI_{diff} [75.65, 201.70]; see Figure S1A). Additionally, in the self-referential condition, participants were faster to respond to emotional (negative: $M = 1772$, $SE = 72.66$; $p = .002$, 95% CI_{diff} [58.61, 262.17]; positive: $M = 1770$, $SE = 50.28$; $p < .001$, 95% CI_{diff} [100.81, 223.99]) than to neutral words ($M = 1932$, $SE = 59.12$; see Figure S1A).

Considering the mean response time during the JOSs ratings, only a main effect of valence was observed, $F(2, 42) = 5.28$, $p = .009$, $\eta^2_p = .20$ [self-reference: $F(1, 21) = 0.24$, $p = .628$, $\eta^2_p = .01$; self-reference x valence: $F(2, 42) = 2.01$, $p = .147$, $\eta^2_p = .09$]. The post-hoc comparisons revealed that participants were faster at conducting metamnemonic judgments in the case of positive ($M = 874$, $SE = 54.87$) relative to neutral words ($M = 925$, $SE = 52.78$; $p = .008$, 95% CI_{diff} [11.90, 91.01]).

Incorrect source memory responses

The 2 x 3 Friedman's ANOVA was statistically significant, $X^2(5) = 12.56$, $p = .028$. After applying the Bonferroni correction for multiple comparisons ($p < .006$), only one comparison remained statistically significant (see Table S5): in the particular case of words encoded in the common condition, positive words were related to a higher proportion of incorrect source responses than neutral words ($p = .003$). This finding fits well with the ANOVA result for the source memory performance in the case of the common condition, as the performance was better for neutral words than for positive words.

Do not know responses

The 2 x 3 repeated-measures ANOVA yielded a main effect of self-reference, $F(1, 21) = 15.51$, $p = .001$, $\eta_p^2 = .43$, and a main effect of valence, $F(2, 42) = 7.56$, $p = .002$, $\eta_p^2 = .27$, but no interaction effect, $F(2, 42) = 2.42$, $p = .127$, $\eta_p^2 = .11$, $\varepsilon = .62$. Specifically, the participants used less the "do not know option" during the recognition test of self-referentially encoded words ($M = .08$, $SE = 0.01$) when compared to words encoded in the common condition ($M = .12$, $SE = 0.02$; $p = .001$, 95% CI_{diff} [0.02, 0.07]). Additionally, irrespective of the encoding condition, negative words ($M = .12$, $SE = 0.01$) were associated with a higher proportion of "do not know" responses relative to both neutral ($M = .08$, $SE = 0.01$; $p = .003$, 95% CI_{diff} [0.01, 0.07]) and positive words ($M = .09$, $SE = 0.01$; $p = .036$, 95% CI_{diff} [0.002, 0.06]).

Misses

The 2 x 3 repeated-measures ANOVA revealed a main effect of self-reference, $F(1, 21) = 44.68$, $p < .001$, $\eta_p^2 = .68$, and an interaction effect, $F(2, 42) = 7.98$, $p = .001$, $\eta_p^2 = .28$ (valence: $F(2, 42) = 1.74$, $p = .188$, $\eta_p^2 = .08$). Irrespective of valence, a greater proportion of misses was observed in the common condition ($M = .23$, $SE = 0.02$) in contrast to the self-referential condition ($M = .13$, $SE = 0.02$; $p < .001$, 95% CI_{diff} [0.07, 0.14]). Moreover, negative words encoded self-referentially ($M = .17$, $SE = 0.02$) received a greater proportion of misses during recognition when compared to both neutral ($M = .10$, $SE = 0.02$; $p < .001$, 95% CI_{diff} [0.03, 0.10]) and positive words ($M = .11$, $SE = 0.02$; $p = .002$, 95% CI_{diff} [0.02, 0.10]).

Correct rejections and false alarms

The repeated-measures ANOVA with the valence factor was statistically significant, $F(2, 42) = 13.98$, $p < .001$, $\eta_p^2 = .40$. A higher proportion of correct rejections was observed for neutral new words ($M = .88$, $SE = 0.02$) relative to both new negative ($M = .81$, $SE = 0.02$; $p < .001$, 95% CI_{diff} [0.04, 0.10]) and positive words ($M = .82$, $SE = 0.03$; $p = .004$, 95% CI_{diff} [0.02, 0.11]). The same effect was observed in the case of false alarms, $F(2, 42) = 14.08$, $p < .001$, $\eta_p^2 = .40$, as new neutral words received lower false alarm rates ($M = .12$, $SE = 0.02$) relative to both new negative ($M = .19$, $SE = 0.02$; $p < .001$, 95% CI_{diff} [0.04, 0.10]) and positive words ($M = .19$, $SE = 0.03$; $p = .004$, 95% CI_{diff} [0.02, 0.11]).

