Universidade do Minho
Escola de Psicologia

Rogério Gonçalves Dalot

Effect of font-size, real-size, and mental imagery on JOLs and memory

June, 2018

Universidade do Minho
Escola de Psicologia

Rogério Gonçalves Dalot

Effect of font-size, real-size, and mental imagery on JOLs and memory

Dissertação de Mestrado

Mestrado Integrado em Psicologia

Trabalho realizado sob orientação do(s)

Professor Doutor Emanuel Pedro Viana Barbas de Albuquerque

Professor Doutor Karlos Luna Ortega

June, 2018

Declaração

Nome: Rogério Gonçalves Dalot

Endereço electrónico: ge.dalot@gmail.com

Telemóvel: 916161903

Número do Cartão de Cidadão: 14347122

Título dissertação: Effect of font-size, real-size, and mental imagery on JOLs and memory recall

Orientação: Professor Doutor Emanuel Pedro Viana Barbas de Albuquerque, Professor Doutor Karlos Luna Ortega

Ano de conclusão: 2018

Designação do Mestrado: Mestrado Integrado em Psicologia

Universidade do Minho, 04/06/2018

Assinatura:____

Contents

Agradecimentos

Agradeço à minha namorada por me ter dado felicidade durante toda esta fase académica, especialmente quando se preocupa com o pH da água que bebe, enquanto come presunto com manteiga e azeitonas.

Agradeço a toda à minha família o apoio infinito. Desde a roupa lavada até ao carinho que tenho sorte de ter. Obrigado Mãe, obrigado Pai e obrigado irmão. Mãe, quando leres isto, podias fazer um incrível bacalhau à brás.

Agradeço à equipa de EPT todo o apoio prestado durante toda as minhas actividades. Agradeço todos os vossos bolos e chocolates disponíveis ao longo da minha fase de estágio e tese. Se bem que podiam reabastecer a água mais frequentemente, só o Vilaça consome metade do stock.

Agradeço a todos os meus amigos, especialmente a geração de 94, todos os bons momentos que fomos vivendo.

Agradeço ao Google a sua sapiência enorme.

Agradeço as tostas americanas do café do Luís, tanto como as suas natas.

Agradeço ao mau e ao bom café. Estou a brincar, não há mau café.

Agradeço os bons momentos de futebol. Força Diogo Dalot!

Agradeço a todas as coisas e pessoas que, de uma maneira ou de outra, me fizeram estar aqui a escrever os agradecimentos, desde o Alberto Constâncio mecânico até ao Jonathan filósofo.

Por último, gostaria de admitir que escrever a tese foi pessoalmente bastante custoso e deixo aqui os meus sinceros agradecimentos aos meus orientadores, Professor Karlos Luna e Professor Pedro Albuquerque. É impressionante a capacidade de trabalho do grupo de memória. Tanto trabalho não costuma estar aliado a tanta qualidade. Obrigado aos dois pela extensa paciência comigo.

Resumo

Um estudante enfrenta diariamente constantes desafios na aprendizagem e um deles é prever se o que estudou está realmente memorizado. Possivelmente, o estudante poderá usar uma pista específica do material estudado para calcular um julgamento de aprendizagem (JOL). Quando um JOL não corresponde à realidade e o estudante não memoriza o material estudado, este sofreu uma ilusão de metamemória. Existem atualmente duas perspetivas, uma baseada em processos experienciais e outra em processos cognitivos teóricos, onde tentam explicar a causa deste fenómeno. Na presente investigação, foram recolhidos JOLs de 60 participantes e feito um free recall test de uma lista de 64 palavras. Esta lista variava no tamanho da fonte (pequeno – 18 pt.; grande – 48 pt.), no tamanho do referente da palavra (pequeno – “agulha”; grande – “elefante”) e imaginabilidade (instruções para imaginar ou sem instruções). Os participantes percecionaram o tamanho da fonte grande como mais fácil de recordar enquanto o tamanho da fonte não teve qualquer efeito na recordação. Julgaram o tamanho de referente pequeno como mais fácil de recordar junto de uma melhoria significativa na memória, no entanto, este efeito foi encontrado apenas para o tamanho da fonte pequeno. Não foram encontrados efeitos na diferença de instruções para imaginar.

