UMinho | 2021

Beatriz Mello

A Race to Remember: The Effect of Prospective Memory Task Difficulty in Incidental Learning

**Universidade do Minho**
Escola de Psicologia

Beatriz Silva Mello

# A Race to Remember: The Effect of Prospective Memory Task Difficulty in Incidental Learning

junho de 2021

Universidade do Minho
Escola de Psicologia

Beatriz Silva Mello

# A Race to Remember: The Effect of Prospective Memory Task Difficulty in Incidental Learning

Dissertação de Mestrado
Mestrado Integrado em Psicologia

Trabalho efetuado sob a orientação do
**Professor Doutor Pedro B. Albuquerque (Universidade do Minho)**
e da
**Professora Doutora Angela Bartolo (Université de Lille)**

junho de 2021

## 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-NãoComercial-SemDerivações
CC BY-NC-ND

https://creativecommons.org/licenses/by-nc-nd/4.0/

## Acknowledgements

Throughout the writing of this dissertation, I have received a great deal of support and assistance and for that, I would first like to thank my advisors Professor Pedro Albuquerque and Professor Angela Bartolo, whose knowledge and experience helped me build a strong body of work that led me to exceed myself. I would also like to thank Patrícia Matos for all the feedback and suggestions made in every step of my dissertation.

Secondly, I would like to thank all participants that took part on this study and the Universities of Minho and Lille for allowing me to develop this work.

I must also thank the Grupo de Investigação em Memória Humana da Universidade do Minho (GIMHUM) for all the support, knowledge and for being an open debate group with a joint interest in contributing for the understanding of memory.

In addition, I would like to thank my family and friends specially, my mom and Cristiana, for the love, support and joined effort in keeping me sane for the past 5 years. I would also like to thank Anaïs and Alex for the good welcoming in the University of Lille and all the fruitful discussions on the most diverse themes in the field of psychology.

Lastly, I would like to thank everyone that I came across in the past 5 years because they made me who I am.

## 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.

Beatriz Silva Mello

(Beatriz Silva Mello)

Uma Corrida a Lembrar: o Efeito da Dificuldade da Tarefa de Memória Prospetiva na Aprendizagem Incidental


## Resumo

A memória prospetiva, capacidade de recordar a intenção de desempenhar uma ação no futuro sem um lembrete permanente, tem sido estudada ao longo dos anos e faz parte do quotidiano. Mas qual o seu efeito na aprendizagem implícita? O presente estudo teve como objetivo responder a esta questão.

Foi manipulada a dificuldade da tarefa de memória prospetiva e, para além de um grupo de controlo que apenas realizou uma tarefa de decisão lexical, o estudo incluiu dois grupos experimentais que adicionalmente realizaram uma tarefa de memória prospetiva com pistas não focais: um realizou uma tarefa de memória prospetiva fácil e outro uma difícil, totalizando 159 participantes. Foi depois realizada uma tarefa de reconhecimento das palavras apresentadas na tarefa de decisão lexical.

O desempenho de uma tarefa de memória prospetiva fácil não teve impacto no desempenho nem da tarefa de decisão lexical nem da tarefa de reconhecimento, mas um aumento da sua dificuldade levou a um pior desempenho na tarefa de decisão lexical, sem impacto no desempenho da tarefa de reconhecimento. Estes resultados suportam estudos prévios e servem de apoio à *delay theory,* uma teoria que explica os mecanismos através dos quais as ações prospetivas são recuperadas para serem desempenhadas.



*Palavras-chave: aprendizagem incidental, delay theory, dificuldade da tarefa de memória prospetiva, memória prospetiva, pistas não focais*

A Race to Remember: The Effect of Prospective Memory Task Difficulty in Incidental Learning


## Abstract

Prospective memory, the ability to remember the intention of performing an action in the future without a permanent reminder, has been studied over the years and is a part of everyday life. But what is its effect on incidental learning? The present study aimed to answer this question.

The difficulty of the prospective memory task was manipulated, and besides a control group that only performed a lexical decision task, it included two experimental groups that additionally performed a prospective memory task with nonfocal cues: one performed an easy prospective memory task and the other a difficult one, in a total of 159 participants. Then, a recognition task of the words presented in the lexical decision task was performed.

The performance of an easy prospective memory task did not impact either the lexical decision task nor the recognition task performances, but an increase of its difficulty led to a worst performance in the lexical decision task, with no impact on the recognition task performance. These results support previous studies and present themselves as one more argument of support to the delay theory, a theory that explains the mechanisms through which prospective actions are retrieved to be performed.



*Keywords: delay theory, incidental learning, nonfocal cues, prospective memory, prospective memory task difficulty,*

## Contents



## List of Illustrations



## List of Tables

## A Race to Remember: The Effect of Prospective Memory Task Difficulty in Incidental Learning

Prospective memory (PM) refers to the individual's ability to remember to perform previously formed action plans (intentions) at the appropriate moment in the future (McDaniel & Einstein, 2000; Rummel & McDaniel, 2019; Scullin et al., 2012). Prospective memory tasks are a daily event in everyone's life, and maybe that is why researchers have been studying PM over the past 25 years (Reese-Melancon et al., 2019). There are several examples of such daily intentions, like remembering to take medicines or return a friends' call. The consequences of someone forgetting to complete a PM task (i.e., performing an omission error) may vary. For instance, forgetting to return a friend's call may not have severe consequences besides the need to apologize, whereas forgetting to take medicines can result in a medical emergency.