Proportion of "yes" responses in the word encoding phase

The 2 x 3 repeated-measures ANOVA showed a main effect of self-reference, $F(1, 21) = 49.39$, $p < .001$, $\eta_p^2 = .70$, a main effect of valence, $F(2, 42) = 116.89$, $p < .001$, $\eta_p^2 = .85$, and an interaction effect, $F(2, 42) = 57.04$, $p < .001$, $\eta_p^2 = .73$, $\varepsilon = .70$. Specifically, for negative ($p < .001$, 95% CI_{diff} [0.41, 0.67]) and neutral words ($p < .001$, 95% CI_{diff} [0.21, 0.41]), the proportion of "yes" responses was higher in the common condition than in the self-referential condition, with no statistically significant differences with respect to positive words ($p = .197$). Additionally, in the self-referential condition, there were statistically significant differences between all valence categories (all $p < .001$), with positive words receiving the highest proportion of "yes" responses, followed by neutral words and then by negative words: positive ($M = .76$, $SE = 0.03$) > neutral ($M = .43$, $SE = 0.03$) > negative ($M = .18$, $SE = 0.04$). In the case of the common condition, positive words ($M = .81$, $SE = 0.03$) also received a greater proportion of "yes" responses relative to both negative ($M = .72$, $SE = 0.04$; $p = .003$, 95% CI_{diff} [0.03, 0.15]) and neutral words ($M = .73$, $SE = 0.04$; $p = .018$, 95% CI_{diff} [0.11, 0.14]); however, no significant differences were observed between the last two valence categories ($p > .05$).

Judgments of source (JOSs)

Concerning the mean JOSs ratings, a main effect of self-reference, a main effect of valence, and an interaction effect were observed [self-reference: $F(1, 21) = 10.04$, $p = .005$, $\eta_p^2 = .32$; valence: $F(2, 42) = 14.29$, $p < .001$, $\eta_p^2 = .41$; source x valence: $F(2, 42) = 5.12$, $p = .010$, $\eta_p^2 = .20$]. In the case of both negative and positive words, participants deemed the source of self-referential words as more memorable than the source of words encoded in the common condition (negative: $p = .008$, 95% CI_{diff} [0.10, 0.59]; positive: $p = .001$, 95% CI_{diff} [0.24, 0.79]; see Table 1); no difference between encoding conditions was found in the case of neutral items ($p = .059$, 95% CI_{diff} [- 0.01, 0.53]). Moreover, positive words studied in the self-referential condition received higher JOSs ratings ($M = 4.34$, $SE = 0.16$) in contrast to both neutral ($M = 3.98$, $SE = 0.16$; $p = .001$, 95% CI_{diff} [0.14, 0.57]) and negative words ($M = 3.95$, $SE = 0.16$; $p = .001$, 95% CI_{diff} [0.16, 0.61]; see Figure S1B). In the case of the common judgment task, there was only a statistically significant difference between positive and negative words, wherein the source of positive words ($M = 3.82$, $SE = 0.17$) was regarded as more memorable than the source of negative words ($M = 3.61$, $SE = 0.16$; $p = .024$, 95% CI_{diff} [0.02, 0.41]; see Figure 1B).

The analysis of metamnemonic accuracy as measured by the gamma correlation suggested that the coefficients were statistically different from zero in specific experimental conditions, namely for negative and neutral self-referential conditions, $t(21) > 3.10$, $p \leq .005$. These results might indicate

some degree of association between judgments of source and source memory recognition scores, and although the coefficients are small and close to zero (see Table 1), the mean gamma coefficients reported here approach the ones already reported in prior studies (e.g., Carroll et al., 1999; Skavhaug et al., 2010). Even though, a 2 x 3 RMANOVA showed no statistically significant main effects, nor an interaction effect [self-reference: $F(1, 21) = 3.74$, $p = .067$, $\eta_p^2 = .15$; valence: $F(2, 42) = 0.13$, $p = .827$, $\eta_p^2 = .01$, $\varepsilon = .79$; self-reference x valence: $F(2, 42) = 0.93$, $p = .404$, $\eta_p^2 = .04$]. Taken together, while there is some evidence supporting that, in given conditions, the metamnemonic predictions of the participants may have some value, there is no compelling evidence in favor of a strong relationship between subjective memory predictions and source memory performance, nor in favor of a difference as a function of source and valence conditions.

