Distinguishing the Impact of Age on Semantic and Non-Semantic Associations in Episodic Memory

Gabriel Jarjat, PhD

Geoff Ward, D.Phil

Pascal Hot, PhD

Sophie Portrat, PhD

Vanessa M. Loaiza, PhD

Manuscript accepted for publication at *Journals of Gerontology, Series B: Psychological Sciences* on 8 January 2020.

Word count: 5942/6000 (excluding references, tables, and figures)

Author Note

Gabriel Jarjat, Université Grenoble Alpes, Université Savoie Mont Blanc. Geoff Ward, University of Essex. Pascal Hot, Université Grenoble Alpes, Université Savoie Mont Blanc, Institut Universitaire de France. Sophie Portrat, Université Grenoble Alpes, Université Savoie Mont Blanc. Vanessa M. Loaiza, University of Essex.

Correspondence concerning this article should be addressed to Vanessa M. Loaiza, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Email: v.loaiza@essex.ac.uk

The pre-registration, materials, data, and analysis scripts for the experiment are available at the Open Science Framework: https://osf.io/hydqw/

Abstract

Objectives. Refreshing, or the act of briefly foregrounding recently presented but now perceptually absent representations, has been identified as a possible source of age differences in working memory and episodic memory. We investigated whether the refreshing deficit contributes to the well-known age-related deficit for retrieving non-semantic associations, but has no impact on existing semantic associations.

Method. Younger and older adults judged the relatedness of stimulus word pairs (e.g., pink–blue or pink–cop) after repeating or refreshing one of the words. During a later source recognition memory test, participants determined whether each item recognized as old was presented on the left or right (non-semantic source memory) and presented in a related or unrelated pair (semantic source memory). The data were analyzed using a hierarchical Bayesian implementation of a multinomial model of multidimensional source memory.

Results. Neither age group exhibited a refreshing benefit to non-semantic or semantic source memory parameters. There was a large age difference in non-semantic source memory, but no age difference in semantic source memory.

Discussion. The study suggests that the nature of the association is most important to episodic memory performance in older age, irrespective of refreshing, such that source memory is unimpaired for semantically meaningful information.

Keywords: attention, multinomial modeling, source memory, semantic memory

A substantial literature suggests age-related declines in long-term *episodic memory* (EM), or memory of personally experienced events (see Zacks et al., 2000, for a review). By contrast, long-term *semantic memory*, or general knowledge about the world, improves with age. One important difference between these systems is that specific, contextual details from the original event are not required for semantic memory, whereas retrieving from EM (e.g., words presented in different locations) involves recalling individual units (henceforth, *item memory*, e.g., the word and/or location) and their specific associations (henceforth, *source memory*, e.g., the association between the word and its location; Chalfonte & Johnson, 1996).

Researchers have observed stronger age deficits in source memory than item memory (Chalfonte & Johnson, 1996; Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995). Source memory is thought to reflect a combination of perceptual, reflective, and attributional decision processes that decline with normal and pathological aging (Mitchell & Johnson, 2009). One such potentially age-deficient process is *refreshing*, or directing attention to recently-presented but no longer perceptually-available representations to keep them active in working memory (WM; see Camos et al., 2018 for a review), in turn reinforcing bindings between memoranda and their source context (Johnson et al., 2002; Loaiza & McCabe, 2012, 2013; Loaiza & Souza, 2018, 2019; Souza et al., 2015). For example, Loaiza and Souza (2019) showed that refreshing increased WM recall in part by reducing binding errors. An age-related refreshing deficit may therefore contribute to impaired source memory if refreshing facilitates content-context binding in WM, that is, the *refreshing deficit hypothesis*.