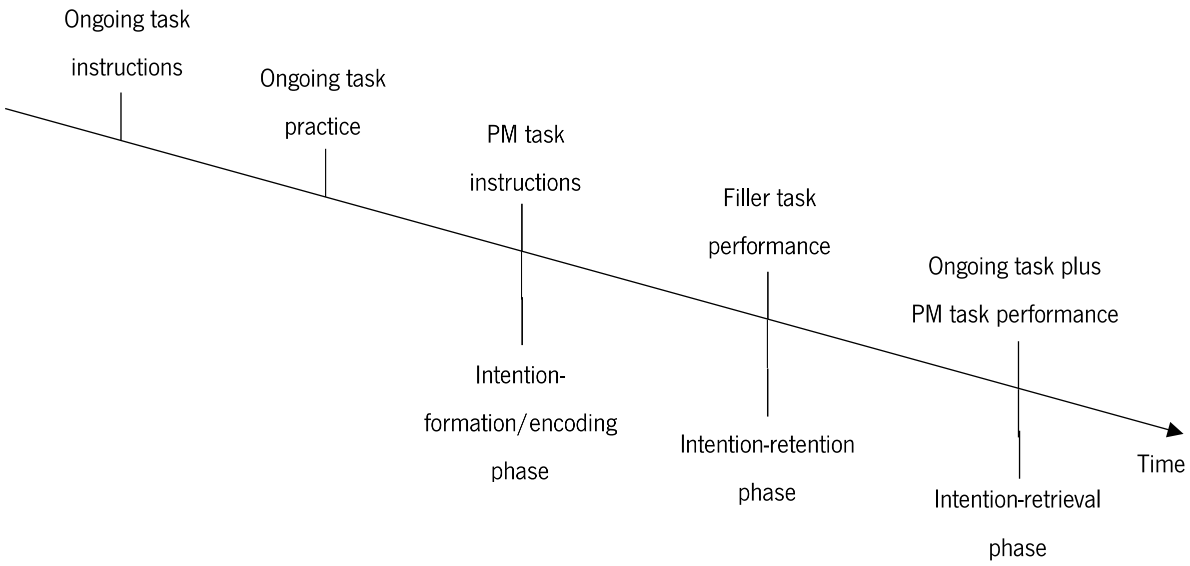
Notably, PM tasks are different from retrospective memory tasks: While in retrospective memory tasks, there is an explicit instruction from the experimenter for participants to engage in a memory search, in PM tasks, there is not (Humphreys et al., 2020). Another important distinction to be made concerns the nature of PM tasks that can be either time-based or event-based initiated: In event-based PM tasks, the action is performed when an event occurs, while in time-based PM tasks, there is a need to remember to act in a given time (Smith, 2003).

Studies in PM tend to use a standard paradigm (Figure 1) to investigate PM-related phenomena. First, participants are instructed to perform an ongoing task (e.g., a lexical decision task, LDT), and the opportunity to practice is given. Then, the PM task is encoded (i.e., intention-formation phase), and a delay is introduced (i.e., intention-retention phase) to prevent participants from rehearsing the intention. Next, participants must perform the PM task embedded in the ongoing task without being reminded of it (i.e., intention-retrieval phase; Einstein & McDaniel, 2005; Rummel & McDaniel, 2019). Participants must perform the intended PM action (e.g., press "V") upon encountering PM targets (i.e., stimuli that serve as a cue to perform the PM targeted action and which are mixed with nontargeted trials of the ongoing task) while performing the ongoing task (Einstein & McDaniel, 2005). This paradigm allows researchers to collect data about, for example, the accuracy in the PM task but also the accuracy and response times (RTs) in the ongoing task (e.g., Einstein et al., 2005; Loft & Humphreys, 2012).

Research on PM aims to understand how prospective memory works and unveiling why sometimes it fails and, by doing so, focuses on studying the cognitive mechanisms that underly PM performance. Theories in the field of PM mainly concern the mechanisms through which the intentions to be performed are retrieved. The monitoring theory (Smith, 2003) states that while completing a PM task and an ongoing task, one is frequently rehearsing intentions, and the environment is continuously being monitored, searching for a cue to perform the PM's targeted action. This permanent monitoring would lead to diminished attention and working memory resources allocated to the ongoing task, resulting in costs (most frequently conceptualized as worst accuracy or larger RTs in the ongoing task performance; e.g., Anderson et al., 2019; Einstein et al., 2005; Reese-Melancon et al., 2019; Smith, 2003). In contrast, the spontaneous retrieval theory (Einstein & McDaniel, 1996) posits that when a PM cue is encountered, the action is brought to mind by a relatively automatic process. That is, the action to be completed "pops into mind" in the presence of the PM cue, thus costs to the ongoing task are not predicted.

**Figure 1.**

*Standard paradigm in PM studies.*




Einstein et al. (2005) proposed another theory that encompasses the previous two, the multiprocess theory, which argues that people rely on different processes under different conditions (Einstein et al., 2005) related to, for example, the nature of the cues (McDaniel & Einstein, 2000). In more detail, PM tasks, and consequently PM cues, have been conceptualized as focal or nonfocal to the ongoing task. Prospective memory cue focality concerns the degree of processing overlap between the ongoing task and the PM task (McDaniel & Einstein, 2000). Focal cues are target events whose processing overlaps with the information required to perform the ongoing task, such as responding to the word *handball* during an LDT. Nonfocal cues are target events in which processing does not overlap with the processing required to perform the ongoing task (Anderson et al., 2019). For instance, responding whether a stimulus, like *handball*, has two syllables or not during an LDT. In a recent meta-analysis conducted by Anderson et al. (2019), researchers concluded that when focal cues are used, there are little to no costs to the ongoing task (as predicted by the spontaneous retrieval theory) but, when nonfocal cues are used, costs on the ongoing task are found (as predicted by the monitoring theory; see Anderson et al., 2019). This phenomenon is called the "focality effect" (Loft & Remington, 2013).