With respect to the electrophysiological data (using a 2 (self *vs*. common) x 3 (negative *vs*. neutral *vs*. positive) x 5 (ROI: frontal *vs*. fronto-central *vs*. central *vs*. centro-parietal *vs*. parietal) x 3 (laterality: left *vs*. middle *vs*. right) RMANOVAs), most of the analyses performed on the four different windows did not yield interaction effects with the experimental conditions (see Supplementary Table S6). Notwithstanding, considering the 300 – 500 ms time window, a self-reference x valence x ROI interaction effect was found, $F(8, 168) = 2.83$, $p = .046$, $\eta_p^2 = .12$, $\varepsilon = .37$. Specifically, follow-up post-hoc tests revealed that at parietal sites the mean amplitude for JOSs associated with positive words studied with the self-referential judgement ($M = 3.89$, $SE = 0.71$) was more positive than the mean amplitude for positive words studied with the common judgment ($M = 2.82$, $SE = 0.65$; $p = .044$, 95% CI_{diff} [0.03, 2.11]; see Figure S2). Moreover, irrespective of the experimental condition, there were statistically significant differences between all ROIs (all $p < .05$), with the mean amplitude being the most positive at parietal sites and the most negative at frontal sites. The only two exceptions to this pattern occurred for neutral words encoded self-referentially and positive words encoded in the common condition for which no significant differences were found between frontal and fronto-central locations ($p > .05$).

References

Bastin, C., Feyers, D., Souchay, C., Guillaume, B., Pepin, J. L., Lemaire, C., ... Salmon, E. (2012). Frontal and posterior cingulate metabolic impairment in the behavioral variant of frontotemporal dementia with impaired autonoetic consciousness. *Human Brain Mapping*, *33*(6), 1268–1278. https://doi.org/10.1002/hbm.21282

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*(4), 548-564. https://doi.org/10.1037/0033-295X.97.4.548

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Castel, A. D., Rhodes, M. G., & Friedman, M. C. (2013). Predicting memory benefits in the production effect: The use and misuse of self-generated distinctive cues when making judgments of learning. *Memory & Cognition*, *41*(1), 28–35. https://doi.org/10.3758/s13421-012-0249-6

Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1121-1136. https://doi.org/10.1037//0278-7393.24.5.1121

Dodson, C. S., Prinzmetal, W., & Shimamura, A. P. (1998). Using excel to estimate parameters from observed data: An example from source memory data. *Behavior Research Methods, Instruments, & Computers, 30*(3), 517-526. https://doi.org/10.3758/BF03200685

Grainger, C., Williams, D. M., & Lind, S. E. (2016). Judgment of learning accuracy in high-functioning adolescents and adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(11), 3570–3582. https://doi.org/10.1007/s10803-016-2895-1

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/13825585.2014.957150.

Müller, B. C. N., Tsalas, N. R. H., van Schie, H. T., Meinhardt, J., Proust, J., Sodian, B., & Paulus, M. (2016). Neural correlates of judgments of learning – An ERP study on metacognition. *Brain Research*, *1652*, 170–177. https://doi.org/10.1016/j.brainres.2016.10.005

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy in internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology, 10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Soares, A. P., Comesaña, M., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., ...Machado, J. (2010). P-PAL: Uma base lexical com índices psicolinguísticos do Português Europeu [P-PAL: A European Portuguese lexical database]. *Linguamática, 2*(3), 67-72.

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Soares, A. P., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., França, P., ...Comesaña, M. (2014). Procura-PALavras (P-PAL): A new measure of word frequency for contemporary European Portuguese. *Psicologia: Reflexão e Crítica, 27*(1), 110-123. https://doi.org/10.1590/S0102-79722014000100013

Soares, A. P., Iriarte, A., Almeida, J. J., Simões, A., Costa, A., Machado, J., ...Perea, M. (2018). Procura-PALavras (P-PAL): A web-based interface for a new European Portuguese lexical database. *Behavior Research Methods, 50*(4), 1461-1481. https://doi.org/10.3758/s13428-018-1058-z

APPENDIX E

SUPPLEMENTARY MATERIALS - CHAPTER VI

In this section, the preliminary behavioral findings that were obtained in the context of the functional magnetic resonance imaging (fMRI) study are briefly presented.