Palavras-chave: metamemória, fluência, crenças, JOLs, tamanho da fonte

Abstract

It is a daily task of a student to predict accurately which material it needs more study effort, in order to fully assure it is properly memorized. Sometimes, students might pick some specific cue to gather information on which material he will remember easier and at this point he calculates a judgement of learning (JOL). When a JOL does not correspond to the actual memory performance, it is called a metamemory illusion. There are two main theories to explain this phenomenon, theory-based and experienced-based processes. In this study, we collected JOLs and passed a free recall test to 60 participants (mainly psychology students) to each one of a 64 list of words. The list varied in font-size (small - 18 pt.; large - 48 pt.), in real size (small - “needle”; large - “elephant”) and in mental imagery (instructions to image each word or no instructions at all). Participants regarded larger font-size as more memorable with no differences in recall (font-size effect). Small real size was regarded as more memorable along with improved recall as well but this effect was only found on small font-size. No more significant effects were found.

Keywords: metamemory, fluency, beliefs, JOLs, font-size

Introduction

When we intend to learn a new material, it is our best intention to memorize as many concepts as possible. However, the ways we ought to learn them might not be the best, like for example, regarding content as really memorable because of any particular cue. One cue that a student might pick to judge the memorability of a content, could be the font-size. When the student would study these larger font-size contents, he probably would, mistakenly assume they are easier to remember (Rhodes & Castel, 2008). This would presumably lead us to study the content in larger font-size for less time, maybe jeopardizing our study to some extent. Additionally, this lesser study time can also serve as another cue for perceived memorability (Mueller, Dunlosky, Tauber, & Rhodes, 2014). The predictions about our learning are called judgments of learning (JOLs) and represent our confidence that we will remember a studied material in the future. We make JOLs when we are studying, so JOLs can be relevant to our daily basis on our education (Mueller & Dunlosky, 2017). Our thoughts, beliefs, and judgments about how our memory works is called metamemory, and if we think a factor affects memory, when in fact it does not, we suffer a metamemory illusion. As Rhodes and Castle (2008) showed, participants would regard larger font-size words as more memorable. However, free recall tests in their studies showed no differences in memory results between large and small font-size words. This serves as a clear example of a metamemory illusion. It is best known as the font-size effect and it has been replicated ever since its discovery (e.g., Mueller & Dunlosky, 2017).

Our metamemory processes cannot predict accurately, all the time, how our memory will perform. Our cognitive processes are not perfect. The reasons behind metamemory cognitive mistakes might have something to do with which cues we reside on. It is a matter that it is still up to debate and at the moment, there are two main approaches to explain how metamemory judgments are made: experience-based and theory-based processes (Besken, 2016). According to theory-based processes, a person’s judgment is based mainly on heuristics and beliefs of how memory works (Besken, 2016). According to experience-based processes, judgments are mainly based in factors streaming from the direct experience with the items. For example, perceptual fluency of materials would create a subjective experience of ease that participants would use to make the JOL (Besken, 2016).

If we take the font-size example from Rhodes and Castle (2008), according to the theory-based explanation, we would argue that higher JOLs are obtained because participants hold the belief that larger font-size facilitates recall. One way to test beliefs of participants is to ask them, without studying any word, what do they think they will remember better, large or small words (Mueller et al., 2014). Results showed that participants gave higher estimates for large words than small words, thus confirming the belief (Mueller et al., 2014). Another relevant experiment would be to test lexical decision times (the time one takes to decide if a group of letters is a word or not). Mueller et al. (2014) collected lexical decision times on words and non-words with large or small font size. Results showed no differences in decision times between small or large words which would indicate that a small word is as fluent as a large word. Thus, Mueller et al., state that theory-based processes (beliefs) would be the most reasonable explanation for the stated experiment.

From an experience-based perspective, the idea is that larger font size creates a subjective experience of ease of encoding or fluency on the participants, which they will use to make a judgement. An interesting study showed that intact words (presented for 2.5s) were regarded as more memorable compared to interfered words (i.e., presented for 83ms and then, letters were substituted by "X" for the remaining 2.417s), while recall was significantly higher for the interfered words (Besken & Mulligan, 2013). This finding shows that one could in fact use, fluency of items, to produce JOLs. Although memory was not affected between intact and interfered conditions, participants suffered a metamemory illusion, regarding intact words as more memorable (Besken & Mulligan, 2013). Later, Besken (2016) found that participants gave higher JOLs to intact words compared with blurred words, also suggesting an effect of fluency on JOLs. Rhodes and Castle (2009) found that participants gave higher JOLs to words heard in high volume relative to words in low volume. Rhodes and Castle (2008) also studied JOLs on an alternate case ("AlTeRnAtE") and concluded that participants would provide lower JOLs because the alternate case is supposedly less fluent. Therefore, there is a reasonable amount of empiric evidence demonstrating the role of fluency on JOLs.