However, Loft and Remington (2013) proposed a more recent theory, inspired by mathematical models of cognitive choice that assume that task decisions are made by accumulating evidence (Heathcote et al., 2015). The delay theory argues that costs (i.e., larger RTs) are not due to monitoring, like Smith (2003) has argued, but due to the building up of information (i.e., evidence) relevant to the ongoing and PM tasks in parallel but at different rates (Anderson et al., 2017). This process continues until a threshold is reached, either from the ongoing task or the PM task, and a response is then given. Put differently, although recognizing that attentional resources are limited, the delay theory argues that costs are not due to divided attention between the PM and the ongoing task but due to the building up of information to complete the task as if the ongoing task response and the PM task response were competing in a race for response selection (Heathcote et al., 2015).

Therefore, according to the delay theory, increased RTs in the presence of a nonfocal PM task arise because the ongoing task response thresholds are strategically set higher to have more time for the PM task response selection (Heathcote et al., 2015; Boywitt & Rummel, 2012). In line with this idea, it was expected that by increasing the time available for participants to process stimuli, PM errors with nonfocal cues should be reduced given that participants – usually instructed to respond as quickly as possible -, can usually tend to respond to the more routine and often faster task (i.e., to the ongoing task; Heathcote et al., 2015). This was what Loft and Remington (2013) found: By delaying task responses, the effect of target focality on PM accuracy was eliminated. This theory is also supported by statistical evidence derived from evidence accumulating models, which assume that what participants do during the ongoing and PM task performances is accumulating evidence regarding each specific stimulus to decide which response to give. Once the amount of evidence accumulated reaches a threshold, the response is given (see Strickland et al., 2019).

More support for this view arises from Loft and Humphreys (2012) study that aimed to understand how participants process the nontargeted items of the ongoing task. More specifically, participants performed an LDT as an ongoing task, a nonfocal PM task (in which participants performed the targeted action whenever they saw an exemplar of a specific category) or a focal PM task (in which participants should perform the targeted action whenever they saw a specific word), and an unexpected recognition task (in which the nontargeted words presented during the LDT were intermixed with new words). Concerning the recognition task data, the authors found higher discrimination in the nonfocal condition when compared to the control group. This result was interpreted as evidence that “participants in the nonfocal condition mapped the semantic features of letter strings onto the semantic features of their PM category” (Loft & Humphreys, 2012, p. 1146). This process would thus explain the results obtained in the recognition task: The mapping of the stimulus' semantic features presented “increased incidental learning of nontargets and enhanced their subsequent recognition” (Loft & Humphreys, 2012, p. 1146).

Many variables can affect a memory test performance, being attention one of them: According to Baddeley et al. (1984), divided attention leads to worse memory performance when compared to encoding information under full attention, and the more demanding the ongoing task processing is, the worse memory performance will be. Thus, in line with the monitoring theory, participants with a PM task to perform during an ongoing task compared to a control group that only performs an ongoing task should perform worst in a memory test since, in that case, both tasks are performed under divided attention. On the other hand, if attentional resources are not divided, like the delay theory argues, the encoding of information should not be by any means affected, leading to no differences in a recognition task performance between a group that performs an ongoing task plus a PM task and a group that only performs an ongoing task. There is yet another possibility: Increasing PM task’s difficulty may improve memory performance by creating a desirable difficulty during the ongoing plus PM tasks performance, enhancing a posterior recognition task performance (Bjork & Bjork, 2011). Using a recognition task, in which old words (i.e., stimuli that have already been presented in the LDT) and new ones are presented to participants, it will be possible to understand the effect of performing a PM task on incidental learning (i.e., any unplanned or unintended learning that develops while engaging in a task or activity; Kelly, 2012) of the nontargeted words of the LDT (i.e., the ongoing task).

The present study aims to understand the impact of having to perform a future intention in incidental learning, more specifically, in a recognition task and thus, helping to unveil the processes that allow the retrieval of an action to be performed. As mentioned before, the literature indicates that there are costs in the ongoing task when using nonfocal cues, but which cognitive process does that reflect? And does the PM task difficulty influence incidental learning? Applying a paradigm similar to that used by Loft and Humphreys (2012), it will be possible to replicate their results and answer the questions above. Besides, the PM task's difficulty was manipulated (by varying the number of different cues to perform the PM targeted action, pressing “V”) to test if the effect of a PM task in incidental learning of the nontargeted items of the ongoing task is independent of its difficulty.

According to the delay theory, since all stimuli are fully attended (i.e., there is no division of attentional resources between the PM task and the LDT), the performance on the recognition task should not be negatively affected by a nonfocal PM task. Adding this to the work of Loft & Humphreys (2012), which showed an enhancement effect on the recognition task performance as an effect of the presence of a nonfocal PM task, we hypothesize that the performance of a PM task will lead to an increase in performance in the surprise recognition task.

Considering the finding of Tyler et al. (1979) that words encountered during high effort tasks are better recalled than those encountered during low effort tasks and adding the concept of "desirable difficulty" (Bjork & Bjork, 2011), we hypothesize that the presence of a difficult PM task will enhance more the recognition task performance when compared to the presence of an easy PM task. Alternatively, in line with the monitoring theory, the division of attentional resources between the PM task and the ongoing task would result in a worst recognition task performance of participants in the experimental conditions when compared to those in the control group. In line with this theory, it was also expected that by increasing PM task's difficulty, more attentional resources would be devoted to it, resulting in a worst performance in the recognition task in the presence of a harder PM task when compared to an easier one.