Accordingly, eighteen Portuguese-speaking participants (9 males and 9 females) recruited from the University of Minho and community took part in the fMRI study. They were aged between 20 and 33 years (M = 26.89, SD = 3.83) and had at least 12 years of formal education (M = 16.47, SD = 2.53). None of the participants reported past or current history of psychiatric or neurological disorders, and they were not taking psychotropic medication. Also, all participants were right-handed and had normal or corrected-to-normal vision.

The pool of stimuli and experimental procedures were generally similar to those described in the context of the event-related potential (ERP) study (see Chapter V). However, some differences are worth noting: whereas five study-test cycles were used in both studies, the number of stimuli in each study and test phases differed. Specifically, in the fMRI study, each study phase was composed of 36 items whereas each test phase was composed of 18 items. Additionally, each study phase was divided in two mini-blocks, one dedicated to self-referential judgments and other dedicated to common judgments, thereby the encoding tasks were constant across the same mini-block. This division in mini-blocks was based on previous fMRI studies with the purpose of reducing the influence of task switching conditions (e.g., Leshikar & Duarte, 2012, 2014). Judgments of source (JOSs) ratings were not collected. The inter-stimulus intervals were jittered between 2 and 8 s during the study phase (M = 3.52 s) and between 2 and 6 s during the test phase (M = 2.34 s). Although the encoding time of each stimulus was 3 s in both ERP and fMRI studies, during the test phase the participants saw the stimulus for 2 s and then responded without time restrictions (i.e., self-paced) in the ERP study, whereas in the case of the fMRI study they had 4 s to see the stimulus and to provide a response. If a keypress was not registered within this temporal window, the subsequent trial would be automatically presented. Also, during the test phase, participants had three options to choose from: “self-description”, “common” or “new”. Thus, due to technical reasons, no “do not know” option was available. A schematic illustration of an experimental trial in both study and test phases of the fMRI study is depicted in Figure S1.

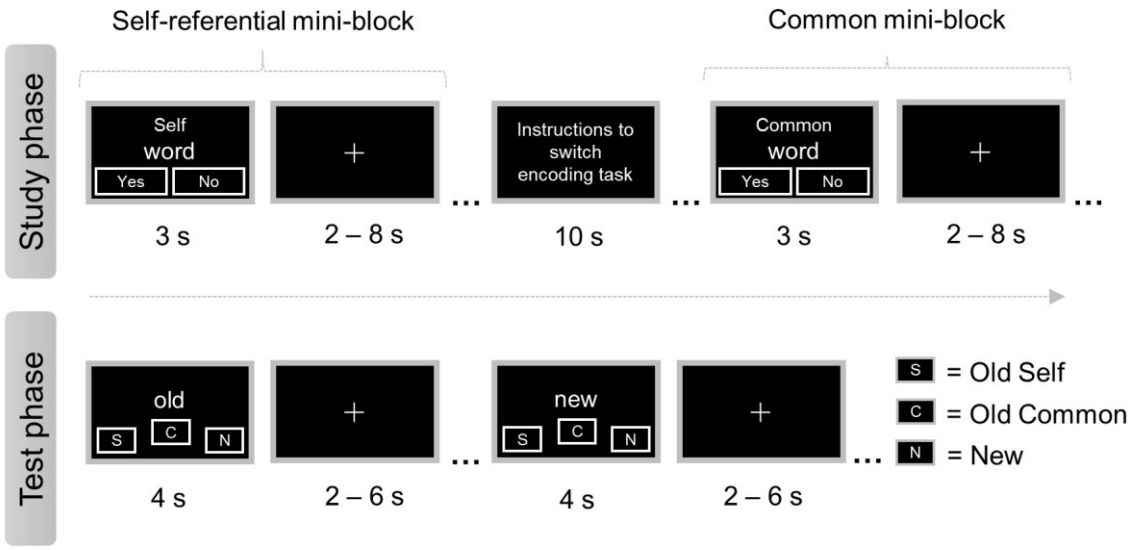


Figure S1. Schematic illustration of an experimental trial from the study and test phases of the fMRI study.