To access one's metamemory judgments, researchers can use JOLs as in the studies above. Researchers use JOLs to access the metamemory judgments of the participants. JOLs can be provided before studying one item (prestudy JOLs), immediately after studying one item (immediate JOLs), or after a delay or a study cycle (delayed JOLs). We should vary the JOLs we ask accordingly on what we want to study.

For example, if we ask for prestudy JOLs, participants have not yet seen the word, so they cannot know if a word is in large font size or not. So, prestudy JOLs are thought to be more suitable to study beliefs. Since the participant has not yet seen the word, the participant would evaluate future memory only accordingly to his beliefs (Mueller et al., 2014). However, immediate JOLs might be a reflection of a combination of beliefs and fluency (Price & Harrison, 2017). Delayed JOLs might also have a combination of both factors, and they have been shown to be more accurate in relation to memory performance than immediate JOLs (the delayed-JOL effect, see Dunlosky & Nelson, 1992; Luna, Martín-Luengo, & Albuquerque, in press). One interesting finding by Price and Harrison (2017) was that, when participants provided both prestudy and immediate JOLs, immediate JOLs had a significantly lower magnitude, indicating that participants were regarding additional information on immediate JOLs to evaluate words and not only beliefs like on prestudy JOLs.

These results are in line with the idea that JOLs are made from many different cues, including theory-based factors and experience-based factors, and that every cue deserves particular attention. Besides font size, several other cues for JOLs have already been studied, such as concreteness (Witherby & Tauber, 2016), physical weight (Alban & Kelly, 2013), mental imagery size (Li et al., 2016), among others. Similar to the font-size effect described above, participants gave higher JOLs to more concrete words than to more abstract words (Witherby & Tauber, 2016). In fact, concreteness actually does present better recall performance (Witherby & Tauber, 2016). Thus, participants would correctly presume an increased memorability in regard to a word like “pencil” than “peace” or “love”. Concerning physical weight, it may lead to a metamemory illusion, because participants provided higher JOLs when they were holding heavier clipboards to write their JOLs, although it did not have any effect on recall (Alban & Kelley, 2013). Regarding mental imagery, participants provided higher JOLs for larger mental images of to-be-remembered words (Chinese characters) with no significant improvement on recall (Li et al., 2016). Therefore, these results should be taken into account while the selection of the list of words, leading researchers to control these variables, such as concreteness, imaginability, or other particular cue that might be significantly affecting JOLs of participants like physical weight. Although the present study does not have anything to do with embodied cues like physical weight, its relevance to the present study lies in the fact that Alban and Kelley (2013), argued that participants might regard something as more important, leading to higher JOLs. This argument, is somewhat analogous to those of the beliefs, being also part of a theory-based perspective. If participants regarded increased physical weight as more important and that led to increased JOLs, it may also holds true with other cues.

1.1 Objectives

In the present study, we aimed to study the following cues on JOLs: the size of the referent of the word (i.e., whether it refers to a large object/animal or to a small one, henceforth referred to as real size), the font size of the words presented, and mental imagery. The font-size effect is already a very known phenomenon, but there is no clear answer as to how font size behaves along with other cues, like the size of the animal/object and the mental image created by a participant. Especially with the real size, since we did not find any study that manipulated this variable. For example, we do not know if words with different real sizes, like “needle” (small real size) or “elephant” (large real size) are rated with different JOLs. Concerning mental imagery, our interest in mental imagery is to study if there is any substantial effects on JOLs or memory recall when creating a mental image.

We hypothesized that the larger font-size of a word, the larger will be the JOLs, with no differences on memory results, replicating the already known font-size effect. There is no precedent for the effect of real size on JOLs, so our interest in that variable was mostly exploratory. According to other known effects, participants might regard words with larger real size as easier to process (fluency) or participants might hold a belief that large things are easier to remember or even maybe regarding larger things as important, as with weight (Alban & Kelley, 2013). Since there is no external indicator of the real size of the word, we did not expect that experience-based factors could affect JOLs for real size. On the other side, if real size has an effect on JOLs, it is more likely that it is due to theory-based factors, like beliefs. Thus, we hypothesized that participants will hold a belief that large real size objects are easier to remember, leading to higher JOLs. We also expect no differences on memory regarding real size.