To conclude, it is possible to affirm that this study provides support to one of these two theories (i.e., the delay theory or the monitoring theory), besides the main goals of testing the effect of a PM task on incidental learning of nontargeted stimuli and if this effect is independent of its difficulty.

## Method

### Participants

Using the software G*Power (Faul et al., 2007), based on a one-way ANOVA, a medium effect size of .25, and a power of .80 (see Loft & Humphreys, 2012), it was established a sample of 159 participants that were randomly assigned to one of three conditions: No-PM (n = 53); Easy-PM (n = 53), and Hard-PM condition (n = 53). A total of 235 participants were collected, however data of 76 participants has not been included in the analyses either because those participants did not reproduce correctly the instructions given (e.g., did not correctly identify the syllables that served as PM cues; n = 74) or did not answer to a minimum of 80% of the LDT stimuli (n=2). All participants were undergraduate students, European Portuguese native speakers, with ages between 18 and 42 (*M* = 20.73; *DP* = 3.01), who received credits for their participation. The local Ethical Committee for Research in Social and Human Sciences approved this study (CEICSH 096/2020).

### Design

The independent variable used was the PM task difficulty manipulated between-participants by varying the number of different syllables that served as prospective cues: in the Easy-PM condition, participants had to perform a single-target PM task (where one syllable served as a cue), and in the Hard-PM condition, a multi-target PM task (where three different syllables served as cues). It has been argued that some PM tasks are harder than others, based on performance and RTs in the ongoing task and, as shown by Humphreys et al. (2020) and Hicks et al. (2005), multi-target PM tasks present lower levels of accuracy and produce robust RTs in the ongoing task. A control group was added, the No-PM condition, in which participants only performed the ongoing task.

The main dependent variable was the recognition task performance accuracy. The data obtained on the recognition task was analysed using hit and false alarm rates, and discrimination index (using the parameter *d'*). We also assessed LDT and PM task accuracies and LDT criterion[1]. In both cases, accuracy was calculated by dividing the number of correct answers on the LDT and PM task, respectively, by the number of answers given (in the case of LDT accuracy) and by the number of PM cues (n=6; in the case of PM accuracy).

### Materials

In total, 270 words were selected from the Minho Word Pool (Soares et al., 2017). The chosen words were nouns, with concreteness and imagery above 3.5, a frequency of over 56 per million, an extension ranging from 4 to 8 letters, and 2 to 4 syllables (Table 1).

**Table 1**

*Means (M), standard deviations (SD), values of F and p for concreteness, imagery, frequency, number of letters and syllables for each set of words*

| | Set A | | Set B | | Set C | | | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *F (2, 267)* | *p* |
| Concreteness | 4.84 | .90 | 4.88 | .92 | 4.85 | .89 | .05 | .95 |
| Imagery | 4.91 | .84 | 4.94 | .89 | 4.45 | .83 | .11 | .89 |
| Frequency | 175.33 | 144.44 | 173.81 | 139.48 | 177.13 | 146.51 | .01 | .99 |
| Letters | 5.90 | 1.28 | 5.94 | 1.31 | 6.06 | 1.24 | .35 | .70 |
| Syllables | 2.62 | .61 | 2.57 | .60 | 2.58 | .58 | .22 | .80 |

[1] The procedure was conducted fully online due to the pandemic situation, and thus it was not possible to have data on RTs like previously planned since this measure is very sensitive to characteristics of the computer used and respective internet connection.

The 270 words were equally divided into three sets (sets A, B and C) of 90 words. In each set, from those 90 words, 90 pseudowords were created by changing the order of the syllables (e.g., *cabeça* to *becaça*) or replacing vowels in each word (e.g., *combate* to *cimbote*), thus resulting in 180 stimuli per set. In each set, six words and six pseudowords were nonfocal PM targets, in a total of 18 words and 18 pseudowords for all three sets. More specifically, in each set, there was one word and one pseudoword with the syllable <nal>, one word and one pseudoword with the syllable <pos>, one word and one pseudoword with the syllable <tar> and three words and three pseudowords with the syllable <sa>.

The sets were used as either words or pseudowords in the LDT or as new words in the recognition task. The assignment of the sets to serve either as words, pseudowords or new words was counterbalanced.

**Procedure**

The experiment was programmed using Qualtrics (Qualtrics, 2020) and was conducted online. Participants were told that the procedure took approximately 45 minutes, and were recommended to be in a quiet place, to turn off all phone and computer notifications and to perform the experiment on a computer. After assigning to the study, participants received the link, by e-mail, to complete the procedure online and the informed consent was presented.

The procedure itself had seven main phases (Figure 2): (1) ongoing task instructions, (2) ongoing task practice, (3) PM task instructions, (4) filler task, (5) ongoing task plus PM task performance, (6) filler task, and (7) recognition task.