The data preparation and analysis for item and source memory performance followed the same steps already described in Chapters III to V. Table S1 presents a summary of the main descriptive statistics regarding the behavioral performance in the fMRI study. Specifically, regarding source memory performance, the 3 (valence: negative *vs*. neutral *vs*. positive) x 2 (source: self *vs*. common) repeated-measures analysis of variance showed a main effect of source, $F(1, 17) = 14.61$, $p = .001$, $\eta_p^2 = .46$, a main effect of valence, $F(2, 34) = 20.44$, $p < .001$, $\eta_p^2 = .55$, and an interaction effect, $F(2, 34) = 13.11$, $p < .001$, $\eta_p^2 = .44$. Bonferroni-corrected post-hoc tests revealed that a source memory advantage of self-referential ($M = 0.67$, $SE = 0.03$) in comparison with common conditions ($M = 0.25$, $SE = 0.05$) was observed in the case of positive words ($p < .001$, 95% CI_{diff} [0.33, 0.55]), but not in the case of the remaining valence categories. More importantly, neutral ($M = 0.58$, $SE = 0.05$; $p = .001$, 95% CI_{diff} [0.11, 0.45]) and positive words ($M = 0.67$, $SE = 0.03$; $p < .001$, 95% CI_{diff} [0.19, 0.55]) studied in the self-referential condition were associated with greater source memory performance than negative words in the same condition ($M = 0.31$, $SE = 0.07$). In the case of the common condition, the only difference emerged between neutral and positive words ($p = .001$, 95% CI_{diff} [0.11, 0.39]), with the first ($M = 0.50$, $SE = 0.05$) being associated with better source memory performance than the second ($M = 0.25$, $SE = 0.05$; see Table S1 and Figure S2B).

Table S1

Descriptive Statistics of the Behavioral Performance for each Experimental Condition (Source x Valence)

	Self – *M* (*SD*)			Common – *M* (*SD*)		
	Negative	Neutral	Positive	Negative	Neutral	Positive
Correct source	0.51 (0.23)	0.73 (0.13)	0.77 (0.10)	0.53 (0.15)	0.60 (0.18)	0.49 (0.16)
Incorrect source	0.20 (0.11)	0.14 (0.09)	0.09 (0.06)	0.19 (0.10)	0.10 (0.07)	0.24 (0.11)
Source miss	0.17 (0.12)	0.12 (0.09)	0.11 (0.10)	0.26 (0.16)	0.26 (0.16)	0.25 (0.15)
Item Pr measure	0.60 (0.29)	0.78 (0.12)	0.69 (0.15)	0.60 (0.20)	0.61 (0.18)	0.56 (0.18)
Item Br measure	0.31 (0.15)	0.42 (0.25)	0.59 (0.26)	0.29 (0.16)	0.27 (0.17)	0.41 (0.19)
SM measure	0.31 (0.28)	0.58 (0.21)	0.67 (0.12)	0.34 (0.20)	0.50 (0.21)	0.25 (0.22)
RT – study phase	1519 (191)	1663 (207)	1499 (213)	1622 (234)	1601 (278)	1627 (346)
RT – correct source	2061 (266)	1851 (280)	1883 (324)	2157 (279)	2136 (386)	2164 (271)

	New negative	New neutral	New positive
CR	0.86 (0.09)	0.89 (0.08)	0.81 (0.11)
FA	0.11 (0.08)	0.09 (0.08)	0.17 (0.10)
RT - CR	1685 (300)	1689 (316)	1733 (284)

Note. CR = Correct Rejections; FA = False Alarms; RT = Response Time; SM = Source Memory.