About mental imagery, Li et al. (2016) in Experiment 1 found that participants regarded to-be-remembered items as more memorable if they have created a larger mental image of the word (Chinese character) than those who created a smaller mental image. In the same study, on Experiment 3, they found two interesting results. First, they recorded the time participants needed to create these images (participants would push a button after the creation of the mental image) and they found that larger mental images took longer time. Second, the difference on JOLs between large and small mental images from Experiment 1 disappeared. Relative to the first result, one can hypothesize that if larger mental images took longer time, it is probably because it demands a greater effort, which means it is less fluent. If it is less fluent, JOLs should be lower than smaller mental images, which did not happen. This suggests that fluency might not be the explaining factor. About the second result there is an interesting argument from the authors. When participants had to push a button until a mental image was created, participants would notice the different amount of times needed to generate those mental images. These times might serve as an experiential fluency cue when making JOLs. It has been shown that faster study times might be regarded as more fluent, leading to higher JOLs (Castel et al., 2007; Mueller et al., 2014). So, a faster study time on smaller mental images might counteract the participant's belief that larger mental images are more memorable. In Experiment 1 participants were not pushing a button, so they did not have any reasonable way of taking into account the time they were providing to each mental image. Being so, beliefs might had been the main cause for higher JOLs to larger mental images. To further explore the contribution of beliefs, Li et al. (2016) conducted an additional experiment using a questionnaire. Results showed that participants regarded larger mental images as more memorable than small mental images.

In the present study we predict that, since we do not think participants will notice any different experiential fluency cue between items, beliefs would mediate as the main influence on JOLs. Thus, we expect JOLs will be higher for those who create a mental image, believing that such thing will benefit memory more than those who do not create mental images. Our hypothesis relative to memory goes in the same direction as the others: no significant differences are expected.

2 Method

2.1 Participants

Sixty students from the University of Minho (M_{age} = 20.57; *SD* = 2.99; 85% Female) were recruited for our study. The majority were studying psychology (*n* = 55) and they gained partial course credits by participating in the experiment. The other 5 participants were studying other subjects and volunteered for the task.

2.2 Design

We conducted a 2 (font size: large 48 pt. or small 18 pt.) x 2 (real size: large or small) x 2 (mental imagery: imagery or no imagery) mixed factorial design. Font size and real size were manipulated within participant and mental imagery was manipulated between participants. We randomly divided participants into two groups of 30 participants each. There were two dependent variables: judgments of learning (JOLs) and memory performance.

2.3 Materials

We used the Minho Word Pool database from Soares et al. (2017) to develop the experimental materials. We filtered the database to obtain nouns which scored more than 5 (scale 1-7) on imageability and concreteness word attributes. Imageability is how easy it is for someone to picture a word and concreteness is how much a word corresponds to something physically real. We only selected words from 6 to 8 letters. After having a list of words with the aforementioned characteristics, we accessed the frequency of each word on the Portuguese contemporary lexicon through a web interface database, called "Procura-Palavras" (P-PAL). Later, font size was counterbalanced between participants. Thus, for example, on one list we would had "airplane" in 18 pt. and on the other list, it would be in 48 pt. font size. Finally, we divided the final list of words into two separate categories, large (i.e., factory) and small (i.e., bee) real size. The final list included 64 words, 32 words to be presented in 18 pt. font size and the other 32 words in 48 pt. Resulting in thirty two small font-size words, compounded of 16 small real size and 16 large real size. The same way goes for the other 32 words of large font-size. The experiment was built on LiveCode, an open-source software, which resulted in an executable application.

2.4 Procedure

All participants read and signed an informed consent. Then, they entered a soundproof cabin with a computer, chair, keypad, and mouse. Before starting the tasks they provided information about sex and age and they read relevant instructions regarding the experiment. The imagery group received instructions to imagine every word they would soon read, as vividly as possible. The non-imagery group was tested with no instructions regarding how they should study the word. Every single word was presented centered in a computer screen with a light-blue background screen for 5 seconds. The list had 2 words at the beginning and at the end of the list serving as buffers to control primacy and recency effects. All words were presented in Arial font. After each word, they would provide JOLs to each word in a scale from 0 (they will not remember the word) to 100 (they will remember the word) spaced between intervals of 10 (0, 10, 20…90, 100) on a total of 11 possible radio-buttons on the computer screen. After they evaluated JOLs on every single word, participants completed two filler tasks for 3 minutes each. One of them consisted on writing cities of Portugal and the other one writing countries of the world, the maximum number of cities and countries they could provide. Finally, they were instructed to recall as many items as possible and input them on the keyboard of the computer for 5 minutes. Each participant took around 20 minutes to complete the experiment.