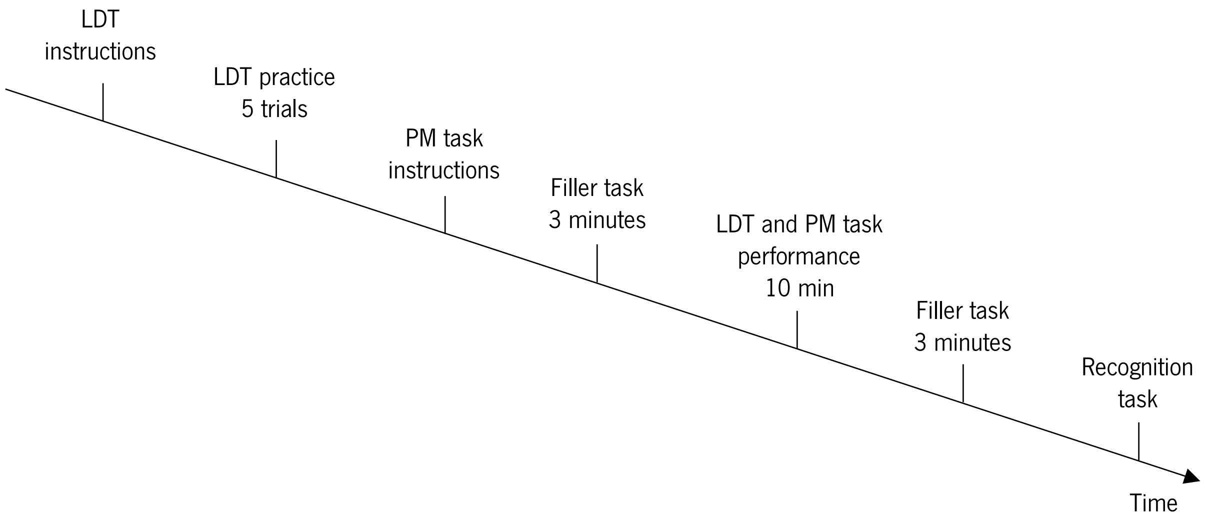
Following the procedure of Loft and Humphreys (2012), all participants read the instructions to perform an LDT (the ongoing task) - they had to decide as quickly and accurately as possible whether a string of letters was a word or a pseudoword and respond by pressing "S" if it was a word or "N" if it was not. After practicing the LDT, except for the No-PM (control) condition, participants received the instructions to perform an additional PM task: In the No-PM condition, there was no PM task; in the Easy-PM condition, participants had to press "V" whenever they saw a stimulus with the syllable <sa> (single-target); and in the Hard-PM condition they had to press "V" whenever they saw a stimulus with the syllable <nal>, or <pos>, or <tar> (multi-target). After these instructions, participants were asked to reproduce (by writing) the instructions of the task they had to perform to ensure that the instructions were understood. Concerning the LDT, an answer was considered correct when the string of letters presented was a word and the participant pressed "S", and when the string of letters showed was a pseudoword, and the participant pressed "N". Concerning the PM task, an answer was considered correct if the participant pressed "V" on the target trial (i.e., in the trial where the PM cue was presented) or in the following trial.

Next, participants performed a semantic fluency task that served as a filler task for 3 minutes, where six categories were presented (e.g., fruit), one at a time, for 30 seconds each, and the instruction was to write, on the text box, words belonging to that specific category. In the LDT that followed, the presentation of the 90 words and the 90 pseudowords was randomized, except that the six cues (3 words and 3 pseudowords) to perform the PM targeted action were presented in a random position between trials 5 and 30, 31 and 60, 61 and 90, 91 and 120, 121 and 150, and 151 and 180.

After another 3-minute delay, during which participants were asked to organize numbers in decreasing order, the recognition task began. The 90 words that were already presented in the LDT were presented in the recognition task plus 90 new words (presented in 3 blocks of 60, 30 old words and 30 new ones, in a total of 180 stimuli). In the recognition task, participants had to press "S" if they believed the word has already been presented and "N" if they thought it has not.

**Figure 2.**

*Procedure of the present study.*




Throughout the experiment, all stimuli were presented in Arial, size 60, in black with a white background. A white screen was presented for 250ms, followed by a fixation cross for 500ms and then the stimuli: In the LDT, the stimuli were presented for 2500ms and in the recognition task for 4000ms (but in the recognition task, the stimuli automatically advanced after an answer was given).

## Results

Accuracy in both LDT and PM task was analysed, as well as the criterion for the LDT (Table 2). Concerning the recognition task, hit rates, false alarm rates, and discrimination were analysed (Figure 3).

The Fife-Schaw law (Fife-Shaw, 2006) was applied

### Lexical Decision Task

Firstly, we examined the accuracy in the LDT, calculated by dividing the sum of hits (i.e., pressing "S" when a word was presented) and correct rejections (i.e., pressing "N" when a pseudoword was presented) by the total number of answers given. We excluded participants that responded to less than 80% of the LDT stimuli (n = 2).

To test the impact of the PM task and its difficulty in the LDT performance, and since normality was not verified, $p < .001$, a Kruskal-Wallis test was performed, followed by Dunn's post-hoc planned comparisons, to determine if there were differences between the three groups.

Lexical decision task accuracy was near ceiling ($M = .93$, $SD = .06$). As illustrated in Table 2, there was a significant difference between conditions, $H(2) = 19.81$, $p < .001$. The Hard-PM group presented a lower LDT accuracy when compared to the No-PM group, $Z = 4.38$, $p < .001$, and when compared to the Easy-PM group, $Z = 2.86$, $p < .001$. No difference was found between the No-PM group and the Easy-PM group, $Z = 1.53$, $p = .06$.

The criterion ($C$) values for the LDT were also analysed to test for the existence of any response bias among the three conditions. In this study, the criterion refers to the minimum evidence used to discriminate between words and pseudowords. The response criterion was calculated by adding the inverse of the normal distribution of the function of false alarms to the inverse of the normal distribution of the function of hits and then dividing by half, $C = .5[z(Hit) + z(False\ Alarm)]$ (Van der Kellen et al., 2008). The ANOVA performed revealed the presence of significant differences between the three conditions, $F(2, 156) = 23.63$, $p = < .001$, $\eta_p^2 = .23$. Tukey's post-hoc comparisons revealed that the Hard-PM group had the highest response criterion, $M =.25$, $SD = .32$, the Easy-PM group had the second highest response criterion, $M = .06$; $SD = .38$, and the No-PM group the lowest one, $M = -.25$, $SD = .45$.

