Journal of Memory and Language 144 (2025) 104673

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Exploring the animacy effect in focal prospective memory tasks: When animates don't stand out

Sara B. Félix [a,b,*], Marie Poirier [b], Josefa N.S. Pandeirada [a]

[a] William James Center for Research, Department of Education and Psychology, University of Aveiro 3810-193 Aveiro, Portugal

[b] Department of Psychology, School of Health and Medical Sciences, City, St-George's, University of London, London, UK

ARTICLE INFO

Keywords:
Adaptive memory
Animacy
Focal tasks
Prospective memory
Proximate mechanisms

ABSTRACT

The animacy effect refers to a memory advantage for animates/living beings as compared to inanimates/nonliving things. So far, the animacy effect has been investigated mostly in retrospective memory. Given that memory serves a future-oriented function, and considering the adaptive significance of animacy, it has been proposed that it should also confer an advantage in prospective memory (i.e., memory for intentions/actions to-be-performed in the future). Recent research reported an animacy effect in nonfocal event-based prospective memory tasks. The present work explored this effect in focal prospective memory. In a series of five studies, conducted in different countries and languages, we employed various ongoing tasks. Across all studies, no differences in prospective memory performance between animates and inanimates were found. This result held in a sign-test including all participants ($N = 408$ young adults) for a more powered analysis. Also, no differences between animates and inanimates were obtained in the baseline and filler trials. These results are discussed considering the mechanisms that have been proposed to explain the effect in retrospective memory tasks, namely attention-prioritization and richness of encoding. Overall, our results are partially explained by the attention-prioritization account of the animacy effect and also provide support for the Multiprocess Framework.

Introduction

The adaptive memory framework posits that our cognitive systems (such as memory) should be tuned to favor the processing of fitness-relevant information, that is, information that is relevant to our survival and/or reproduction (Cosmides & Tooby, 1992; Nairne et al., 2017). An example of an adaptive mnemonic tuning is the *animacy effect*, which reflects a remembering advantage for animates (living beings) than inanimates (nonliving things). Indeed, it would have been, and continues to be, adaptive to prioritize remembering animates, as they are considered to be more fitness-relevant than inanimates. Indeed, animates could represent friends (kin), foes (competitors, predators), prey, potential sexual mates, among other possibilities (Nairne et al., 2017, 2013).

It is well established that animacy plays a relevant role in various domains, including language and memory. For example, people tend to recognize strings of letters as corresponding to words faster when they represent animates, as opposed to inanimates (Bonin et al., 2019; Ferré et al., 2023; but see Radanović et al., 2016). In retrospective memory, animacy has been shown to be one of the best predictors of free recall (Aka et al., 2023, 2021; Madan, 2021; Nairne et al., 2013). Also, the animacy mnemonic advantage has been found in different forms of memory (e.g., working memory: Daley et al., 2020; meta-memory/judgements of learning: DeYoung & Serra, 2021; Li et al., 2016) and using several procedures (e.g., immediate and delayed recall: Félix, Pandeirada, et al., 2019; incidental and intentional learning: Félix & Pandeirada, 2024; Félix, Pandeirada, et al., 2019; Gelin et al., 2017; Komar et al., 2023b; in full and divided attention conditions: Leding, 2019; Rawlinson & Kelley, 2021). Despite the robustness of the effect, there are some conditions in which the effect has not been consistently obtained (e.g., cued-recall: Mah et al., 2023; Popp & Serra, 2016; VanArsdall et al., 2015; recognition: Félix & Pandeirada, 2024; Komar et al., 2023a; Leding, 2020; Rawlinson & Kelley, 2021).

Memory serves several valuable functions, extending beyond simply recalling past information and experiences. Notably, memory has a future-oriented focus (Schacter et al., 2007) which allows us to retrieve information stored in the past to aid solving problems in the present, and helps to anticipate and plan for future events (Klein et al., 2010; Schacter et al., 2007). This is a crucial ability for our adaptiveness (e.g., remembering to avoid a place where I saw a lion previously, Nairne &

* Corresponding author at: Department of Education and Psychology, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal.
E-mail address: sara.felix@ua.pt (S.B. Félix).

https://doi.org/10.1016/j.jml.2025.104673
Received 13 November 2024; Received in revised form 28 May 2025; Accepted 26 June 2025
Available online 18 July 2025

Pandeirada, 2008; or to avoid cheaters, Schaper et al., 2022). In this context, prospective memory (PM) – the ability to remember and execute future intentions or actions – is thought to be central (Einstein & McDaniel, 1990). Combining these two adaptive components – the animacy variable and prospective memory – we hypothesized that animacy would also impact PM performance.

Examples of everyday PM tasks are remembering to deliver a message to a friend when encountering him/her, to go to a medical appointment, or to take medication; many of these involve interactions with other people (i.e., animates; Brandimonte & Ferrante, 2008; Loft et al., 2019). The success or failure of PM tasks can have significant consequences for individuals themselves (such as health issues resulting from undermedication or overmedication), and their relationships with others (for instance, forgetting to call friends on their birthday, or to initiate safety routines on an airplane; Nowinski et al., 2003), thus having clear adaptive implications.

In everyday life, people form intentions and make plans to be performed sometime in the future. However, people are also often occupied with (and busily engaged in) various tasks that need to be interrupted to fulfil a prospective memory intention. For example, while driving home and noticing an approaching supermarket [PM target], one may remember needing to buy bread and decide to stop and make that purchase [PM response].

Researchers have investigated PM in the laboratory using procedures that mimic these circumstances. There are several types of PM tasks (for an overview, see Kvavilashvili & Ellis, 1996); our work focused on event-based tasks, in which the presence of an event or target signals the moment to perform the future-intention/PM response. In a typical event-based PM paradigm, participants are tasked to provide an alternative response (the PM response) whenever the PM targets are presented, while being busily engaged in an unrelated ongoing task. For example, while pressing the M or Z keys for responding 'word' or 'nonword' during a lexical decision task (ongoing task), participants should press F1 (PM response) whenever the target word *tortoise* appears (McDaniel & Einstein, 2000). The trials involving the target words are the *target trials*, whereas those regarding the ongoing task are named the *filler trials*. Sometimes a baseline phase is included, where participants perform the ongoing task only, providing a reference point about their performance on that task (*baseline trials*).

Moreover, PM tasks can vary in terms of focality. A task is focal when the cognitive processes required to detect the PM targets overlap with those involved in performing the ongoing task. On the other hand, it is considered nonfocal when such an overlap is absent, introducing additional demands on the available cognitive resources (Einstein et al., 2005).[1] For example, if the ongoing task is a lexical decision task, a focal PM task would be to provide a different response whenever a target word is presented (e.g., *tortoise*), as both the ongoing and the PM task involve the processing of words as a whole. In the nonfocal version of this example, a PM response would be required whenever, for instance, a target syllable is presented (e.g., *tor*). In this case, as the identification and response to the target word no longer require reading the whole word, but rather emphasize another aspect of the stimuli, less overlap exists on the processes involved in performing the ongoing (lexical decision task) and the PM (syllable identification) task; in fact, some competition for the available cognitive resources might occur between the two tasks.

There are several theories and models aiming to explain PM functioning (for a recent overview, see Bayen et al., 2024; Rummel & Kvavilashvili, 2023). One of the most prominent and robust models is the (Dynamic) Multiprocess Framework (McDaniel & Einstein, 2000; Scullin et al., 2013; Shelton & Scullin, 2017). According to this framework, participants rely more heavily on a bottom-up spontaneous identification of the targets in focal tasks (i.e., the recognition of the targets occurs in a more automatic way; Cona et al., 2014), while nonfocal tasks require top-down effortful processes (e.g., strategic monitoring) to detect the PM targets.

The animacy variable has been usually overlooked in the PM literature. Still, animate (e.g., *tortoise, butterfly*) and/or inanimate words (e.g., *dormitory, towel*) have been used frequently as targets in event-based PM tasks (e.g., Chen et al., 2014; Einstein & McDaniel, 2005). In other studies, participants were asked to provide the PM response whenever targets from a specific category were presented, including categories of animate items (e.g., four-footed animals) and inanimate things (e.g., metals, furniture; Moyes et al., 2019; Zuber et al., 2019). However, the potential influence of animacy on PM performance was not considered.

An ultimate mechanism rationale provides the evolutionary argument for why humans should have a PM memory tuning for animates: the enhancement of fitness. Recently, Félix et al. (2024) reported, for the first time, that using animate words as targets for the PM task improved performance as compared to using inanimate targets. This finding of an animacy advantage in PM held in the three nonfocal PM studies that used different procedures and in different languages, showing the generality of the findings. Nonfocal PM tasks call for the maintenance of controlled attention (e.g., monitoring) to detect the PM targets. However, relying on controlled processes would be costly for (and, thus, compromise) the performance of ongoing/everyday activities (Anderson et al., 2019; Einstein et al., 2005; Horn, Bayen, & Smith, 2011; Smith, 2003). Therefore, in everyday tasks, people tend to rely more routinely on spontaneous retrieval processes which facilitate PM performance (and reduce PM failures). These last processes are often studied in the laboratory through focal PM procedures (McDaniel & Einstein, 2007). The present work aimed to study the animacy effect in focal PM tasks. In a series of five PM studies, animate (e.g., *monkey*) and inanimate (e.g., *phone*) target words were used. Given the fitness-relevance of animates and our most common reliance on focal PM tasks, a PM animacy advantage would be expected in this type of tasks as well.

Other elements support this same prediction. For example, the performance between focal and nonfocal PM tasks have been reported to be correlated (Zuber et al., 2016, $r = .36$, $p < .001$). Moreover, the correlation between PM and free recall (a type of memory in which the animacy effect has been obtained reliably, e.g., Nairne et al., 2013) tends to be positive. Yet, while some studies reported a significant, although weak, correlation of both focal and nonfocal PM performance with free recall (Harris & Menzies, 1999; Zuber et al., 2016; all rs $\leq .21$, highest $p < .05$), other studies found no such association (McDaniel & Einstein, 1993).[2] Additionally, both types of PM tasks load into related latent constructs; for example, they both rely on executive functioning and controlled attention, although each one relating to specific executive functions: while focal tasks are more related to inhibition, nonfocal tasks relate to updating and shifting (Zuber et al., 2016).

Within the domain of proximate mechanisms of the animacy effect in retrospective memory – those that focus on explaining *how* the animacy effect occurs (in complement to *why* it occurs) – two have gathered stronger empirical support: the attention-prioritization and the richness of encoding accounts. The *attention-prioritization account* states that

[1] Other classifications to "focal/nonfocal" tasks have been proposed. Some examples are the "task appropriate/inappropriate processing" (Maylor, 1996), the "bottom-up/top-down" processes (McDaniel et al., 2015), "automatic/controlled" processes (Cona et al., 2016; Gilbert et al., 2013), "noncompetitive/competitive" processes used to perform the ongoing and the PM tasks (Gilbert et al., 2013). We opted to use the focal/nonfocal terminology as it aligns with the studies that inspired our experiments, our previous work on nonfocal PM, and the literature used to discuss our results.