2.5 Data analysis

We analyzed data through JASP, a free open-source software. We ran repeated measure ANOVAs (Analysis of Variance) and independent t-tests. We also calculated Goodman-Kruskal gamma correlations to test how much accurate the magnitude of JOLs are relative to actual recall performance.

3 Results

A 2 (font size) × 2 (real size) × 2 (mental imagery) mixed design ANOVA was used to analyze our results. Font size and real size were manipulated within-subject and mental imagery was manipulated between-subject.

3.1 Judgments of Learning

There was a statistically significant effect on font size, $F(1, 58) = 20.22$, $p < .001$, $\eta_p^2 = .26$. Participants gave higher JOLs to large font size ($M = 43.59$, $SD = 15.70$) compared to small font size ($M = 39.53$, $SD = 15.53$). There was also a significant effect on real size, being that small real size ($M = 42.66$, $SD = 15.66$) received higher JOLs compared to large real size ($M = 40.46$, $SD = 15.56$), $F(1, 58) = 6.13$, $p = .016$, $\eta_p^2 = .09$.

Mean JOLs regarding font size and real size are shown in Figure 1. There was no significant differences between the imagine *(M* = 41.43, *SD* = 23.85) and not-imagine conditions *(M* = 42.68, *SD* = 24.35), *F(*1, 58) = 0.01, *p* = .950, $\eta_p^2 < .01$. There was no statistically significant effect found for font size and mental imagery interaction, *F(*1, 58) = 0.40, *p* = .529, $\eta_p^2 = .01$, neither between real size and mental imagery, *F(*1, 58) = 0.752, *p* = 0.390, $\eta_p^2 = .03$, and neither between font size and real size, *F(*1, 58) = 0.36, *p* = 0.550, $\eta_p^2 = .01$. There was also no statistically significant effect between font size, real size and mental imagery interaction, *F(*1, 58) = 1.225, *p* = 0.273, $\eta_p^2 = .02$

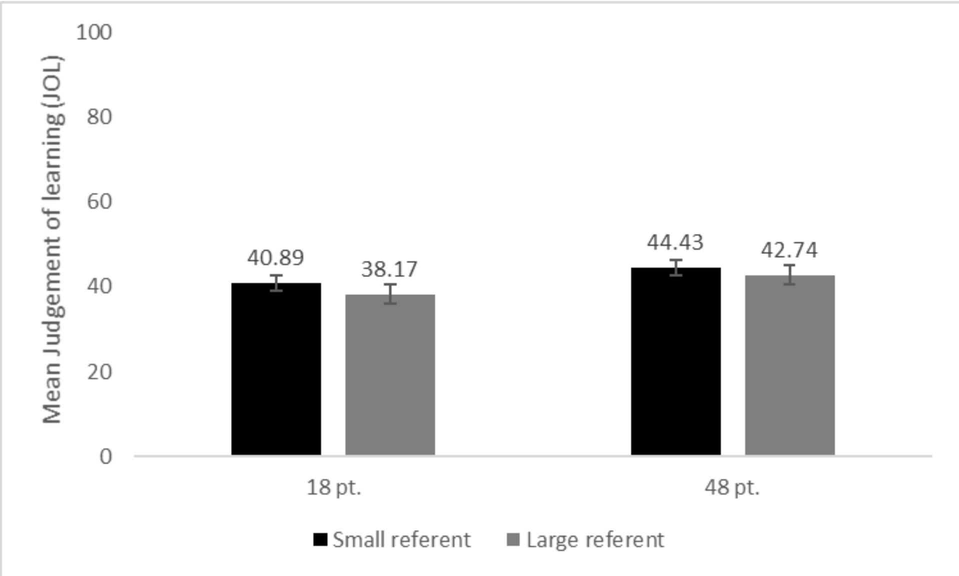

Figure 1. Mean JOLs scores by real size and font size.

3.2 Memory

Table 1 – Mean proportions (standard deviation) of recall for font size, real size, and instructions to imagine.

	18 pt.		**48 pt.**	
	Large real size	**Small real size**	**Large real size**	**Small real size**
Imagine	.24 (.14)	.32 (.17)	.27 (.17)	.28 (.11)
Not Imagine	.25 (.12)	.34 (.16)	.30 (.15)	.32 (.12)

All results are showed in Table 1. Proportion of recalled words are shown in Figure 2 for only font size and real size. There was no statistically significant effect on font size.