The analysis of item memory recognition demonstrated a main effect of source, $F(1, 17) = 11.70$, $p = .003$, $\eta_p^2 = .41$, a main effect of valence, $F(2, 34) = 4.88$, $p = .014$, $\eta_p^2 = .22$, and also an interaction effect, $F(2, 34) = 4.35$, $p = .045$, $\eta_p^2 = .20$, $\varepsilon = .59$. Further analysis of the interaction revealed that the benefit of self-referential (*vs*. common) conditions was verified for neutral ($p = .001$, 95% CI_{diff} [0.08, 0.24]) and positive words ($p < .001$, 95% CI_{diff} [0.07, 0.20]), but not for negative words (see Table S1 and Figure S2A). Also, whereas there were no differences between valence categories in the common condition, neutral words ($M = 0.78$, $SE = 0.03$) were better recognized than emotional words in the self-referential condition (negative: $M = 0.60$, $SE = 0.07$; $p = .040$, 95% CI_{diff} [0.01, 0.35]; positive: $M = 0.69$, $SE = 0.04$; $p = .006$, 95% CI_{diff} [0.02, 0.15]). Considering the measure of response bias, both main effects (source: $F(1, 17) = 21.57$, $p < .001$, $\eta_p^2 = .56$; valence: $F(2, 34) = 17.08$, $p < .001$, $\eta_p^2 = .50$) and the interaction effect were statistically significant ($F(2, 34) = 3.93$, $p = .043$, $\eta_p^2 = .19$, $\varepsilon = .76$). As response bias values above 0.50 are indicative of a more liberal criterion of response while values below 0.50 are indicative of a more conservative criterion of response (Snodgrass & Corwin, 1988), results showed that participants were more conservative in their response (i.e. tendency to respond "new" when unsure) for words encoded in the common (*vs*. self-referential) condition, in the specific case of neutral ($p = .002$, 95% CI_{diff} [0.06, 0.25]) and positive words ($p = .001$, 95% CI_{diff} [0.08, 0.27]; see Table S1). Also, irrespective of the encoding task, negative (self-reference: $p = .001$, 95% CI_{diff} [0.12, 0.44]; common: $p = .009$, 95% CI_{diff} [0.03, 0.21]) and neutral words (self-reference: $p = .015$,

95% CI_{diff} [0.03, 0.30]; common: $p < .001$, 95% CI_{diff} [0.07, 0.22]) were associated with a more conservative response criterion than positive words (see Table S1).