[2] Despite the positive and significant correlation between the performance in PM and retrospective memory tasks, it is important to note that these memory types differ in many aspects. For example, while the former are characterized by a self-initiated retrieval of intentions (Craik, 1986), the later place participants in a retrieval mode (McBride & Workman, 2017).

animates capture attention more automatically than inanimates (e.g., New et al., 2007; but see Hagen & Laeng, 2016). Prior studies have assumed that, if animates are processed more automatically than inanimates, then, the animacy effect should be larger under divided attention conditions than in full attention ones. However, contrary to expectations, larger animacy advantages have been obtained in full (*vs.* divided) attention conditions (Leding, 2019; Rawlinson & Kelley, 2021). Transposing those findings to the present work: as animates afford cognitive advantages, including in less demanding tasks, one would expect to obtain an animacy advantage when animates (*vs.* inanimates) are presented as the PM targets. According to the *richness of encoding account,* animates have more associates in one's mental lexicon, and allow the generation of more related ideas than inanimates (e.g., Rawlinson & Kelley, 2021; but see Komar et al., 2024). Therefore, animates (*vs.* inanimates) are assumed to form narrower semantic clusters (Räling et al., 2017; Serra & DeYoung, 2023; Wulff et al., 2022); that would enhance the generation of ideas and create stronger associations among animate concepts, as the semantic memory nodes associated with each animate would be closer to each other. The same occurs with typical (*vs.* atypical) elements of a category (e.g., *apple* is a typical element of the category 'fruits', while *date* is an atypical one). Indeed, typical targets improved PM performance in focal tasks, compared to atypical targets (Nowinski & Dismukes, 2005; Penningroth, 2005). In the same vein, one could predict better PM performance for animate (*vs.* inanimate) PM targets, although prior studies have suggested that animacy and typicality effects occur independently and do not interact with each other (Penningroth, 2005; Räling et al., 2017).

No strong predictions were made about the animacy effect for the baseline or the filler trials. However, as it has been suggested that animates capture attention more automatically than inanimates (e.g., Bugaiska et al., 2019), the former could "absorb" participants' attention during the ongoing task (Zuber et al., 2016). Therefore, one could anticipate a decline in the performance of the ongoing task (baseline and filler trials) and/or longer response times (RTs) in the animate compared to the inanimate item trials (Horn, Bayen, Smith, et al., 2011).

Study 1: Lexical decision ongoing task

In this Study, participants performed a lexical decision task (LDT), which has been commonly used in focal PM studies as the ongoing task (e.g., Gilbert, 2015; Scullin et al., 2010; Smith, 2003). For a set of words (PM targets), participants were instructed to perform an alternative task (i.e., the PM task). Importantly, one of those targets was an animate and the other an inanimate, allowing us to explore the potential influence of animacy in PM performance.

Method

Participants

Despite the long-term discussion on how to interpret effect sizes as either *small, medium,* or *large,* it is usually accepted that effect sizes less than *small-to-medium* are likely to be questioned for theoretical, practical, and explanatory purposes (Funder & Ozer, 2019). Thus, all the present studies comprised samples sizes able to detect, at least, a small-to-medium animacy effect size. Using G*Power 3.1.9.4 (Faul et al., 2007), considering an $\alpha = .05$ (two-tailed), power = .85, and aiming to obtain a small-to-medium effect size in the PM target trials ($dz = 0.35$), N was set to at least 76 participants.

Following the recommendations for high-quality data collection on MTurk (https://www.mturk.com/; Cobanoglu et al., 2021; Hauser et al., 2019), the following pre-screeners were used: English as first language, approval rate > 95 %, age 18–35 years old, and USA location. A total of 114 participants responded to the study in exchange for $1.70 for a 15-minute participation. From those, 32 were excluded because they: did not perform any PM response ($n = 14$); asked their data to be deleted / did not pay attention to the task ($n = 5$); did not encode the PM target (i.e., did not recognize at least one of the PM targets and did not provide any PM response to that target; $n = 7$); did not meet the MTurk requirements to obtain compensation ($n = 2$); provided no responses to more than 50 % of the trials ($n = 1$); had low performance in the ongoing task (below the Grand Mean − 3 *SD*; $n = 2$); or due to a technical problem ($n = 1$). The final sample included 82 English native speakers (Mean age = 25.91 years; *SD* = 3.03; 52.4 % females, 47.6 % males). Each version of the task included 20 to 21 participants.

Materials

The stimuli and their characterization are available in SM 1. A total of 24 words (half animate, half inanimate) were selected from VanArsdall and Blunt (2022). Four of them (two animates, two inanimates) were used in practice trials, while the remaining 20 were used in the PM phase (filler trials). Additionally, two pairs of words (composed of an animate and an inanimate) were selected to be used as the PM targets (target trials). Each set of PM targets (*sister/wallet; canary/lock*) were presented to different participants for the sake of generalization of the results. All animate and inanimate words were matched along several variables (e.g., concreteness, age of acquisition; see SM 1). For each word (except for the target words), a pseudoword was generated through Wuggy (Keuleers & Brysbaert, 2010), following the English orthographic and phonotactic rules. The two types of pseudowords (i.e., those generated from animate and inanimate words) were also matched on several variables (see SM 1).

Procedure

In all Studies, data were collected online, through Qualtrics, unless otherwise specified. After consenting to participate in the study, participants were presented with the instructions for the LDT: a string of letters was shown one at a time; the participants' task was to decide whether each string represented a word (e.g., *wind*) or not (e.g., *wike*), by clicking on the M or Z keyboard keys. The response keys for the LDT were counterbalanced across different versions of the task (Z for word, M for pseudoword, and vice-versa; e.g., Scullin et al., 2010; Smith, 2010). Each participant was presented with one of the two sets of PM targets (*sister/wallet*; *canary/lock*); this produced four different versions of the task. Participants were randomly assigned to one of these versions.

Participants started by responding to eight practice trials to get acquainted with the task (cf. Radanović et al., 2016; Smith et al., 2007). Each practice stimulus was presented once, in a random order, for a maximum of 2500 ms or until a response was given. Participants received feedback for their accuracy (the words "correct" or "incorrect" were presented) and speed (the message "please, respond faster" was displayed if no response was provided within the 2500 ms); the feedback message remained on the screen for 1000 ms (as in Bowden et al., 2021).

After the practice phase, the PM instructions were presented. Participants were instructed that, in addition to the lexical decision task, they should perform another task (the PM task): whenever any of two specific words were presented, they should press the SPACEBAR (instead of the M/Z keys); otherwise, they should perform the LDT. They were also informed that they would no longer receive feedback on their responses. Then, they were presented the two target words (an animate and an inanimate) for one minute and asked to memorize them.

To prevent participants from rehearsing the PM instructions, a 2-minute distractor task followed this memorization period: a 3D mental rotation task (Ganis & Kievit, 2015). Two images were presented side-by-side, for 7500 ms; participants had to decide if the images were equal (only differing in their rotation angle) or if they were different images. Responses were provided by selecting either the "yes" or "no" option displayed on the screen.

Next, the PM phase started without further mentioning the PM instructions. This phase contained 84 trials: 20 filler words (10 animates and 10 inanimates), 20 filler pseudowords (generated from the filler words) and two target words (an animate and an inanimate). All stimuli were presented twice, in a random order, with the only constraint that each word and pseudoword appeared once in each half of the PM phase. By presenting each stimulus twice (both fillers and targets) we ensured that the target words were not perceived as different or discrepant (because of their repeated presentation) relative to the filler words (Whittlesea & Williams, 2001). The position of the target trials was fixed for all participants, as, for example, in Smith (2010); these occupied the 15^{th}, 35^{th}, 55^{th}, and 75^{th} positions. In each of these positions, an animate or an inanimate target was randomly presented (with the constraint that an animate and an inanimate target were presented in each half of the task). After a fixation cross (300 ms), each stimulus was displayed in lowercase on the screen until a response was given or until 2500 ms had elapsed (as in Matos et al., 2020; Radanović et al., 2016). The PM target trials represented 4.8 % of the total trials, as in previous studies (Scullin et al., 2012; Smith, 2010).

Upon completing the PM phase, participants recalled the instructions they were provided for the task. A PM target recognition test followed; one at a time, and in a random order, both targets and four lures (half animates, half inanimates; from the PM phase) were presented in a random order to each participant. Participants were asked to decide whether each word corresponded to one of the PM targets or not (yes/no response).

Participants also provided sociodemographic information (age, gender, and native language) for sample characterization purposes. Also, they responded to the "honesty questions" regarding whether they paid attention and responded honestly to the task (Rouse, 2015), and provided optional feedback about the task. Finally, participants were thanked and debriefed. The experiment took about 15 minutes.

Data analyses (studies 1–5)

Unless otherwise specified, in all Studies 1 to 5, we proceeded as follows: the trials immediately after the PM target trials were excluded from the data analyses (e.g., Matos et al., 2020; Strickland et al., 2020); missing responses in all types of trials were scored as incorrect responses (as in Scullin et al., 2010); and, data analyses were conducted using SPSS 25 (IBM, 2017). Performance was analyzed using repeated measures ANOVAs, with Animacy (animate *vs.* inanimate words) and Type of Trial as within-subject variables. The Greenhouse-Geisser correction was used when applicable and multiple comparisons were performed along with the Bonferroni correction. Additionally, Bayesian analyses were conducted on JASP 0.16.3 (JASP Team, 2022) to provide further evidence for the null and alternative hypotheses (Dienes, 2014).[3] The results of the final target recognition task implemented in Studies 1–5 are presented in SM 2. Also, as usual, the performance in focal PM procedures was very high (sometimes at ceiling levels; e.g., Rummel et al., 2017) which might have prevented the effect of the animacy variable from standing out. To explore this possibility, we considered the performance on the filler trials to split the sample into high and low performing groups based on the median value (i.e., median-split); such variable (performing group: high *vs.* low) was then included in a mixed ANOVA as a between-subjects variable. More detailed arguments for these analyses and the corresponding results from each experiment (Studies 1–5) are reported in SM 2. The analyses on the excluded participants and RTs are available in SM 4 and SM 5, respectively.

Results

The main dependent variable was the proportion of correct responses (i.e., press the correct key — Z/M for the filler trials and SPACEBAR for the PM trials). As our main interest concerned the animate/inanimate status of the *words*, the analyses included only word trials (as in Rummel et al., 2017).[4] Fig. 1 depicts the results. A 2 (Animacy: animate *vs.* inanimate words) X 2 (Type of Trial: filler *vs.* target) repeated measures ANOVA was performed. The only significant main effect was the Type of Trial, $F(1, 81) = 43.36$, $p < .001$, $\eta_p^2 = .349$, $BF_{Inclusion} = 1.87x10^6$, revealing that the performance on the target trials was significantly lower than that obtained for the filler trials. Both the Animacy main effect, $F(1, 81) = 1.37$, $p = .242$, $BF_{Inclusion} = 0.18$, and the Animacy X Type of Trial interaction were non-significant, $F(1, 81) = 0.11$, $p = .746$, $BF_{Inclusion} = 0.15$.

Interim discussion

Contrary to our hypothesis, no animacy effect was found in the PM target trials in this focal procedure. Said finding was further corroborated by the Bayesian statistic. Additionally, no influence of animacy was found in the LDT performance (nor in RTs; see SM 5). Previous studies have also not found an effect of animacy in lexical decision latencies (Radanović et al., 2016; but see Bonin et al., 2019). The results of the median-split analysis based on performance, reported in SM 3, followed the results just described.

As stated above, no animacy effect was obtained in this task, which used a LDT as the ongoing task. However, considering that LDTs rely more on lexical/orthographic processing (Izura & Hernández-Muñoz, 2017), one could argue that a semantic-related ongoing task would be more appropriate for investigating the influence of animacy (a semantic variable) in PM. Therefore, Studies 2 and 3 employed semantic-related ongoing tasks.