Participants did not recall better large font-size words *(M* = .29, *SD* = .12), compared to small font-size words *(M* = .29, *SD* = .13), *F(*1, 58) < 0.01, *p* = .959, η_p^2 < .01. There was a significant effect on real size, being that small real-size words were better recalled *(M* = .31, *SD* = .13), compared to large real-size words *(M* = .26, *SD* = .13), *F(*1, 58) = 12.69, *p* < .001, η_p^2 = .18. There were no differences in memory recall regarding mental imagery, *F(*1,58) = 0.758, *p* = .387, $\eta_p^{2\,=}$ .013. Participants who had instructions to imagine did not recall better *(M* = .28, *SD* = .12) than those who did not have any instructions *(M* = .30, *SD* = .11, *F(*1, 58) = 0.76, *p* = .387, η_p^2 = .01. There was a statistically significant effect between font size and real size, *F(*1, 58) = 7.28, *p* = .009, η_p^2 = .11. To further analyze this interaction, we ran paired samples t-tests. There was significant differences in recall for real size with small font size, *p* < .001, *d* = -.570, but no differences with large font size, *p* = .531, *d* = -.081 (see Figure 2). There was no statistically significant effect found for font size and mental imagery, *F(*1, 58) = 0.46, *p* = .501, η_p^2 < .01, neither between real size and mental imagery, *F(*1,58) = 0.19, *p* = .668, η_p^2 < .03. No statistically significant effect between font size, real size and mental imagery, *F(*1, 58) < 0.01, *p* = .968, η_p^2 < .01.

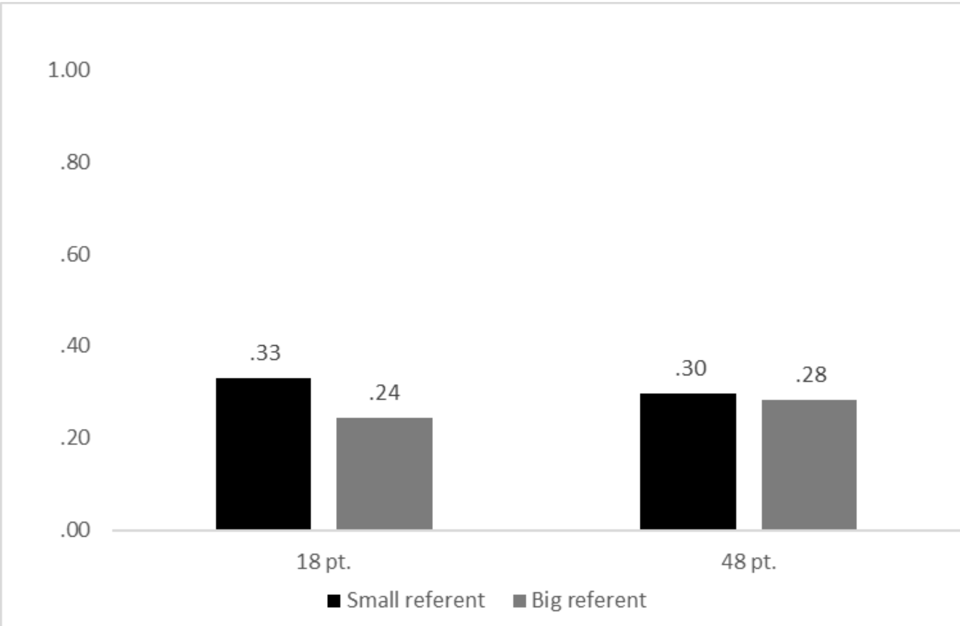

Figure 2. Proportion of the recalled words by font size and real size

3.3 Gamma correlations

The relationship between JOLs and recall was examined by calculating the Goodman-Kruskal gamma for each participant. Goodman-Kruskal gamma is an index of the correspondence between JOLs and accuracy, that is, if JOLs are related to memory performance. Gamma tends to 1 when participants' JOLs are accurate and predict later memory performance. If participants tend to wrongly predict their memory performance, meaning that their JOLs do not correspond to their subsequent recall, gamma will tend to -1. If there are no correspondence at all between JOLs and recall, gamma value would tend to 0.

Gamma for the full sample was significant and positive (*M* = .24, *SD* = .29), $t(59) = 6.30$, $p < .001$, $d = .83$, which means that JOLs can predict memory to some extents. In this case, participants, tend to recall more those words rated with higher JOLs, and tend to recall less words rated with lower JOLs.