These results show that participants in the No-PM group were less conservative, that is, needed less information to make a decision on each stimulus, than the ones in the Easy-PM and the latter were less conservative than the Hard-PM group.

**Prospective Memory Task**

Accuracy in the PM task was analysed to test for differences between the Easy-PM condition and the Hard-PM condition to access if both tasks were of different difficulty.

Correct PM task responses occurred when a participant pressed the "V" in the target trial or the following one (n = 1). Prospective memory accuracy was calculated by dividing the number of correct answers to PM cues by the number of responses given to the six PM cues, either correct or incorrect.

Results are presented in Table 2 and, as predicted, the two tasks (i.e., Easy-PM and Hard-PM) showed to be of different difficulties, $t(104) = 5.68$, $p < .001$, *Cohen's d* = 1.10, 95% CI [.69,1.51]. Prospective memory accuracy was higher in the Easy-PM condition, $M = .68$, $SD = .30$, than in the Hard-PM condition, $M = .36$, $SD = .27$. This result showed that the Hard-PM task used (i.e., multi-target PM task) was harder to perform than the Easy-PM task (i.e., single target PM task).

**Table 2**

*Means (M) and standard deviations (SD) of lexical decision task accuracy (LDT accuracy), prospective memory accuracy (PM accuracy), and lexical decision task criterion (LDT criterion) per condition*

| | No-PM | | Easy-PM | | Hard-PM | | |
|---|---|---|---|---|---|---|---|
| Measures | *M* | *SD* | *M* | *SD* | *M* | *SD* | *p* |
| LDT accuracy | .95 | .07 | .93 | .07 | .92 | .04 | < .001 |
| LDT criterion | - .25 | .45 | .06 | .38 | .26 | .32 | < .001 |
| PM accuracy | - | - | .68 | .30 | .36 | .27 | < .001 |

**Recognition Task**[2,3]

To understand the effect of PM difficulty on the incidental learning of the LDT nontargeted items, the main purpose, hit rates, false alarm rates and discrimination index (d') in the recognition task were analysed (Figure 3). A hit occurred whenever a participant pressed "S" as an answer to a word presented in the recognition task that was already presented in the LDT. When the participant pressed "S" in the presence of a new word (i.e., a word that was not presented in the LDT but was identified as such), this was taken as a false alarm. The discrimination index was calculated by subtracting the inverse of the normal distribution of the function of false alarms from the inverse of the normal distribution of the function of hits, $d' = z(Hit) - z(False\ Alarm)$ (Van der Kellen et al., 2008).

[2] Three outliers were included in the analysis because their inclusion does not significantly change the means.

[3] Prospective memory target words were not included in these analyses because of a possible enhanced distinction effect (n=6).

Regarding hit rates, no differences across conditions were found, $F(2, 156) = 2.72$, $p = .07$, $\eta_p^2 = .03$, nor in the false alarm rates, $F(2, 156) = .05$, $p = .95$, $\eta_p^2 = .00$. There were no significant differences between groups in the discrimination index, $F(2, 156) = 2.04$, $p = .13$, $\eta_p^2 = .03$.

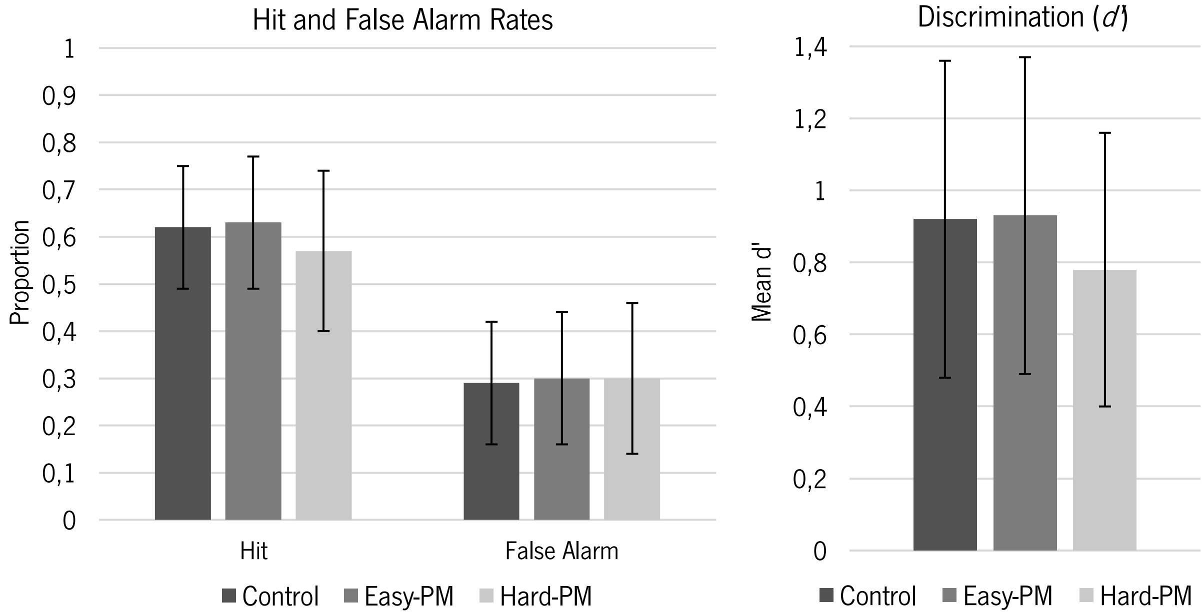



Figure 3. *Recognition task performance in terms of hit and false alarm rates (left) and d' (right) per condition.*

## Discussion

The present study aimed the understanding of the impact of PM intentions in incidental learning, more specifically in a recognition task, intending to contribute to the understanding of the processes that allow the retrieval of an action to be performed. In order to understand the effect of a PM task in incidental learning and if this effect is impacted by PM task's difficulty, participants performed a PM task embedded in an LDT. After a delay, a surprise recognition task, where the words previously presented in the LDT were intermixed with new words, was introduced for participants to distinguish between old and new words, thus testing incidental learning of those old words.