Study 2: Anagram-solving ongoing task

This Study employed an anagram-solving ongoing task, as previously used in other focal PM studies (Lee & McDaniel, 2013). In this Study, we used the list of words used in Félix et al.'s Study 1b (2024), in which a PM animacy effect was found using a nonfocal PM task.

Method

Participants

A sample size of at least 81 participants was pre-registered, calculated through G*Power 3.1.9.4 (Faul et al., 2007), with $\alpha = .05$, power $(1-\beta) = 0.85$, aiming to obtain a small-medium effect size in the target trials ($dz = 0.34$). Participants were recruited at the University of Aveiro ($N = 136$). They participated in-person, in exchange for course credits or for entering a prize draw. Forty-eight participants were excluded because they: did not provide any PM response ($n = 27$); were non-European-Portuguese native speakers ($n = 9$); their performance in the ongoing task (filler trials) was lower than the Grand Mean − 3 *SD* ($n = 6$); failed to recognize the PM targets and to provide PM responses to those targets ($n = 2$); did not respond to 50 % or more of the trials ($n = 1$); or indicated their data should be excluded / did not pay attention to the task ($n = 1$). Two additional participants were excluded: one due to a technical problem which prevented data registration, and another because he/she was non-naïve to the aims of the study. Even though we pre-registered the exclusion of participants aged below 18 years old, we included two participants aged 17 years old, after obtaining proper

[3] We report $BF_{Inclusion}$ for quantifying the strength of the evidence of a particular effect (Rouder et al., 2017). $BF_{Inclusion}$ assumes values above or below 0, indicating evidence for the inclusion or not of a variable in the model, respectively. *T*-tests report BF_{10}, providing evidence for the likelihood of a specific result under the alternative (H1) or the null hypothesis (H0; Wagenmakers et al., 2011). For example, $BF_{10} \geq 3$ indicates that the results are at least three times more likely under H1 than under H0, while $BF_{10} \leq 1/3$ provides evidence in favour of the null hypothesis (Wagenmakers et al., 2011, 2018).

[4] Of note, the mean proportion of correct responses for the pseudowords, created directly from the animate and inanimate words, was, in both cases, of 0.97 (*SEM* = 0.01).

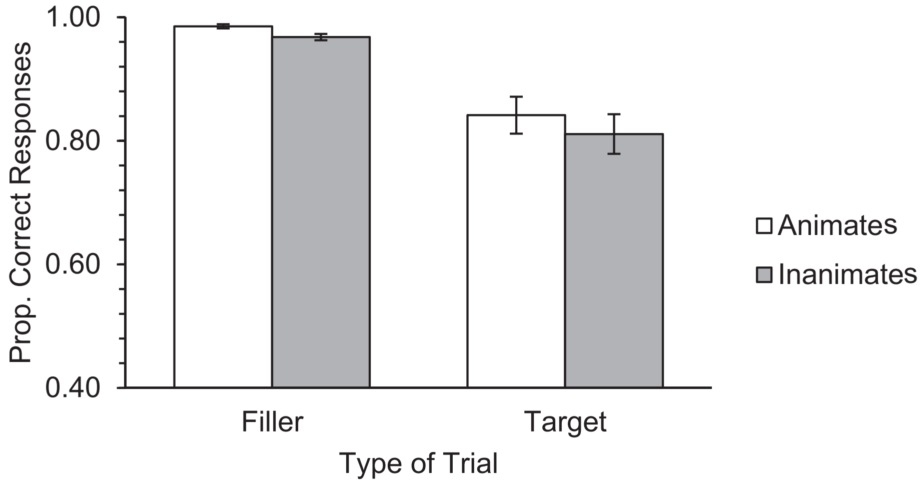

Fig. 1. Mean Performance Obtained in the Word Filler and Target Trials (PM Task) in Study 1. Error Bars Represent Standard Errors of the Mean.

written consent from their tutors as well as from the participants themselves. The final sample was composed of 88 European Portuguese native speakers (Mean age = 21.01 years; *SD* = 4.48; 81.8 % females and 18.2 % males). Twenty to 24 participants responded to each version of the task.

Materials

As mentioned before, in this experiment we used the same material as in Study 1b from Félix et al. (2024). Two words were selected for the practice phase, and 16 other words for the baseline phase. For the PM phase, another 24 filler words and two pairs of target words were used. Words (half animate, half inanimate, in all phases) were selected from Félix et al. (2020) and equated along several variables (available in SM 1). For the present Study, words (e.g., *pomba* [dove]) were converted into anagrams by changing the order of two adjacent letters in each word (e.g., *pobma*), as in Lee and McDaniel (2013). Two different anagrams were created for each word (by switching two different letters) to allow presentation of the word twice during the PM phase without repeating the corresponding anagram.

Procedure

After consenting to participate, the ongoing task instructions were presented; participants were tasked to solve anagrams (e.g., POBMA → POMBA [*dove*]), type the response, and then press ENTER to move on to the next anagram. Anagrams were presented one at a time, for 15 s or until a response was given.

After two practice trials, participants performed 16 baseline trials (with the anagram-solving task only). Then, they were presented the PM instructions: in addition to performing the anagram-solving task, they were instructed to press the SPACEBAR whenever the anagram of either two specific targets (an animate and an inanimate) was presented (instead of solving the anagram; PM response). Then, they were presented the two target words for one minute and asked to memorize them. As in Study 1, a 2-minute 3D mental rotation distractor task followed.

Then, the PM phase began without further mentioning the PM instruction. In this phase, each word was presented twice, each time in a different anagram, totalizing 52 trials (48 filler and 4 target trials). The PM targets were presented in trials 11th, 24th, 36th, and 51st, as in Study 1b from Félix et al. (2024). Words in the practice and baseline phases were presented in the same/fixed order for all participants. There were two predetermined presentation orders in the PM phase to ensure that, in each position of the list, an animate and an inanimate item was presented an equal number of times across participants (both in filler and target trials). We also used one out of two sets of PM targets (*atleta [athlete]* and *camisa [shirt]* / *cavalo [horse]* and *janela [window]*) in each of these versions. Therefore, there were four versions of the task to which participants were randomly assigned. Finally, as in Study 1, participants responded to the instructions recall task, the PM target recognition test, a sociodemographic questionnaire, honesty questions, and provided optional feedback on the study. Then, they were thanked and debriefed. The experiment lasted approximately 16 minutes.

Results

Fig. 2 shows the obtained results on the proportion of correct responses. For the baseline and filler trials, this corresponds to the correctly solved anagrams, whereas for the target trials it refers to the proportion of correct PM responses. A 2 (Animacy: animates *vs.* inanimates) X 3 (Type of Trial: baseline *vs.* filler *vs.* target) repeated measures ANOVA was conducted. Only the Type of Trial main effect reached significance, $F(1.20, 104.53) = 8.41$, $p = .003$, $\eta_p^2 = .088$, $BF_{Inclusion} = 51.12$. Follow-up paired *t*-tests, conducted along with the Bonferroni correction ($p < .05/3 \sim .0167$) revealed that the performance in the filler trials was significantly higher than that obtained in the baseline, $t(87) = 6.11$, $p < .001$, $dz = 0.65$, $BF_{10} = 434744.57$, and target trials, $t(87) = 3.40$, $p = .001$, $dz = 0.36$, $BF_{10} = 23.14$. The difference obtained between the latter two was non-significant, $t(87) = 1.20$, $p = .235$, $BF_{10} = 0.07$. The Animacy main effect, $F(1, 87) = 2.63$, $p = .109$, $BF_{Inclusion} = 0.23$, as well as the interaction were non-significant, $F(1.45, 126.43) = 0.78$, $p = .424$, $BF_{Inclusion} = 0.08$.

Interim discussion

Once again, the animacy effect was not obtained in a focal PM task, this time using a semantic-related ongoing task. The average performance on the baseline and target trials was similar to that obtained in prior studies, but the performance in the filler trials was higher (e.g., Lee & McDaniel, 2013). The median-split analysis revealed the same pattern of results as reported above (see SM 3).

Study 3: Word-rating ongoing task

Studies 1 and 2 failed to obtain a PM animacy advantage. In Study 1, we used a LDT task which focused more on lexical aspects of the words. The anagram-solving task used in Study 2 required semantic processing, but other non-semantic variables might actually be more prominent in such task (e.g., phonemic and mental spatial manipulation of letter position; Nioka et al., 2008). Therefore, in Study 3, we further explored the animacy effect in PM using another semantic ongoing task frequently employed in focal PM studies: a word-rating task (e.g., Kliegel et al., 2007, 2004; McDaniel et al., 2004).

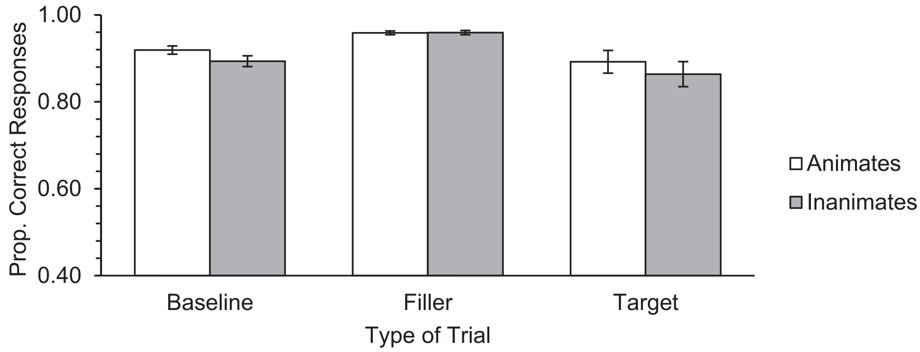

Fig. 2. Mean Performance Obtained in the Baseline, Filler (Ongoing Task), and Target Trials (PM Task), in Study 2. Error Bars Represent Standard Errors of the Mean.

Method

Participants

Sample size was calculated as in Study 2. A total of 126 students were recruited from several Portuguese universities. From those, 42 were excluded because they: provided no PM responses (n = 15); had participated previously in other PM tasks in our laboratory (n = 9); were not European Portuguese native speakers (n = 9); asked their data to be deleted / did not pay attention to the task (n = 4); failed to recognize the PM targets and to provide PM responses to those targets (n = 2); failed both attention checks (n = 1); or did not respond to 50 % or more of the trials (n = 2). The final sample included 84 European Portuguese native speakers (Mean age = 23.86 years; SD = 7.68; 60.7 % females, 39.3 % males; although one participant did not disclose his/her age we included him/her in the data analysis, as the study was shared among university students who are typically aged between 18 and 35 years old). Each version of the task was responded to by 20 to 23 participants.

Materials

Twenty-four words (12 animates and 12 inanimates) were selected from Félix et al. (2020). Considering that the ongoing task was a word-rating task (pleasantness and subjective frequency rating tasks), we selected words with low, medium and high pleasantness mean ratings (as per data collected in a pleasantness-rating task reported in Félix & Pandeirada, 2024), as well as with low, medium and high subjective frequency (Soares et al., 2017). Additionally, two pairs of PM target words (each including one animate and one inanimate) were selected to be used in each version of the task. Two additional words (an animate and an inanimate) were used for the practice trials. Words were equated on several variables (see SM 1).