We also compared gamma between each condition within each variable, being that there were no significant differences between large-font (*M* = .27, *SD* = .30), and small-font (*M* = .22, *SD* = .37), $p = .334$, $d = .13$. No significant differences between large (*M* = .27, *SD* = .30) and small real size (*M* = .22, *SD* = .37), $p = .716$, $d = -.05$ and no significant differences between Imagine (*M* = .22, *SD* = .32) and Not Imagine (*M* = .26, *SD* = .26), $p = .539$, $d = -.11$. Regarding each one of our variables, the gamma, on every condition is statistically different from zero, $p < .001$, except on Imagine condition, where the result was $p = .001$.

4 Discussion

Our main objectives were replicating the font-size effect while trying to explore the effect of the real size and to understand the role of mental imagery. We did replicate the font-size effect and regarding real size, we had interesting results that were different from our hypothesis. Besides, we did not find any statistically significant effect on mental imagery, which also goes against our predictions.

The font-size effect has shown to be a very consistent metamemory illusion in literature. In the present study it was found higher JOLs scores for larger font size, in accordance to our hypothesis, being one more relevant replication to the already known font-size effect, reported along the present study (Luna et al., in press; Mueller et al., 2014; Mueller & Dunlosky, 2017; Rhodes & Castel, 2008). The present study is a relevant replication because it has a unique set of variables, not yet been studied in conjunction.

An interesting, unexpected result is the real size of the words that seemed to affect

JOLs and memory. Being that small real-size words were regarded as more memorable and were better recalled than large real size words, it goes against our initial hypothesis, on both JOLs and recall. Participants are able to make a correct metamemory judgment, so possible explanations might point towards memory processes and not metamemory. Also, an interesting detail is that words in small font size did differ significantly on JOLs and memory between small and large real size, but not with large font size. Why did recall and JOLs scores on real size only differ with small font size? What happens with large font size that makes this real-size effect disappears? One simple explanation lies in the possible improved memory, that small real size facilitates cognitive processes and participants are just able to predict it, whether based on fluency or beliefs, providing higher JOLs. One alternative explanation is that when participants see words in large font size, this sense of “large” category is not congruent to the real-size of the object if its real size is small. This might provoke a cognitive dissonance that might disturb cognitive processes. On the other side, a large font size word is congruent with a large real-size word in the category of “large” or “big”. Thus, words would be regarded as more memorable and better recalled, whether they were a result of a belief of the participant or a familiar/fluent experience in terms of metamemory and a result of a lesser load on cognitive processes in terms of memory performance.

Mental imagery did not show any significant difference which goes against our hypothesis. This could be because the procedure was not efficient on one of the following situations or a combination of them: (1) Participants did not really imagine the words. Participants might have completely disregard the instruction, especially those that had high levels of motivation. Some participants, when informally asked how the experiment had gone, explicitly told that they adopted another strategy, which did not include imagining those words. They thought they would memorize better that way. The participants in the Not Imagine condition, might have followed a similar path, ending both groups to have adopted similar strategies. (2) Participants did imagine the words on both conditions. If participants in the Not Imagine condition, ended up imagining the words, maybe trying to have a better recall, this would also ended up with both groups having the same conditions. (3) Participants did imagine the initial words but forgot to imagine the following ones. The instructions were embedded in a portion of text which could be unintentionally disregarded somewhere in the experiment. Future procedures should have the above issues taken into account carefully, to improve control to the experiment. For example, in situation (3), maybe there should be a slide remembering to imagine the word (i.e. “Imagine the following word”) before each word is presented. In an exploratory matter, it could be interesting to have a third condition with instructions to not imagine a word. For example, “Do not Imagine a/an”, and in a next slide, “elephant”. There could be a paradoxically effect, where participants would indeed imagine each word they were instructed not do so. There is a possibility that mental imagery does not have any effect on JOLs or memory but, since it is really difficult to control such variable, it is too soon to advocate such conclusion.

Gamma scores were all significantly different from zero, which shows that participants tend to be consistent on recall accordingly to the JOLs provided. In addition, no differences were found between each condition within each variable, which suggests that no variable is more predictive of better correspondence between JOLs and memory recall.

Font-size effect is a robust phenomenon of the metacognition processes and it is today a well-known established effect. However, there is still work to do to better understand its causes. In which circumstances, belief and fluency perspectives, do a better job explaining the font-size effect? Both sides have reasonable arguments and it seems very difficult to reach the main causes for each study results. Some argue that the answer must lie in-between. The answer is far from being a simple one to attain. There is an interesting theory, the analytic-processing (AP) theory, adapted by Mueller and Dunlosky (2017), basing on former propositions. AP theory advocates that participants, when forming a JOL, enter in an analytical, problem-solving cognitive stance to search for relevant information to deliver an accurate JOL. In this case, this problem-solving state consisting in processing information leads one to form beliefs about how his memory will then perform. Participants can either have a-priori beliefs or form them on-line, during the experiment (Mueller & Dunlosky, 2017). This theory is not exclusively about beliefs, although it establishes beliefs as the main explaining factor to how JOLs are developed.