First, it was expected that the performance of a PM task during an LDT (i.e., the encoding phase of the words then presented in the recognition task) would lead to enhanced discrimination in the recognition task and that a harder PM task would create a "desirable difficulty", thus enhancing more the discrimination when compared to the possible enhancement effect of an easy PM task. However, no differences between conditions were found concerning discrimination in the recognition task, showing that the performance of a nonfocal PM task and its difficulty did not impact incidental learning. In contrast with Loft and Humphreys (2012) and Humphreys et al. (2020) findings, there was no enhancement effect in the recognition task performance for the experimental groups This apparent inconsistency can be explained by a methodological difference: In the aforementioned studies, detecting the nonfocal PM cues required semantic processing (i.e., participants had to perform the PM intended action whenever they saw an example of a certain category), while in the present study the detection of the nonfocal PM cues required more of a perceptual processing (i.e., performing the PM intended action in the presence of a syllable).

More specifically, lexical decisions require lexical access but no initiation of a search of a semantic network, but a nonfocal PM task of a semantic nature does require both processes to occur (Humphreys et al., 2020). The existence of a nonfocal PM task based on semantic features (i.e., categories) may have led, in those two studies, to different processing of the nontargeted LDT trials in the sense that participants initiated a semantic search for all stimuli (as suggested by Humphreys et al., 2020). But, in this study, since the nonfocal PM task was not semantic, semantic search was not initiated, only the lexical access, and thus the recognition task performance was not enhanced by the presence of the nonfocal PM task. The present study supports the findings and the respective explanations of Loft and Humphreys (2012) and Humphreys et al. (2020) regarding semantic processing as responsible for the enhancement effect found in their studies.

Regarding LDT accuracy, the hard-PM group had lower accuracy than the easy-PM group and the no-PM group. Arguably, this result could have occurred because participants in the hard-PM group were monitoring while performing both the LDT and the PM task, but since no differences were found on the recognition task performances, it is reasonable to suppose that during the LDT, all stimuli were equally attended, unlike the monitoring theory suggests. This idea might be supported by Boywitt and Rummel (2012) study that applied a diffusion model to PM functioning. A diffusion model sees binary decisions as a result of a set of cognitive processes that can be partitioned into four parameters: the drift rate ($v$), which concerns the rate of information uptake per time (it indicates the efficiency at information processing, the higher the drift rate, the more efficient the participant is in processing information); the upper threshold ($a$), also called *response criterion*, which is the minimum information necessary to decide on what response to give (the higher the criterion, the more conservative the behavior); and, lastly, the starting point ($z$) of the diffusion model, and the reaction time constant ($t_0$), which is the non-decisional reaction time component (i.e., all non-decisional parts of information processing like response execution; Boywitt & Rummel, 2012).

Boywitt and Rummel (2012) found that participants with a more demanding PM task to perform set the decision criteria/threshold higher, so more information is needed to select the action to be performed when compared to participants with a less demanding PM task at hand and the control group (that only performs an ongoing task). This is explained by the fact that participants with a harder PM task to perform, during the intention-formation phase (i.e., during the PM task instructions), perceive the task as more demanding than the ones with an easier PM task, leading to the adoption of different strategies to fulfill the PM intended action. Thus, participants who perform a hard-PM task are more conservative than those performing an easy-PM task (Boywitt & Rummel, 2012). Applying this idea to the present study, participants in the hard-PM group needed more information to make a response decision than participants in the other two groups. This idea is corroborated by the finding that the hard-PM group did present a higher decision criterion (*C*) than the easy-PM and the no-PM groups (see Table 2).

The same authors also found that under a more demanding PM task, the reaction time constant is bigger when compared to a less demanding one (Boywitt & Rummel, 2012) and again, applying this finding to the present study, participants in the hard-PM group needed more time to execute the selected response than participants in the other two groups.

In contrast to Boywitt and Rummel (2012) study, the present one had a time limit to answer each stimulus (2.5 s), and the instruction was to answer as quickly and accurately as possible, so participants in the hard-PM condition that had pre-set a higher threshold, may have lowered it to be able to respond to each stimulus on time, leading them to be less cautious in selecting the answer to give which, in turn, lead to a worst performance in the LDT when compared to the easy-PM group and the no-PM group so, a speed-accuracy trade-off occurred (Boywitt & Rummel, 2012) – in order to be able to respond faster, the accuracy was worst. Besides, from the time available to respond to stimuli, participants in the hard-PM condition needed more time for the nondecisional component and thus, leaving less time available for response selection. These two factors might explain the worst LDT and PM task performances of the hard-PM group when compared to the easy-PM and the no-PM groups.

To better substantiate this explanation, it would be interesting, in the future, to replicate this study in a lab context to be able to analyse data on RTs. It would also be interesting to replicate the present study but altering the time available to answer each stimulus (i.e., giving more time to participants to respond) to test if this speed-accuracy trade-off would or not occur in the hard-PM group. This would allow testing if participants in this condition performed worst in the LDT because of the time available to respond to each stimulus.