Procedure

Data collection was held both online (n = 3) and in-person (n = 81)[5] using LimeSurveyUA. After consenting to participate, the instructions for the word-rating ongoing task were presented: words were presented, one at a time for 5000 ms, and the participants' task was to rate each word in terms of pleasantness or subjective frequency, using 5-point rating scales. The trials for pleasantness ratings and subjective frequency ratings were intermixed and were identified by a prompt (see Supplemental Fig. 1 for an illustration of each of those trials).

The Portuguese word-rating instructions, created from preexisting ones (Nairne et al., 2007; Soares et al., 2017), were as follows (English translation). SM 6 presents the complete rating instructions used in this Study (in European Portuguese).

Pleasantness: Each word can describe things that are pleasant or unpleasant for you. You should rate each word according to how the word makes you feel. The rating will be made on a scale ranging from 1 (very unpleasant) to 5 (very pleasant).

(Subjective) Frequency: The number of times we are confronted with a given word in our daily lives varies from word to word. You should rate each word in relation to the number of times you read, use, or hear it in your daily life. The rating will be made on a scale ranging from 1 (I have never come into contact with this word) to 5 (I encounter this word several times a day).

The ratings were given by pressing the "1" to "5" number keys on the keyboard. As in Kliegel et al. (2007), participants were asked to provide a fast, honest, and intuitive rating; additionally, instructions stated that there were no right or wrong answers, and encouraged participants to try to use all values in the scale.

Participants began by responding to four practice trials which ran as described. Also, words were presented twice, each time being rated on a different dimension (pleasantness or subjective frequency).

Next, the PM instructions were presented, asking participants to provide a PM response (i.e., to press "0" instead of providing the word-rating) whenever any of the target words was presented during the word-rating task; for the remaining (filler) words, participants should perform the word-rating task. Then the two PM target words (one animate, one inanimate) were presented for one minute, and participants were asked to memorize them. As in Study 1, the 2-minute 3D mental rotation distraction task followed.

The PM phase then started without further mentioning the PM instructions. Words were presented in the same order for every participant (pseudo-randomized order), ensuring a balanced number of animate and inanimate words, and of rating dimension (pleasantness / subjective frequency), in each quarter of the list. As stated before, trials presenting each rating dimension were intermixed throughout the task. Every word was presented twice, once in either half of the list, each time in a different to-be-rated dimension, totalizing 52 trials. As in Kliegel et al. (2007), there were at least 10 trials between the repetition of a given word. The target trials were in the 5th, 25th, 37th, and 51st positions (one target in each quarter of the list). A comparable proportion of target trials (~8 %) was used in similar studies (e.g., Kliegel et al., 2007).

In this Study, aiming to gather more insight about the data collected online, an attention check was presented after the practice trials ("Have you ever walked on Mars?" – yes/no response), and another after the PM task ("Indicate which is the second word of this sentence: How many colors are there in the Portuguese flag?") (Rouse, 2015). Upon completing the task, participants underwent the same general procedure as in Studies 1 and 2 and reported whether they had participated in previous PM studies from our lab. Finally, participants were thanked and debriefed. The experiment lasted approximately 26 minutes.

[5] A sample of at least 81 participants was calculated *a priori*. Still, beyond the participants recruited in-person (n = 81), we opted to also include participants recruited online (n = 3). We re-ran all the analyses excluding those 3 online participants, and the same results were found as when including them: Subjective frequency ratings (animates > inanimates) $t(80) = 4.20$, $p < .001$, dz = 0.47, BF_{10} = 275.65; Pleasantness ratings (animates < inanimates) $t(80) = -4.42$, $p < .001$, $dz = -0.49$, BF_{10} = 588.42; PM performance (animates ~ inanimates) $t(80) = 0.90$, $p = .372$, dz = 0.10, BF_{10} = 0.18.

Results[6]

Trials with missing responses were not considered in the data analysis (more information in SM 2). Paired *t*-tests were conducted to compare the subjective frequency and pleasantness ratings for the filler trials, as well as PM performance for the target trials, between animates and inanimates. Animate words received, on average, higher subjective frequency ratings ($M = 2.82$; $SD = 0.47$) than inanimate words ($M = 2.65$; $SD = 0.42$), $t(83) = 4.00$, $p < .001$, $dz = 0.44$, $BF_{10} = 147.82$. The opposite pattern was obtained for pleasantness, with animates ($M = 3.22$; $SD = 0.32$) rated, on average, as less pleasant than inanimates ($M = 3.41$; $SD = 0.38$), $t(83) = -4.39$, $p < .001$, $dz = -0.48$, $BF_{10} = 535.28$. Regarding the PM trials, performance was very high for both types of stimuli (animates: $M = 0.94$; $SD = 0.18$; inanimates: $M = 0.91$; $SD = 0.26$), with the small descriptive difference not being statistically significant, $t(83) = 0.90$, $p = .372$, $BF_{10} = 0.18$.

Interim discussion

We obtained unexpected rating differences between animates and inanimates given they were matched *a priori* on the rated dimensions (see SM 1). Such differences might be related to an unbalance on the type of rating that items received in their first and second presentation. Specifically, animates were more likely to be rated on pleasantness in their first presentation, likely increasing their sense of familiarity on their second presentation (and the opposite for the inanimate items). However, we do not anticipate this issue to impact our results of main interest.

In this Study, using a task that more directly appealed to semantic processing, we also did not obtain an animacy advantage in focal PM, contrary to our prior findings using nonfocal tasks (Félix et al., 2024). This pattern of results was also obtained in the median-split analyses (SM 3).

Looking at the overall pattern of results, we attempted to identify aspects of these focal tasks that differ from those of the nonfocal tasks in which we obtained the effect. Besides the focality of the task, the just presented studies differ from those employed by Félix et al. (2024) on the type of ongoing task. In those studies, the ongoing tasks involved, to some degree, a working-memory component which is cognitively more demanding than the verbal-related tasks used in the current studies. Also, the high levels of performance obtained in the current tasks suggest they were easy to solve, allowing effortless processing of both animate and inanimate PM targets, thus reducing the potential for animacy to play an effect. Moreover, prior research has revealed that attentional and working-memory resources are recruited differently in focal and nonfocal PM tasks (Brewer et al., 2010). To address this possibility, Studies 4 and 5 were conducted using a different ongoing task.

Study 4: 2-back ongoing task

In Studies 4 and 5 we used PM procedures with working-memory-related ongoing tasks. If the animacy effect in PM depends on using a more demanding ongoing task (verbal- *vs.* working-memory-related), one would expect a significant animacy effect in PM target trials in these Studies. This Study employed a 2-back ongoing task (Kliegel & Jäger, 2006) which complies with such criteria.

Method

Participants

Sample size was calculated as in Study 1. A total of 115 participants were recruited online: 42 via Prolific (https://www.prolific.com/, paid £3.00 for a 20-minute participation), and 73 through the City, University of London (currently, City-St. George's, University of London) SONA system (in exchange for course credits). The Prolific pre-screeners used were: location (UK), age (18–35 years old), being currently a student (for comparability with the participants from SONA), and approval rate (95–100 %).

Thirty-seven participants were excluded from data analysis because they: did not recall the task instructions or the PM response key, or did not recognize the targets and did not perform any PM response to the non-recognized targets ($n = 20$, 12 of those from SONA); asked their data to be excluded / did not pay attention to the task ($n = 9$, eight of them from SONA); did not respond to more than 50 % of the trials ($n = 4$, all from SONA); had a performance in the ongoing task (filler trials) lower than the Grand Mean − 3 *SD* ($n = 3$, all from SONA); or failed both attention checks ($n = 1$ from SONA). The final sample included 78 participants, all fluent English speakers (33 of them recruited on Prolific; Mean age = 21.72 years; $SD = 4.92$; 67.9 % identified themselves as females, 29.5 % as males, and 2.6 % preferred not to answer to this question). Sixty-seven participants were English native speakers or English bilinguals; the remaining 11 participants were non-English natives or preferred not to disclose their native language, but all reported being highly fluent in English (answering 4 or 5, in a 5-point scale, ranging from "1 – Not fluent at all" to "5 – Very fluent"). Each version of the task was responded to by 18 to 21 participants.

Materials

Words (half animates, half inanimates) were selected from VanArsdall and Blunt (2022), and matched on several variables (see SM 1). Six words were selected for the practice trials, 38 for the baseline, 66 for the filler trials, and two pairs to be used as targets.

Procedure

After consenting to participate, participants were provided the 2-back task instructions. One at a time, words were presented, and the participants' task was to decide whether the presented word was (or not) the same as that presented two words before, by pressing the Y (for *Yes*) or the N (for *No*) keys. An image illustrating the task was also provided with the instructions. Participants were instructed to respond as fast and accurately as possible. Throughout the task, each word was presented for 2000 ms, or until a response was provided (Schnitzspahn et al., 2014). As in similar research (e.g., Schnitzspahn et al., 2022, 2014), there were 20 % hits (i.e., *Yes* responses) in each phase of the study (practice, baseline and PM phase). Word presentation was previously pseudo-randomized, ensuring that, in each phase, each quarter of the task contained a balanced number of animate and inanimate words (and a balanced number of hit-responses for animates and inanimates).

Following similar research (e.g., Schnitzspahn et al., 2022, 2014), participants performed 10 practice trials to get acquainted with the task. Also, they were instructed to respond correctly on at least 80 % of the trials to move on to the task; otherwise, they re-read the instructions and repeated the practice phase once again (about 7 % of participants performed < 80 % in the second repetition; still they were not excluded as their performance in the PM phase did not motivate their exclusion). Furthermore, participants received feedback on each practice trial: if a correct response was given within the 2000 ms display time, a green check was presented; otherwise, a red cross was displayed (for 100 ms). At the end of the practice phase, a message showed up, indicating if the minimum performance of 80 % had been achieved. This procedure ensured participants understood the instructions, were engaged, and complied with the task.

Then, a reminder of the instructions was presented, and participants

[6] We pre-registered a repeated measures ANOVA (Animacy X Type of Trial) on the proportion of correct responses as the dependent variable. Nonetheless, as this was a word-rating ongoing task, there were no correct/incorrect responses for the filler trials. Alternatively, we conducted *t*-tests on the ratings to explore unanticipated differences between animates and inanimates.

performed the baseline phase with the 2-back task only (60 trials, as in related research, Ballhausen et al., 2017; Schnitzspahn et al., 2022, 2014). During this phase no feedback was provided.

Next, the PM instructions were presented, indicating that, in addition to the 2-back task, they should perform another task: to press the SPACEBAR (instead of Y/N) whenever either of the two target words was presented. An image illustrating the task was also presented. Then, participants were instructed to memorize the two target words (one animate, one inanimate). Each target word was presented for 5000 ms. To check whether the targets were correctly encoded, participants were asked to type the target words they were just asked to memorize.[7] If recall of the targets was incorrect, the target words were again displayed for 5000 ms each, and the recall task was repeated (as in Horn & Bayen, 2015). If incorrect responses were given again, a new presentation of the target words occurred for one last opportunity for memorization, and no recall followed. No participants needed a second presentation of the targets.

To prevent participants from rehearsing the PM instructions, they performed two distractor tasks: a spot the five differences task (six pairs of images were presented for 30 s each, totalizing 3 min), and a 2-minute 3D mental rotation task, as described in Study 1. In this Study we opted to increase the distractor period aiming to reduce the high performance levels.