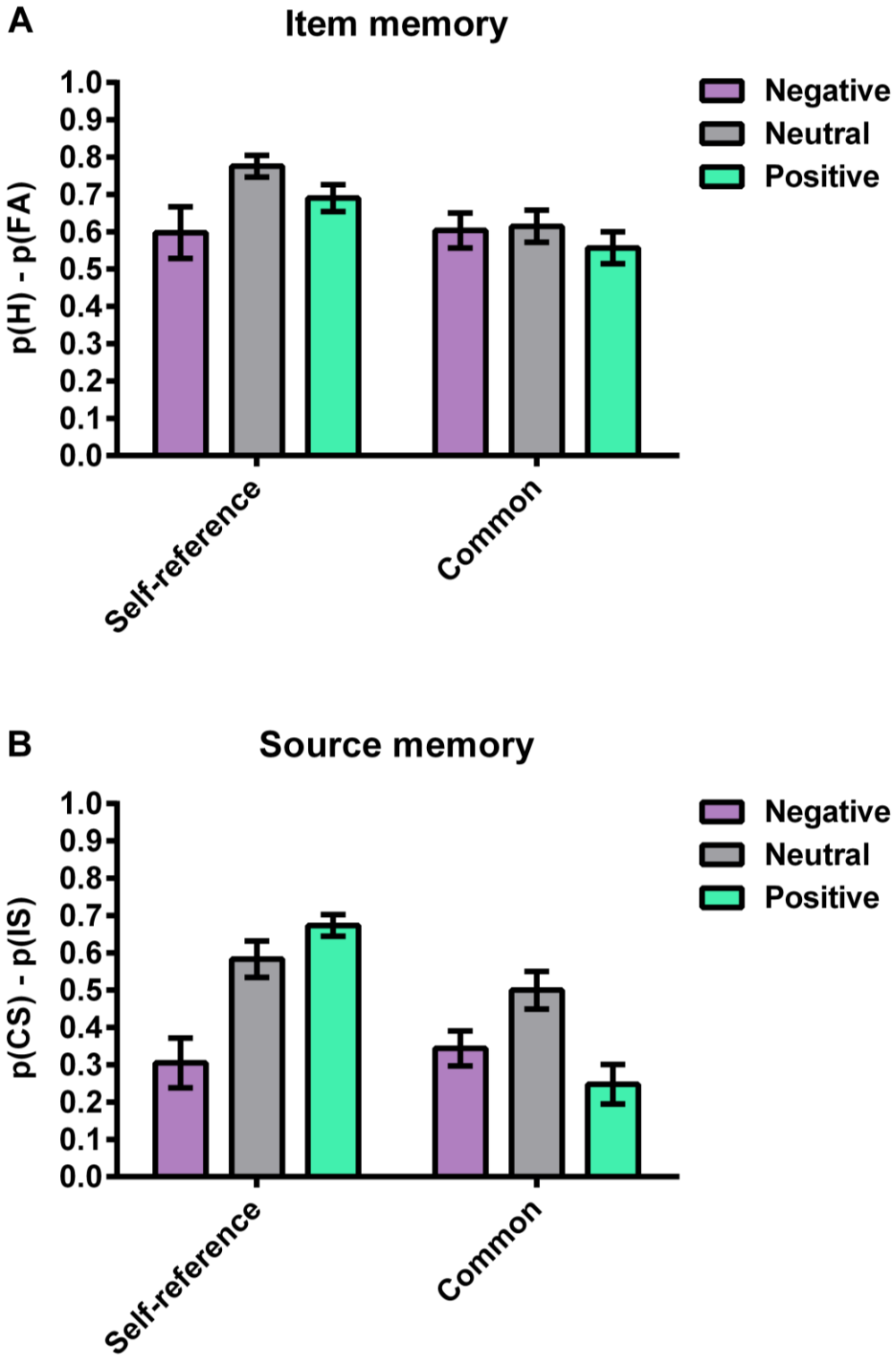


Figure S2. Behavioral findings for item and source memory (panel A and B, respectively). The performance is plotted on the y-axis as a function of emotion (negative/neutral/positive) and source (self-reference/common). Note: the error bars correspond to the standard errors of the means; CS = Correct Source; FA = False Alarms; H = Hits; IS = Incorrect Source.

References

Leshikar, E. D., & Duarte, A. (2012). Medial prefrontal cortex supports memory accuracy for self-referenced items. *Social Neuroscience*, *7*(2), 126–145. https://doi.org/10.1080/17470919.2011.585242

Leshikar, E. D., & Duarte, A. (2014). Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(1), 236–252. https://doi.org/10.3758/s13415-013-0198-y

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

APPENDIX F

APPROVAL BY THE SUBCOMMISSION OF LIFE AND HEALTH SCIENCES - UMINHO ETHICS COMMISSION

Universidade do Minho

SECVS

Subcomissão de Ética para as Ciências da Vida e da Saúde

Identificação do documento: SECVS 105/2016

Título do projeto: *Exploring the modulatory role of emotion in source memory attributions - Behavioral and electrophysiological evidence*

Investigador(a) responsável: Professora Doutora Adriana Sampaio, Centro de Investigação em Psicologia (CIPsi), Escola de Psicologia, Universidade do Minho; Professora Doutora Ana P. Pinheiro, Centro de Investigação em Psicologia (CIPsi), Escola de Psicologia, Universidade do Minho e Faculdade de Psicologia, Universidade de Lisboa; e a aluna Diana Cristina Rodrigues Pereira, Centro de Investigação em Psicologia (CIPsi), Escola de Psicologia, Universidade do Minho

Subunidade orgânica: Centro de Investigação em Psicologia (CIPsi), Escola de Psicologia, Universidade do Minho

PARECER

A Subcomissão de Ética para as Ciências da Vida e da Saúde (SECVS) analisou o processo relativo ao projeto intitulado *"Exploring the modulatory role of emotion in source memory attributions - Behavioral and electrophysiological evidence".*

Os documentos apresentados revelam que o projeto obedece aos requisitos exigidos para as boas práticas na experimentação com humanos, em conformidade com o Guião para submissão de processos a apreciar pela Subcomissão de Ética para as Ciências da Vida e da Saúde.

Face ao exposto, a SECVS nada tem a opor à realização do projeto.

Braga, 29 de setembro de 2016.

A Presidente

MARIA CECÍLIA DE LEMOS PINTO ESTRELA LEÃO

Digitally signed by MARIA CECÍLIA DE LEMOS PINTO ESTRELA LEÃO
DN: c=PT, o=Cartão de Cidadão, ou=Cidadão Português, ou=Assinatura Qualificada do Cidadão, sn=DE LEMOS PINTO ESTRELA LEÃO, givenName=MARIA CECÍLIA, serialNumber=BI014512203, cn=MARIA CECÍLIA DE LEMOS PINTO ESTRELA LEÃO
Date: 2016.09.30 14:34:47 +01'00'

Maria Cecília de Lemos Pinto Estrela Leão