As we can see in the results, participants in the hard-PM condition continued to be more conservative than the ones in the other conditions. This result is consistent with the delay theory which suggests that individuals strategically set higher ongoing task response thresholds to delay ongoing task responses and allow more time for PM response selection (Heathcote et al., 2015). This can indicate that even though participants in the hard-PM condition may have lowered their pre-set criterion to respond to stimuli in the available time, they still had a higher criterion than the other two groups. This idea is supported by the absence of differences in LDT accuracy between the easy-PM group and the no-PM group and the presence of a difference between the criterion adopted by the two groups: Probably participants in the easy-PM group set the criterion higher to be able to select a response, but they did not have to lowered it because the pre-set criterion was low enough for participants to make a decision on the answer to give, unlike participants in the hard-PM group.

The present study supports the previous findings and explanations of Loft and Humphreys (2012) and Humphreys et al. (2020) regarding the enhancement effect in a recognition task as a result of the performance of a nonfocal PM task. This effect is dependent on the type of processing (i.e., perceptual or semantic) that the used nonfocal PM task requires. Besides, the results obtained can be explained using the application of the diffusion model to PM processes made by Boywitt and Rummel (2012). Lastly, by introducing a recognition task in the standard PM paradigm, the study is able to support the rationale behind the delay theory explanation for costs in the presence of a nonfocal PM task.

## References


Anderson, F. T., McDaniel, M. A., & Einstein, G. O. (2017). Remembering to remember: An examination of the cognitive processes underlying prospective memory. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (2nd ed., pp. 451-463). Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.21049-3

Anderson, F. T., Strube, M. J., & McDaniel, M. A. (2019). Toward a better understanding of costs in prospective memory: A meta-analytic review. *Psychologic Bulletin, 145*(11), 1053-1081. https://doi.org/10.1037/bul0000208

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113*(4), 518-540. https://doi.org/10.1037/0096-3445.113.4.518

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W..Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (2nd ed., pp 55-64). Worth Publishers.

Boywitt, C. D., & Rummel, J. (2012). A diffusion model analysis of task interference effects in prospective memory. *Memory & Cognition*, *40*(1), 70-82. https://doi.org/10.3758/s13421-011-0128-6

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 115–142). Psychology Press.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science, 14*(6), 286-290. https://doi.org/10.1111/j.0963-7214.2005.00382.x

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General, 134*(3), 327-342. https://doi.org/10.1037/0096-3445.134.3.327

Faul, F., Erdfelder, E., Lang, A. G., & 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

Fife-Schaw, C. (2006). Levels of Measurement. In G. M. Breakwell, S. Hammond, C. Fife-Schaw, & J. A. Smith (Eds.), Research methods in psychology (pp. 50–63). Sage Publications, Inc.

Heathcote, A., Loft, S., & Remington, R. W. (2015). Slow down and remember to remember! A delay theory of prospective memory costs. *Psychological Review, 122*(2), 376-410. https://doi.org/10.1037/a0038952

Hicks, J. L., Marsh, R. L., & Cook, G. I. (2005). Task interference in time-based, event-based, and dual intention prospective memory conditions. *Journal of Memory and Language, 53*(3), 430-444. https://doi.org/10.1016/j.jml.2005.04.001

Humphreys, M. S., Li, Y. R., Burt, J. S., & Loft, S. (2020). How semantic processing affects recognition memory. *Journal of Memory and Language*, *113*, [104109]. https://doi.org/10.1016/j.jml.2020.104109

Kelly, S. W. (2012). Incidental Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (1st ed., pp. 1517-1518). Springer. https://doi.org/10.1007/978-1-4419-1428-6_366

Loft, S., & Humphreys, M. (2012). Enhanced recognition of words previously presented in a task with nonfocal prospective memory requirements. *Psychonomic Bulletin & Review, 19*, 1142-1147. https://doi.org/10.3758/s13423-012-0303-1

Loft, S., & Remington, R. W. (2013). Wait a second: Brief delays in responding reduce focality effects in event-based prospective memory. *Quarterly Journal of Experimental Psychology, 66*(7), 1432-1447. https://doi.org/10.1080/17470218.2012.750677

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14*, 127-144. https://doi.org/10.1002/acp.775

Qualtrics (2020). Qualtrics (February 2021) [computer software]. Qualtrics.

Reese-Melancon, C., Harrington, E. E., & Kytola, K. L. (2019). How did I remember to do that? Self-reported strategy use for laboratory prospective memory tasks. *Memory, 27*(9), 1224-1235. https://doi.org/10.1080/09658211.2019.1645180

Rummel, J., & McDaniel, M. A. (2019). Introduction. In J. Rummel, & M. A. McDaniel (Eds.), *Prospective memory* (1st ed., pp 1-9). Routledge.

Strickland, L., Loft, S., & Heathcote, A. (2019). Evidence accumulation modeling of event-based prospective memory. In J. Rummel, & M. A. McDaniel (Eds.), *Prospective memory* (1st ed., pp 78-94). Routledge.

Scullin, M. K., Bugg, J. M., & McDaniel, M. A. (2012). Whoops, I did it again: Commission errors in prospective memory. *Psychology and Aging, 27*(1), 46-53. https://doi.org/10.1037/a0026112

Soares, A.P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. M. (2017). The Minho Word Pool: Norms for imagery, concreteness, and subjective frequency for 3,800 European Portuguese words. *Behavior Research Methods, 49*(3), 1065-1081. https://doi.org/10.3758/s13428-016-0767-4

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory and Cognition 29*(3), 347-361. https://doi.org/10.1037/0278-7393.29.3.347

Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory, 5*(6), 607-617. https://doi.org/10.1037/0278-7393.5.6.607

Van der Kellen, D., Nunes, L. D., & Garcia-Marques, L. (2008). Sensibilidade e bom senso: Princípios fundamentais da teoria de detecção de sinal na investigação em Psicologia. *Laboratório de Psicologia, 6*(1), 75-91. https://doi.org/10.14417/lp.694