The PM phase followed without further mentioning the PM task instructions. In this phase, there were 100 trials (96 filler and 4 target trials), maintaining a similar proportion of target trials as in previous research (Schnitzspahn et al., 2022, 2014). There were two target words (an animate and an inanimate, presented twice). The PM target trials were presented in the 24th, 47th, 75th, and 96th trials. There were two pairs of PM targets (*monkey* and *phone*; *frog* and *paper*). There were two predetermined presentation orders of the PM targets to ensure that, in each target-position of the list, an animate and an inanimate item was presented. Therefore, there were four versions of the task to which participants were randomly assigned. The filler trials were presented in the same order for all participants. The animate and inanimate hit trials (i.e., those with a correct Yes response) were preceded and followed by an animate and an inanimate filler word an equal number of times. There was roughly the same number of hit animate/inanimate trials in each quarter of the task (both in the baseline and PM phases). In the baseline and filler trials, 22 and 30 words, respectively, appeared twice. Half of these appeared in hit trials and half in correct rejection trials, preventing hits from occurring only for repeated words.

An attention check was presented after the baseline phase ("Can you fly with invisible wings?" – press Y/N for Yes/No), and another after the PM phase ("How many triangles can you see in this figure?").

After the PM phase, participants recalled the PM instructions, followed by the target recognition test. Then, they were asked to recall the PM response key. Participants also provided sociodemographic information (age, sex, native language; non-English native speakers were also asked to indicate how fluent they were in English), responded to the honesty questions, and were given the possibility of providing feedback about the task. Finally, they were thanked and debriefed (total duration of the procedure: approximately 20 minutes).

Results

Fig. 3 shows the results. A 2 (Animacy: animate *vs.* inanimate words) X 3 (Type of Trial: baseline *vs.* filler *vs.* target) repeated measures ANOVA was conducted. Only the Type of Trial main effect was significant, $F(1.13, 86.91) = 16.89$, $p < .001$, $\eta_p^2 = .180$, $BF_{Inclusion} = 24546.76$. Follow-up paired *t*-tests revealed that the performance in the target trials was lower than in the baseline, $t(77) = -4.49$, $p < .001$, dz = -0.51, $BF_{10} = 1608.01$, or the filler trials, $t(77) = -3.84$, $p < .001$, $dz = -0.44$, $BF_{10} = 155.04$; with no differences between the latter two types of trials, $t(77) = 2.08$, $p = .041$, $BF_{10} = 1.72$ (following the Bonferroni correction). Both the Animacy main effect, $F(1, 77) = 2.15$, $p = .147$, $BF_{Inclusion} = 0.22$, and the interaction were non-significant, $F(1.06, 81.42) = 1.54$, $p = .219$, $BF_{Inclusion} = 0.20$.[8]

Interim discussion

Again, no advantage was found for the animate target trials even when using a working-memory updating task. This task was used to explore whether our results were being influenced by the type of ongoing task (rather than the focality of the PM task) given that we found a PM animacy advantage using related tasks in nonfocal procedures. The obtained results suggest that the animacy effect in PM depends more on the overlap between the processes involved in the ongoing task and those required to detect the PM target (i.e., focality of the task), than on the type of ongoing task used. Nonetheless, considering that performance still approached ceiling levels (limiting any meaningful conclusion; cf. median-split analyses in SM 3), we conducted Study 5, with a more demanding working-memory updating ongoing task.

Study 5: 3-back ongoing task

Aiming to lower performance levels and further test the influence of the ongoing task in our results, in this Study we employed a 3-back ongoing task, as used in previous work (e.g., West et al., 2005).

Method

Participants

Sample size was calculated as in Study 1. Participants were recruited on Testable (https://www.testable.org), using the following prescreeners: age (18–35 years old), English as first language, location (USA, UK, Australia, New Zealand, Ireland, or Canada), and approval rate (95–100 %). A total of 102 participants were recruited. From those, 26 were excluded because they: failed to recall the PM response key or to recognize the targets and did not perform any PM response to said target ($n = 21$); had an ongoing task performance (filler trials) below the Grand Mean − 3 *SD* ($n = 4$); or missed responses on more than 50 % of all trials ($n = 1$). The final sample was composed of 76 participants (75 were English native speakers or English bilinguals, another participant reported being highly fluent in English; Mean age = 28.55 years; *SD* = 5.08; 52.6 % identified themselves as females, 46.1 % as males, and 1.3 % preferred not to answer). Considering that, in Studies 1–3, including participants who failed to make any PM responses did not affect the main results (see SM 4), we opted to include those participants in the final 76-participant sample ($n = 9$).[9] Other studies also included said participants (e.g., Horn & Bayen, 2015). Each version of the task contained between 18 and 21 participants (see SM 2 for additional analyses

[7] This procedure was included in this and the next study aiming to decrease the number of excluded participants due to not recognizing the target words at the end of the experiment. A similar procedure was used in other studies that used a *n*-back ongoing task (Chen et al., 2014; Kliegel & Jäger, 2006).

[8] To explore the potential influence of sample provenience on the PM performance, we conducted a 2 (Animacy: Animate *vs.* Inanimate targets) X 2 (Group: SONA *vs.* Prolific). The results did not reveal a Group main effect, $F(1,76) = 2.17$, $p = .145$, nor Animacy X Group significant interaction, $F(1,76) = 0.28$, $p = .597$.

[9] A 2 (Animacy: animate vs. inanimate words) X 3 (Type of Trial: baseline vs. filler vs. target) repeated measures ANOVA was conducted excluding the 9 participants who did not provide any PM responses. The analysis included 67 participants. The results followed the same pattern as reported in the main text: Only the Type of Trial main effect was significant, $F(1.12, 74.00) = 13.20$, $p < .001$, $\eta_p^2 = .167$. Both the Animacy main effect, $F(1, 66) = 0.90$, $p = .346$, and the interaction were nonsignificant, $F(1.10, 72.66) = 0.74$, $p = .477$.

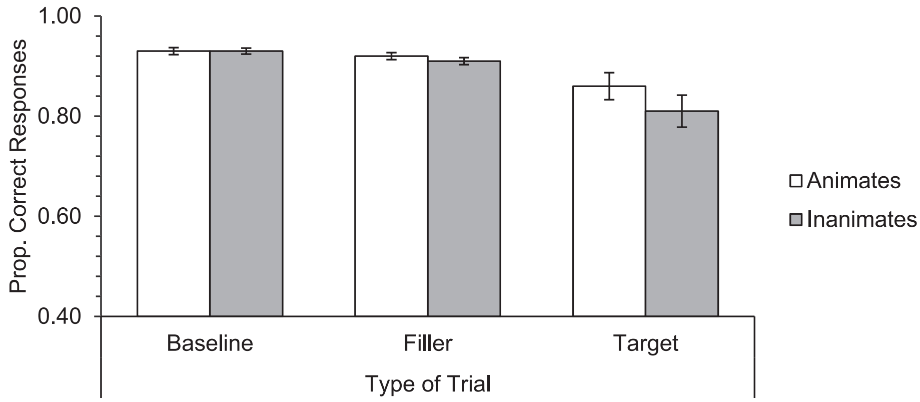

Fig. 3. Mean Performance Obtained in the Baseline, Filler (Ongoing Task), and Target Trials (PM Task), in Study 4. Error Bars Represent Standard Errors of the Mean.

regarding task version).

Materials and procedure

We used the same words and procedure as in Study 4, except that: after consenting to participate, participants performed a 3-back task and words were presented for 1500 ms or until a response was given, aiming to lower performance. Therefore, they were instructed to press Y (for *Yes*) or N (for *No*) if the presented word was the same as the one presented three words before, or not. For the PM task, in addition to the 3-back task, participants were instructed to press the SPACEBAR whenever any target word was presented. The experiment lasted about 19 minutes.

Results

Fig. 4 presents the results. As intended, the procedure used in the present Study was effective in lowering participants' performance (about 5 % to 8 % in baseline and filler trials), particularly on target trials (a difference of about 17 % as compared to the prior Study).

As in Study 4, the 2 (Animacy) X 3 (Type of Trial) repeated measures ANOVA only revealed a significant Type of Trial main effect, $F(1.09, 81.89) = 26.49$, $p < .001$, $\eta_p^2 = .261$ $BF_{Inclusion} = 5.30\text{x}10^7$. Follow-up paired *t*-tests revealed, again, that the performance in the target trials was significantly lower than in the baseline, $t(75) = -5.14$, $p < .001$, $dz = 0.59$, $BF_{10} = 2.61\text{x}10^6$, or the filler trials, $t(75) = -5.30$, $p < .001$, $dz = 0.61$, $BF_{10} = 7.54\text{x}10^6$. No difference in performance was obtained between the baseline and the filler trials, $t(75) = -1.15$, $p = .255$, $BF_{10} = 0.22$. The Animacy main effect, $F(1, 75) = 0.94$, $p = .334$, $BF_{Inclusion} = 0.13$, and the interaction were non-significant, $F(1.12, 83.70) = 0.69$, $p = .503$, $BF_{Inclusion} = 0.05$.

Interim discussion

The procedure employed in Study 5 was effective in reducing the overall levels of performance. Although including participants who made no PM responses lowered overall PM performance, the procedure was still effective in reducing performance levels even when those participants were excluded (cf. SM 4). The now obtained performance levels are closer to those obtained using nonfocal procedures in which an animacy advantage was obtained (Félix et al., 2024). Still, and despite using an ongoing task involving working memory, no animacy advantage was obtained in this focal PM study.

Overall sign-test

In the above five studies, we found no evidence for an animacy PM effect, as corroborated by Bayesian statistics. Although all experiments – and their sample sizes – had been pre-registered, we opted to conduct a sign test across all participants from Studies 1 to 5 for a more powered analyses. The animacy effect in PM target trials remained non-significant ($z = -1.73$, $p = .083$; PM performance animate > inanimate, $n = 77$; animate < inanimate, $n = 56$; animate = inanimate, $n = 275$).

General discussion

Animacy plays an important role in many cognitive domains, including memory (Nairne et al., 2013, 2017). Indeed, an animacy advantage has been reported for many mnemonic processes (e.g., free recall), in different languages, age-groups, and with a breadth of to-be-remembered stimuli (words, nonwords, and pictures), in various tasks and with multiple procedures. Recently, Félix et al. (2024) reported that this animacy advantage extends to prospective memory, when using nonfocal tasks. The aim of the present work was to explore, for the first time, the animacy effect in focal PM tasks.

In a series of five studies, using different ongoing tasks, we found no animacy effect in focal PM tasks. Bayesian statistics supported these conclusions. It is interesting that although the PM animacy effect was not statistically significant in any of the present focal PM studies, there was a descriptive tendency for a higher PM performance towards animates. However, a sign test, conducted across all participants from Studies 1 to 5 remained non-significant.

Additionally, considering that Study 2 used the same materials as in Study 1b of Félix et al. (2024) – in which an advantage was found for animate targets in a nonfocal task – it is unlikely that these results relate to the specific materials used. Therefore, the absence of the animacy advantage in focal PM tasks is more likely due to processing differences introduced by the focal/nonfocal PM procedures. Other studies have also reported differential effects of specific variables (e.g., emotional valence) in focal *vs.* nonfocal tasks (Hostler et al., 2018). These findings are informative to the overall animacy effect and also to the PM functioning itself.

As stated in the Introduction, according to the Multiprocess Framework (McDaniel & Einstein, 2000), the processes recruited to perform PM tasks depend on their focality. While focal tasks rely more on automatic processes and spontaneous identification of the targets, nonfocal tasks require more effortful processes (e.g., strategic monitoring) to detect the PM targets. Consequently, the focal PM tasks likely require less attentional resources than the nonfocal. This difference relates to one of the most prominent explanations for the animacy effect: the attention-prioritization account.