Regarding both theories, experience-based processes and theory-based processes, our study is not conclusive to favor any of them. Even if we had questionnaires or lexical decision times to access the root cause of the JOLs, it would still be debatable. It is continuous struggle to investigators, to know which perspective is the main cause. Remind the study where words were presented in “AlTeRnAtE” case (Rhodes & Castel, 2008). Authors argued that this happened because alternate case is less fluent so they would be regarded as less memorable. Mueller et al. (2017) challenged that argument and argued that larger font-size words are associated with higher JOLs because of a belief that larger font-size words are more memorable. When words are presented in alternate case this belief would be disrupted and participants would form another one which consists that words in an alternate case are more difficult to memorize.

It is a valid argument in Mueller et al. (2014) study, which shows the complexity of the issue and the tireless effort of all the researchers to understand the causes for the metamemory illusions. Even though one may construct an experiment with results apparently favoring fluency, one could still argue that beliefs are still the leading reason for such effects. Another reasonable explanation is that both theories may be accountable.

In the present study, font-size effect remains intact in such unique study and small real size seemed to have a significant increase in both memory performance and JOLs. We did not find similar studies to support our results regarding the real-size effect. Mental imagery lack of effects cannot be conclusive because the procedure could not guarantee that participants were, in fact, imagining each word. Future studies should take this into account and further analyze the real size variable, which seems to have an interesting say on metamemory and or memory.

5 References

Alban, M. W., & Kelley, C. M. (2013). Embodiment meets metamemory: Weight as a cue for metacognitive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1628–1634. https://doi.org/10.1037/a0032420

Besken, M. (2016). Picture-perfect is not perfect for metamemory: Testing the perceptual fluency hypothesis with degraded images. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1417–1433. https://doi.org/10.1037/xlm0000246.

Besken, M., & Mulligan, N.W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition, 41*, 897-903. https://https://doi.org/10.3758/s13421-013-0307-8.

Castel, A.D., McCabe, D.P., & Roediger, H.L (2007). Illusions of competence and overestimation of associative memory for identical items: Evidence from judgments of learning. *Psychonomic Bulletin & Review, 14*, 107-111. https://doi.org/10.3758/BF03194036.

Dunlosky, J., & Nelson, T.O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, *20*, 374-80.

JASP Team (2018). JASP (Version 0.8.5)[Computer software].

Li, T., Hu, X., Zheng, J., Su, N., Liu, Z., & Luo, L. (2016). The influence of visual mental imagery size on metamemory accuracy in judgment of learning. *Memory, 25*, 244-253. https://doi.org/10.1080/09658211.2016.1153658.

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

Mueller, M. L., & Dunlosky, J. (2017). How beliefs can impact judgments of learning: Evaluating analytic lprocessing theory with beliefs about fluency. *Journal of Memory and Language*, *93*, 245–258. https://doi.org/10.1016/j.jml.2016.10.008.

Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language, 70*, 1–12. https://doi.org/10.1016/j.jml.2013.09.007.

Price, J., & Harrison, A. (2017). Examining what prestudy and immediate judgments of learning reveal about the bases of metamemory judgments. *Journal of Memory and Language, 94*, 177–194. https://doi.org/10.1016/j.jml.2016.12.003.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology General*, *137*, 615–625. https://doi.org/10.1037/a0013684.

Rhodes, M.G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review, 16*, 550-554. https://doi.org/10.3758/PBR.16.3.550.

Soares, A.P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. M. (2017). Norms for imageability, concreteness and subjective frequency for 3,800 Portuguese words. *Behavior Research Methods, 49*, 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., Machado, J., & Comesaña, M. (2014). Procura-PALavras (P-PAL): Uma nova medida de frequência lexical do Português Europeu contemporâneo [Procura-PALavras (P-PAL): A new measure of word frequency for contemporary European Portuguese]. *Psicologia: Reflexão e Crítica, 27*, 1-14.

Witherby, A. E., & Tauber, S. K. (2016). The concreteness effect on judgments of learning: Evaluating the contributions of fluency and beliefs. *Memory & Cognition, 45*, 639-650. https://doi.org/10.3758/s13421-016-0681-0.