The *attention-prioritization account* suggests that animates are processed more automatically than inanimates and, thus, their attention/detection threshold is lower (Bugaiska et al., 2019; Loucks et al., 2023; New et al., 2007; but see Komar et al., 2023b; Rawlinson & Kelley, 2021). In nonfocal PM tasks, which rely more on strategic and effortful

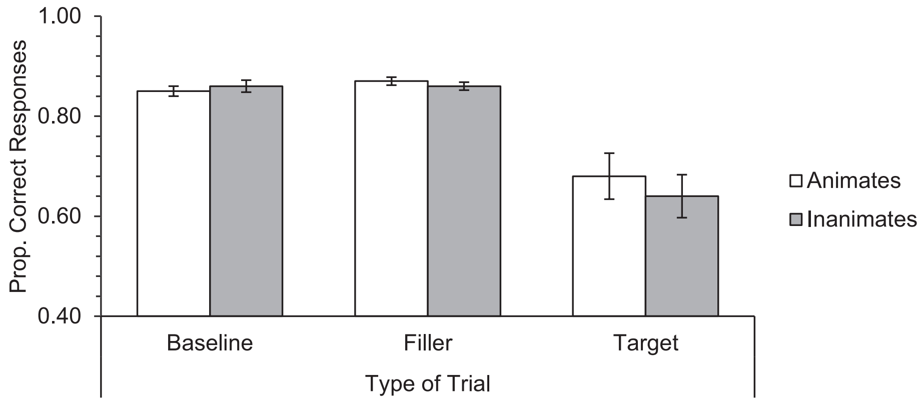

Fig. 4. Mean Performance Obtained in the Baseline, Filler (Ongoing Task), and Target Trials (PM Task), in Study 5. Error Bars Represent Standard Errors of the Mean.

retrieval, the attentional bias toward animate stimuli—processed at lower attention thresholds than inanimates—can facilitate the detection of animate PM targets. This leads to a significant animacy advantage, as reported by Félix et al. (2024). In contrast, as focal PM tasks depend more on automatic or spontaneous retrieval of the targets, PM performance overall is facilitated. The results here reported and prior research are consistent with this idea. In this context, the attentional boost associated with animacy no longer enhances performance as both types of targets are easily identified, resulting in a non-significant animacy advantage. This finding, along with the lack of an RT difference between the baseline and filler trials, lends support to the differentiation between focal and nonfocal tasks proposed by the Multiprocess Framework (McDaniel & Einstein, 2000).

The attentional account would also predict effects of the animacy variable on ongoing task performance (baseline and filler trials). In particular, the automatic attention drawn by animate stimuli may impair performance on these stimuli relative to those involving inanimates. However, across studies, no animate *vs.* inanimate differences were obtained on these trials. Response time data have also been used in the literature as an indicator of the attentional mechanisms associated with animacy. For example, the response times in a color-naming Stroop task were found to be longer when words referred to animates than to inanimates (e.g., Bugaiska et al., 2019). Although this dependent variable was not of main interest in this work, the corresponding data, available in SM 5, revealed no effect of animacy on the response times of the baseline and filler trials. In sum, although the present PM results fit the attention-prioritization account, it does not explain the data from the baseline and filler trials (see also Rawlinson & Kelley, 2021). Thus, overall, the influence of animacy in PM may not be explained solely by the attention-prioritization account.

The other proximate mechanism that has been mostly explored in the animacy effect is the *richness of encoding* account. According to this account, animates tend to be better recalled because they naturally lead to the generation of more ideas and/or have more features than inanimates (Meinhardt et al., 2020; Rawlinson & Kelley, 2021; but see Bonin et al., 2022, Komar et al., 2024). Such ideas or features potentially work as retrieval cues which would improve memory for animates, particularly in free recall tasks (but not in other tasks, such as cued-recall; Popp & Serra, 2016). This account could predict the non-occurrence of the animacy effect in focal PM tasks as well; specifically, the multiple ideas related to animate targets (*vs.* inanimates) would mislead the target-PM intention association, hindering PM performance for animate targets (McDaniel & Einstein, 2000). In this vein, Sugimori and Kusumi (2008) reported that unfamiliar targets (i.e., words with low values of familiarity and, thus, fewer associations), compared to familiar targets, improved PM performance when a spontaneous retrieval process was used (see also McDaniel & Einstein, 1993). However, considering that PM performance for inanimates (words with fewer associates) did not supplant that of animate targets, our findings are not consistent with the richness of encoding account.

The present results may also relate to research regarding the animacy effect in recognition memory. Event-based prospective memory tasks require a component of recognition memory, as participants need to recognize the targets that signal the moment to perform the PM task (McDaniel & Einstein, 2007). Previous research has reported that the animacy effect tends to be found in *remember* responses – those that rely in recollection processes – but not in *know* responses – which relate to familiarity processes (Bonin et al., 2014; Komar et al., 2023a; Rawlinson & Kelley, 2021; see also Félix & Pandeirada, 2024). Also, studies have suggested that there is an overlap between the processes (and even the neural correlates) used for *remember*-responses and performance in nonfocal PM tasks, and those used in *know*-responses and focal PM tasks (Cona et al., 2014; Curran, 2000; McDaniel et al., 2004). Our present results seem to parallel this pattern: animacy enhances memory when recollection is required but may have no (or smaller) effect when retrieval relies on familiarity.

Throughout this work, different proportions of target trials (in relation to the number of filler trials) were used. In Studies 1, 4 and 5, target trials corresponded to 4–5 % of the total trials of the PM phase, while in Studies 2 and 3 that corresponded to approximately 8 %. These resembled ratios of target/filler trials used in previous research (e.g., Scullin et al., 2012; Smith, 2010). Still, the different ratio of target/filler trials did not seem to influence the expression of the animacy effect (despite the different ongoing tasks that were used across studies). That is consistent with the idea that the ratio of target trials to filler trials does not influence greatly the performance in focal PM tasks (Anderson et al., 2019).[10]

Only some of the studies here presented included a baseline phase. Said phase is particularly relevant for exploring the costs imposed on the ongoing task by the embedded PM task. As this was not the main aim of our work, this decision allowed us to create shorter procedures aiming for more successful completion of the tasks. Nevertheless, the RT data presented in SM 5 revealed, as expected in focal procedures (e.g., Anderson et al., 2019; Scullin et al., 2010; but see Smith, 2003; Smith et al., 2007), minimal costs (measured by a slowing in RTs) between the baseline trials (ongoing task only) and the filler trials (PM task embedded in the ongoing task). Said results further attest to the focal nature of the PM tasks here reported.

In each study we only used two different PM targets (one animate and one inanimate) to prevent a high cognitive load and, consequently,

[10] A lower proportion of target trials could reduce participants' monitoring strategies, asseverating the focal nature of the present tasks; conversely, targets frequently displayed could warn/remind participants of the PM task, leading them to pay extra attention to detect the PM targets, and rely more on monitorization and rehearsal (Anderson et al., 2019).

low levels of PM performance (Anderson et al., 2019). Still, we opted to use different sets of possible targets to increase the generalizability of our results. It is also noteworthy that the present results were obtained using different ongoing tasks and in two languages (Portuguese and English), which further reinforces the consistency of the present findings.

One potential limitation of the present Studies concerns the levels of PM performance, which, as expected in focal PM tasks (cf. Loft & Humphreys, 2012), was higher than in our previous work with nonfocal tasks (Félix et al., 2024). The high PM performance (above 80 % in Studies 1–4) could be hindering an animacy advantage in PM. Even though we were able to bring performance levels down in Study 5 (PM performance ~ 66 %), an animacy advantage was still not observed. Nevertheless, other studies have reported significant effects in focal PM tasks even with high performances. For example, Sugimori and Kusumi (2008) obtained a PM increase caused by unfamiliar *vs.* familiar targets, with PM performances of 90 % and 83 %, respectively. To further explore this possibility, we conducted median-split analyses (see SM 3), comparing the PM performance of the "high" and "low" performance groups (considering their results on the filler trials). The obtained pattern of results was similar to that reported in the respective Results sections (except for Study 4, which was likely a random result, not replicated in Study 5 with a similar – yet more demanding – ongoing task). These results, corroborated with the Bayesian analyses, provide strong evidence against an animacy advantage in PM focal tasks. Therefore, it is unlikely that the high-performance levels account for the absence of the effect in our studies.

In conclusion, across a series of five well-powered studies, execution of the intended prospective action was not different between animate and inanimate targets, contrary to the findings for nonfocal PM tasks. These results align with the attentional-prioritization proximate mechanism of the animacy effect and support the Multiprocess Framework. Yet not all results (baseline and filler trials) are explained by such account. Therefore, the present findings call for continuing research on the underlying mechanisms of the animacy effect in (prospective) memory.

Author's note

Some of the Studies here reported were partially presented at ESCoP Conferences (2022, Lille, France; 2023, Porto, Portugal) and the International APPE-SEPEX Meeting (2022, Faro, Portugal). Studies 4 and 5 were conducted while SBF was a Visiting Scholar at City, University of London.

Ethics statements

All studies followed the Declaration of Helsinki. All participants provided their consent prior to data collection. Study 1 was revised by the Purdue's Institutional Review Board (Project Number: 1301013109). Studies 2 and 3 were approved by the Ethics and Deontology Committee of the University of Aveiro (Ref: 34/2019). Studies 4 and 5 were approved by the Psychology Ethics Committee of the City, University of London (Refs: ETH2223-1163; ETH2223-1327).

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT 3.0 and 4.0 to review parts of the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

CRediT authorship contribution statement

Sara B. Félix: Investigation, Formal analysis, Data curation, Conceptualization, Writing – review & editing, Writing – original draft, Methodology. **Marie Poirier:** Writing – review & editing, Supervision, Funding acquisition. **Josefa N.S. Pandeirada:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Sara B. Felix reports financial support was provided by Foundation for Science and Technology. Josefa N. S. Pandeirada reports financial support was provided by Foundation for Science and Technology. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Portuguese Foundation for Science and Technology through a doctoral fellowship to SBF (SFRH/BD/145097/2019, COVID/BD/153450/2024), a grant to JNSP (CEECIND/01914/2017), and multiannual funding to the WJCR (UIDB/04810/2020). Data collection of Study 4 was partly funded by the J. Frank Yates Student Conference Award 2022 granted by the Psychonomic Society, United States to SBF. The City, University of London, United Kingdom funded Study 5.

The authors also thank to Dr. Pedro Bem-Haja for his help with the Bayesian analyses, Dr. Michelle E. Coverdale for contributions in programming and reviewing Study 1, Rui Lebre and Roberto Villas for their help in programming Study 3. Additionally, the authors thank the feedback provided to parts of this work by Dr. Katharina Schnitzspahn, Dr. Matthias Kliegel, Dr. Ute Bayen, and Fabian Gümüsdagli. The authors also thank Dr. James S. Nairne for his comments on the final draft of this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2025.104673.

Data availability

All studies were pre-registered on AsPredicted prior to data collection, except for Study 1 as, at the time the study was conducted, preregistration was not a common practice (Study 2: https://aspredicted.org/1WH_CT6; Study 3: https://aspredicted.org/1N2_BHT; Study 4: https://aspredicted.org/XKZ_H1X; Study 5: https://aspredicted.org/XNF_GVZ). The materials used in the present studies are available as Supplementary Materials (SM). The trial-level data and analysis codes are available via Open Science Framework: https://osf.io/c78qf/?view_only=None.

References

Aka, A., Bhatia, S., & McCoy, J. (2023). Semantic determinants of memorability. *Cognition, 239*, Article 105497. https://doi.org/10.1016/j.cognition.2023.105497

Aka, A., Phan, T. D., & Kahana, M. J. (2021). Predicting recall of words and lists. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 47*, 765–784. https://doi.org/10.1037/xlm0000964

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

Ballhausen, N., Schnitzspahn, K. M., Horn, S. S., & Kliegel, M. (2017). The interplay of intention maintenance and cue monitoring in younger and older adults' prospective memory. *Memory and Cognition, 45*, 1113–1125. https://doi.org/10.3758/s13421-017-0720-5

Bayen, U. J., Rummel, J., Ballhausen, N., & Kliegel, M. (2024). Prospective memory. In M. J. Kahana, & A. D. Wagner (Eds.), *The Oxford Handbook of Human Memory* (pp. 1453–1484). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190917982.001.0001.

Bonin, P., Gelin, M., & Bugaiska, A. (2014). Animates are better remembered than inanimates: Further evidence from word and picture stimuli. *Memory and Cognition, 42*, 370–382. https://doi.org/10.3758/s13421-013-0368-8

Bonin, P., Gelin, M., Dioux, V., & Méot, A. (2019). "It is alive!" evidence for animacy effects in semantic categorization and lexical decision. *Applied PsychoLinguistics, 40* (4), 965–985. https://doi.org/10.1017/S0142716419000092

Bonin, P., Thiebaut, G., Bugaiska, A., & Méot, A. (2022). Mixed evidence for a richness-of-encoding account of animacy effects in memory from the generation-of-ideas paradigm. *Current Psychology, 41*, 1653–1662. https://doi.org/10.1007/s12144-021-02666-8

Bowden, V. K., Smith, R. E., & Loft, S. (2021). Improving prospective memory with contextual cueing. *Memory and Cognition, 49*, 692–711. https://doi.org/10.3758/s13421-020-01122-5

Brandimonte, M. A., & Ferrante, D. (2008). The social side of prospective memory. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 347–366). Taylor & Francis Group.

Brewer, G. A., Knight, J. B., Marsh, R. L., & Unsworth, N. (2010). Individual differences in event-based prospective memory: Evidence for multiple processes supporting cue detection. *Memory and Cognition, 38*(3), 304–311. https://doi.org/10.3758/MC.38.3.304

Bugaiska, A., Grégoire, L., Camblats, A.-M., Gelin, M., Méot, A., & Bonin, P. (2019). Animacy and attentional processes: Evidence from the Stroop task. *The Quarterly Journal of Experimental Psychology, 72*, 882–889. https://doi.org/10.1177/1747021818771514

Chen, X., Wang, Y., Liu, L., Shi, H., Wang, J., Cui, J., Shum, D. H. K., & Chan, R. C. K. (2014). The effect and mechanisms of implementation intentions on prospective memory in individuals with and without schizotypal personality features. *Memory, 22*, 349–359. https://doi.org/10.1080/09658211.2013.792841

Cobanoglu, C., Cavusoglu, M., & Turktarhan, G. (2021). A beginner's guide and best practices for using crowdsourcing platforms for survey research: The case of Amazon Mechanical Turk (MTurk). *Journal of Global Business Insights, 6*, 92–97. https://doi.org/10.5038/2640-6489.6.1.1177

Cona, G., Bisiacchi, P. S., & Moscovitch, M. (2014). The effects of focal and nonfocal cues on the neural correlates of prospective memory: Insights from ERPs. *Cerebral Cortex, 24*, 2630–2646. https://doi.org/10.1093/cercor/bht116

Cona, G., Bisiacchi, P. S., Sartori, G., & Scarpazza, C. (2016). Effects of cue focality on the neural mechanisms of prospective memory: A meta-analysis of neuroimaging studies. *Scientific Reports, 6*, 25983. https://doi.org/10.1038/srep25983

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford University Press.

Craik, F. I. M. (1986). A_functional_account_of_age_differences_in_memory. In F. Klix, & H. Hangendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (pp. 409–422). Elsevier.

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

Daley, M. J., Andrews, G., & Murphy, K. (2020). Animacy effects extend to working memory: Results from serial order recall tasks. *Memory, 28*, 157–171. https://doi.org/10.1080/09658211.2019.1699574

DeYoung, C. M., & Serra, M. J. (2021). Judgments of learning reflect the Animacy advantage for memory, but not beliefs about the effect. *Metacognition and Learning, 16*, 711–747. https://doi.org/10.1007/s11409-021-09264-w

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16*, 717–726. https://doi.org/10.1037/0278-7393.16.4.717

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science, 14*, 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*, 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*, 175–191. https://doi.org/10.3758/BF03193146

Félix, S. B., & Pandeirada, J. N. S. (2024). The animacy (bias) effect in recognition: Testing the influence of intentionality of learning and retrieval quality. *Memory, 32*, 889–900. https://doi.org/10.1080/09658211.2024.2362755

Félix, S. B., Pandeirada, J. N. S., & Nairne, J. S. (2019). Adaptive memory: Longevity and learning intentionality of the animacy effect. *Journal of Cognitive Psychology, 31*, 251–260. https://doi.org/10.1080/20445911.2019.1586716

Félix, S. B., Pandeirada, J. N. S., & Nairne, J. S. (2020). Animacy norms for 224 European Portuguese concrete words. *Análise Psicológica, 38*, 257–269. https://doi.org/10.14417/ap.1690

Félix, S. B., Poirier, M., Nairne, J. S., & Pandeirada, J. N. S. (2024). The breadth of animacy in memory: New evidence from prospective memory. *Psychonomic Bulletin & Review, 31*, 1323–1334. https://doi.org/10.3758/s13423-023-02406-y

Ferré, P., Comesaña, M., Félix, S. B., & Pandeirada, J. N. S. (2023). Do animacy and emotions go hand-by-hand on word processing? *23rd EScoP Conference, 115*.

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in Psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2*, 156–168. https://doi.org/10.1177/2515245919847202

Ganis, G., & Kievit, R. (2015). A new set of three-dimensional shapes for investigating mental rotation processes: Validation data and stimulus set. *Journal of Open Psychology Data, 3*, e3.

Gelin, M., Bugaiska, A., Méot, A., & Bonin, P. (2017). Are animacy effects in episodic memory independent of encoding instructions? *Memory, 25*, 2–18. https://doi.org/10.1080/09658211.2015.1117643

Gilbert, S. J. (2015). Strategic offloading of delayed intentions into the external environment. *The Quarterly Journal of Experimental Psychology, 68*, 971–992. https://doi.org/10.1080/17470218.2014.972963

Gilbert, S. J., Hadjipavlou, N., & Raoelison, M. (2013). Automaticity and control in prospective memory: A computational model. *PLoS One, 8*(3), Article e59852. https://doi.org/10.1371/journal.pone.0059852

Hagen, T., & Laeng, B. (2016). The change detection advantage for animals: An effect of ancestral priorities or progeny of experimental design? *I-Perception, 7*(3), 1–17. https://doi.org/10.1177/0956797610388814

Harris, L. M., & Menzies, R. G. (1999). Mood and prospective memory. *Memory, 7*, 117–127. https://doi.org/10.1080/741943717

Hauser, D. J., Paolacci, G., & Chandler, J. (2019). Common concerns with MTurk as a participant pool: Evidence and solutions. In F. R. Kardes, P. M. Herr, & N. Schwarz (Eds.), *Handbook in research methods in consumer Psychology* (pp. 319–337). Routledge.

Horn, S. S., & Bayen, U. J. (2015). Modeling criterion shifts and target checking in prospective memory monitoring. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 41*, 95–117. https://doi.org/10.1037/a0037676

Horn, S. S., Bayen, U. J., & Smith, R. E. (2011). What can the diffusion model tell us about prospective memory? *Canadian Journal of Experimental Psychology, 65*, 69–75. https://doi.org/10.1037/a0022808

Horn, S. S., Bayen, U. J., Smith, R. E., & Boywitt, C. D. (2011). The multinomial model of prospective memory: Validity of ongoing-task parameters. *Experimental Psychology, 58*, 247–255. https://doi.org/10.1027/1618-3169/a000091

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*, 1578–1596. https://doi.org/10.1080/02699931.2017.1423280

IBM. (2017). IBM SPSS Statistics for Windows. *IBM Corp.*

Izura, C., & Hernández-Muñoz, N. (2017). The role of semantics in Spanish word recognition: An insight from lexical decision and categorization tasks. *Open Linguistics, 3*, 500–515. https://doi.org/10.1515/opli-2017-0025

JASP Team. (2022). JASP (Version 0.16.3)[Computer software]. https://jasp-stats.org/.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*, 627–633. https://doi.org/10.3758/BRM.42.3.627

Klein, S. B., Robertson, T. E., & Delton, A. W. (2010). Facing the future: Memory as an evolved system for planning future acts. *Memory and Cognition, 38*(1), 13–22. https://doi.org/10.3758/MC.38.1.13

Kliegel, M., Guynn, M. J., & Zimmer, H. (2007). The role of noticing in prospective memory forgetting. *International Journal of Psychophysiology, 64*, 226–232. https://doi.org/10.1016/j.ijpsycho.2006.09.007

Kliegel, M., & Jäger, T. (2006). The influence of negative emotions on prospective memory: A review and new data. *International Journal of Computational Cognition, 1*, 1–17.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2004). Importance effects on performance in event-based prospective memory tasks. *Memory, 12*, 553–561. https://doi.org/10.1080/09658210344000099

Komar, G. F., Mieth, L., Buchner, A., & Bell, R. (2023a). Animacy enhances recollection but not familiarity: Convergent evidence from the remember-know-guess paradigm and the process-dissociation procedure. *Memory and Cognition, 51*, 143–159. https://doi.org/10.3758/s13421-022-01339-6

Komar, G. F., Mieth, L., Buchner, A., & Bell, R. (2023b). The animacy effect on free recall is equally large in mixed and pure word lists or pairs. *Scientific Reports, 13*, 11499. https://doi.org/10.1038/s41598-023-38342-z

Komar, G. F., Mieth, L., Buchner, A., & Bell, R. (2024). Manipulations of richness of encoding do not modulate the animacy effect on memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 50*, 580–594. https://doi.org/10.1037/xlm0001249

Kvavilashvili, L., & Ellis, J. (1996). Varieties of intention: Some distinctions and classifications. In M. A. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective Memory: Theory and Applications* (pp. 23–51). Psychology Press.

Leding, J. K. (2019). Adaptive memory: Animacy, threat, and attention in free recall. *Memory and Cognition, 47*, 383–394. https://doi.org/10.3758/s13421-018-0873-x

Leding, J. K. (2020). Animacy and threat in recognition memory. *Memory and Cognition, 48*, 788–799. https://doi.org/10.3758/s13421-020-01017-5

Lee, J. H., & McDaniel, M. A. (2013). Discrepancy-plus-search processes in prospective memory retrieval. *Memory and Cognition, 41*(3), 443–451. https://doi.org/10.3758/s13421-012-0273-6

Li, P., Jia, X., Li, X., & Li, W. (2016). The effect of animacy on metamemory. *Memory and Cognition, 44*, 696–705. https://doi.org/10.3758/s13421-016-0598-7

Loft, S., Dismukes, K., & Grundgeiger, T. (2019). Prospective memory in safety-critical work contexts. In J. Rummel, & M. A. McDaniel (Eds.), *Prospective memory* (pp. 170–185). Routledge.

Loft, S., & Humphreys, M. S. (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

Loucks, J., Reise, B., Gahite, R., & Fleming, S. (2023). Animate monitoring is not uniform: Implications for the animate monitoring hypothesis. *Frontiers in Psychology, 14*, Article 1146248. https://doi.org/10.3389/fpsyg.2023.1146248

Madan, C. R. (2021). Exploring word memorability: How well do different word properties explain item free-recall probability? *Psychonomic Bulletin & Review, 28*, 583–595. https://doi.org/10.3758/s13423-020-01820-w
Mah, E., Grannon, K. E. L., Campbell, A., Tamburri, N., Jamieson, R. K., & Lindsay, S. D. (2023). A direct replication and extension of Popp and Serra (2016, experiment 1): Better free recall and worse cued recall of animal names than object names, accounting for semantic similarity. *Frontiers in Psychology, 14*, Article 1146200. https://doi.org/10.3389/fpsyg.2023.1146200
Matos, P., Santos, F. H., & Albuquerque, P. B. (2020). When we must forget: The effect of cognitive load on prospective memory commission errors. *Memory, 28*, 374–385. https://doi.org/10.1080/09658211.2020.1726399
Maylor, E. A. (1996). Age-related impairment in an event-based prospective-memory task. *Psychology and Aging, 11*, 74–78. https://doi.org/10.1037/0882-7974.11.1.74
McBride, D. M., & Workman, R. A. (2017). Is prospective memory unique? A comparison of prospective and retrospective memory. In B. H. Ross (Ed.), Psychology of learning and motivation (Vol. 67, pp. 213–238). Elsevier. Doi: 10.1016/bs.plm.2017.03.007.
McDaniel, M. A., & Einstein, G. O. (1993). The importance of cue familiarity and cue distinctiveness in prospective memory. *Memory, 1*, 23–41. https://doi.org/10.1080/09658219308258223
McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14*, S127–S144. https://doi.org/10.1002/acp.775
McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. SAGE Publications.
McDaniel, M. A., Guynn, M. J., Einstein, G. O., & Breneiser, J. E. (2004). Cue-focused and reflexive-associative processes in prospective memory retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30*(3), 605–6014. https://doi.org/10.1037/0278-7393.30.3.605
McDaniel, M. A., Umanath, S., Einstein, G. O., & Waldum, E. R. (2015). Dual pathways to prospective remembering. *Frontiers in Human Neuroscience, 9*, 392. https://doi.org/10.3389/FNHUM.2015.00392/BIBTEX
Meinhardt, M. J., Bell, R., Buchner, A., & Röer, J. P. (2020). Adaptive memory: Is the animacy effect on memory due to richness of encoding? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 46*, 416–426. https://doi.org/10.1037/xlm0000733
Moyes, J., Sari-Sarraf, N., & Gilbert, S. J. (2019). Characterising monitoring processes in event-based prospective memory: Evidence from pupillometry. *Cognition, 184*, 83–95. https://doi.org/10.1016/j.cognition.2018.12.007
Nairne, J. S., & Pandeirada, J. N. S. (2008). Adaptive memory: Remembering with a stone-age brain. *Current Directions in Psychological Science, 17*, 239–243. https://doi.org/10.1111/j.1467-8721.2008.00582.x
Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology, 33*, 263–273. https://doi.org/10.1037/0278-7393.33.2.263
Nairne, J. S., VanArsdall, J. E., & Cogdill, M. (2017). Remembering the living: Episodic memory is tuned to animacy. *Current Directions in Psychological Science, 26*, 22–27. https://doi.org/10.1177/0963721416667711
Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N. S., Cogdill, M., & LeBreton, J. (2013). Adaptive memory: The mnemonic value of animacy. *Psychological Science, 24*, 2099–2105. https://doi.org/10.1177/0956797613480803
New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences, 104*, 16598–16603. https://doi.org/10.1073/pnas.0703913104
Nioka, S., Tracy, J. I., Raines, S., Bunce, S., & Chance, B. (2008). Anagram problem-solving and learning in anterior prefrontal cortex. *Journal of Innovative Optical Health Sciences, 1*, 157–175. https://doi.org/10.1142/S1793545808000194
Nowinski, J. L., & Dismukes, K. (2005). Effects of ongoing task context and target typicality on prospective memory performance: The importance of associative cueing. *Memory, 13*, 649–657. https://doi.org/10.1080/09658210444000313
Nowinski, J. L., Holbrook, J. B., & Dismukes, K. (2003). Human memory and cockpit operations: An ASRS study. In *12th International Symposium on Aviation Psychology* (pp. 888–893).
Penningroth, S. L. (2005). Effects of attentional demand, cue typicality, and priming on an event-based prospective memory task. *Applied Cognitive Psychology, 19*, 885–897. https://doi.org/10.1002/acp.1116
Popp, E. Y., & Serra, M. J. (2016). Adaptive memory: Animacy enhances free recall but impairs cued recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 42*, 186–201. https://doi.org/10.1037/xlm0000174
Radanović, J., Westbury, C., & Milin, P. (2016). Quantifying semantic animacy: How much are words alive? *Applied PsychoLinguistics, 37*, 1477–1499. https://doi.org/10.1017/S0142716416000096
Räling, R., Hanne, S., Schröder, A., Keßler, C., & Wartenburger, I. (2017). Judging the animacy of words: The influence of typicality and age of acquisition in a semantic decision task. *The Quarterly Journal of Experimental Psychology, 70*(10), 2094–2104. https://doi.org/10.1080/17470218.2016.1223704
Rawlinson, H. C., & Kelley, C. M. (2021). In search of the proximal cause of the animacy effect on memory: Attentional resource allocation and semantic representations. *Memory and Cognition, 49*, 1137–1152. https://doi.org/10.3758/s13421-021-01154-5
Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods, 22*, 304–321. https://doi.org/10.1037/met0000057
Rouse, S. V. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior, 43*, 304–307. https://doi.org/10.1016/j.chb.2014.11.004
Rummel, J., & Kvavilashvili, L. (2023). Current theories of prospective memory and new directions for theory development. *Nature Reviews Psychology, 2*, 40–54. https://doi.org/10.1038/s44159-022-00121-4
Rummel, J., Smeekens, B. A., & Kane, M. J. (2017). Dealing with prospective memory demands while performing an ongoing task: Shared processing, increased on-task focus, or both? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*, 1047–1062. https://doi.org/10.1037/xlm0000359
Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews. Neuroscience, 8*, 657–661. https://doi.org/10.1038/nrn2213
Schaper, M. L., Horn, S. S., Bayen, U. J., Buchner, A., & Bell, R. (2022). Adaptive prospective memory for faces of cheaters and cooperators. *Journal of Experimental Psychology. General, 151*, 1358–1376. https://doi.org/10.1037/xge0001128
Schnitzspahn, K. M., Plessow, F., Kirschbaum, C., Wong, Y. H., & Kliegel, M. (2022). Acute psychosocial stress impairs intention initiation in young but not older adults. *Psychoneuroendocrinology, 135*, Article 105593. https://doi.org/10.1016/j.psyneuen.2021.105593
Schnitzspahn, K. M., Thorley, C., Phillips, L., Voigt, B., Threadgold, E., Hammond, E. R., Mustafa, B., & Kliegel, M. (2014). Mood impairs time-based prospective memory in young but not older adults: The mediating role of attentional control. *Psychology and Aging, 29*, 264–270. https://doi.org/10.1037/a0036389
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
Scullin, M. K., McDaniel, M. A., & Shelton, J. T. (2013). The Dynamic Multiprocess Framework: Evidence from prospective memory with contextual variability. *Cognitive Psychology, 67*, 55–71. https://doi.org/10.1016/j.cogpsych.2013.07.001
Scullin, M. K., McDaniel, M. A., Shelton, J. T., & Lee, J. H. (2010). Focal/nonfocal cue effects in prospective memory: Monitoring difficulty or different retrieval processes? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*, 736–749. https://doi.org/10.1037/a0018971
Serra, M. J., & DeYoung, C. M. (2023). Within-pair factors might explain the inconsistent effects of animacy on paired-associates recall. *Psychonomic Bulletin & Review, 30*, 688–699. https://doi.org/10.3758/s13423-022-02184-z
Shelton, J. T., & Scullin, M. K. (2017). The dynamic interplay between bottom-up and top-down processes supporting prospective remembering. *Current Directions in Psychological Science, 26*, 352–358. https://doi.org/10.1177/0963721417700504
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*, 347–361. https://doi.org/10.1037/0278-7393.29.3.347
Smith, R. E. (2010). What costs do reveal and moving beyond the cost debate: Reply to Einstein and McDaniel. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*, 1089–1095. https://doi.org/10.1037/a0019183. in press.
Smith, R. E., Hunt, R. R., McVay, J. C., & McConnell, M. D. (2007). The cost of event-based prospective memory: Salient target events. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*, 737–746. https://doi.org/10.1037/0278-7393.33.4.734
Soares, A. P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. (2017). The Minho word pool: Norms for imageability, concreteness and subjective frequency for 3800 Portuguese words. *Behavior Research Methods, 49*, 1065–1081. https://doi.org/10.3758/s13428-016-0767-4
Strickland, L., Loft, S., & Heathcote, A. (2020). Investigating the effects of ongoing-task bias on prospective memory. *The Quarterly Journal of Experimental Psychology, 73*, 1495–1513. https://doi.org/10.1177/1747021820914915
Sugimori, E., & Kusumi, T. (2008). Relationship between cue word activation and prospective memory performance. *Psychological Reports, 102*, 317–327. https://doi.org/10.2466/pr0.102.1.317-327
VanArsdall, J. E., & Blunt, J. R. (2022). Analyzing the structure of animacy: Exploring relationships among six new animacy and 15 existing normative dimensions for 1,200 concrete nouns. *Memory and Cognition, 50*, 997–1012. https://doi.org/10.3758/s13421-021-01266-y
VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., & Cogdill, M. (2015). Adaptive memory: Animacy effects persist in paired-associate learning. *Memory, 23*, 657–663. https://doi.org/10.1080/09658211.2014.916304
Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., & Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*, 58–76. https://doi.org/10.3758/s13423-017-1323-7
Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of Psi: Comment on Bem (2011). *Journal of Personality and Social Psychology, 100*, 426–432. https://doi.org/10.1037/a0022790
West, R., Krompinger, J., & Bowry, R. (2005). Disruptions of preparatory attention contribute to failures of prospective memory. *Psychonomic Bulletin & Review, 12*, 502–507. https://doi.org/10.3758/BF03193795
Whittlesea, B. W. A., & Williams, L. D. (2001). The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 27*, 14–33. https://doi.org/10.1037/0278-7393.27.1.14

Wulff, D. U., Hills, T. T., & Mata, R. (2022). Structural differences in the semantic networks of younger and older adults. *Scientific Reports, 12*, 21459. https://doi.org/10.1038/s41598-022-11698-4

Zuber, S., Kliegel, M., & Ihle, A. (2016). An individual difference perspective on focal versus nonfocal prospective memory. *Memory and Cognition, 44*, 1192–1203. https://doi.org/10.3758/s13421-016-0628-5

Zuber, S., Mahy, C. E. V., & Kliegel, M. (2019). How executive functions are associated with event-based and time-based prospective memory during childhood. *Cognitive Development, 50*, 66–79. https://doi.org/10.1016/j.cogdev.2019.03.001