The classic refreshing paradigm compares the relative long-term benefit of cueing participants to think back to a just-recently presented word (i.e., *refresh*) to re-presenting the word (i.e., *repeat*) immediately after it has disappeared (Johnson et al., 2002). The paradigm also often includes a condition of reading the words once without repetition/refreshing (i.e., *read*) that is sometimes used as a baseline to compare the effect of refreshing rather than the repeat condition. We focus on refresh-repeat comparisons because they more appropriately determine whether refreshing provides unique memorial benefits beyond mere perceptual repetition. Johnson and colleagues have observed that younger adults' later recognition memory was better for refreshed versus repeated words (Johnson et al., 2002; Raye et al., 2002; Johnson et al., 2005), although this has not consistently been demonstrated (Bartsch, Loaiza, Jäncke, et al., 2019; Grillon et al., 2008; Johnson et al., 2004; Raye et al., 2008). Conversely, older adults typically do not show a refreshing benefit (Johnson et al., 2002, 2004; Raye et al., 2008), thereby indicating that deficient refreshing in older age may contribute to age-related impairments in EM. However, the evidence for a specific refreshing deficit in older age is inconsistent (Baumans et al., 2012; Duarte et al., 2013; Jarjat et al., 2019; Souza, 2016), thus emphasizing that more investigation is warranted to understand if and how refreshing contributes to age-related impairments in memory performance.

The aforementioned work has focused primarily on general recognition memory performance, without distinguishing between item memory and source memory. However, as explained previously, some have argued that refreshing specifically operates on content-context bindings (e.g., Loaiza & Souza, 2019; Souza, 2016). Refreshing may therefore be uniquely important for source memory, although this has not yet been demonstrated. Nevertheless, refreshing may not be relevant for all types of source memory given that not all types of source memory show age deficits. Indeed, some work has indicated a small or null source memory deficit for meaningfully-related compared to unrelated information (Mohanty et al., 2016; Naveh-Benjamin, 2000). Further work indicates that arbitrary features (e.g., font size, position of studied words on the screen) are less likely to contribute to older adults' recollective experience compared to younger adults (Boywitt et al., 2012; Kuhlmann & Boywitt, 2016). Given that refreshing is argued to promote content-context binding and that the age-related source memory deficit is greatest for information that lacks semantic meaningfulness, it may be that refreshing is most important for building novel, arbitrary associations that do not already have strong representation in semantic memory (Loaiza et al., 2015). If so, older adults may show less of a refreshing deficit to source memory when they can rely on their relatively greater semantic memory.

To investigate whether impaired refreshing specifically contributes to the age-related source memory deficit for novel associations, younger and older participants performed a two-dimensional source memory task (Meiser & Bröder, 2002). One dimension drew upon existing associations in semantic memory (henceforth, *semantic source memory*) and the other concerned associations for which existing semantic memory has no relevance (henceforth, *non-semantic source memory*). Specifically, either semantically related (e.g., *pink – blue*) or unrelated (e.g., *pink – cop*) pairs of words were presented, and after the words disappeared from the screen, the refreshing manipulation was implemented. Following previous work (Johnson et al., 2002), participants either repeated or refreshed one of the words of each pair when prompted with either the original word re-appearing or an asterisk appearing in the corresponding original frame, respectively. Thus, repeating served as the control condition in order to examine the benefits of refreshing beyond the mere priming of repetition. During a later recognition test, participants distinguished new words (*lures*) from previously studied words (*targets*) and, for those detected as old, (1) whether the word had originally been displayed on the left- or right-hand side of the screen (i.e., non-semantic source memory), and (2) whether the word came from a related or unrelated pair (i.e., semantic source memory).

We fit the data to a multinomial processing tree (MPT) model for multidimensional joint source memory (Meiser, 2014; Meiser & Bröder, 2002; Figure 1) to examine the impact of refreshing on parameters representing semantic source memory (e_{ij}^{Rel}), non-semantic source memory (e_{ij}^{Loc}), and joint source memory (d_{ij}) for both younger and older adults.[1] If refreshing

[1] Although we focus on these predictions here for the sake of brevity, note that the full list of research questions and their respective predictions can be found on the Open Science Framework (OSF) and their corresponding

specifically promotes novel content-context bindings, then non-semantic source memory should be greater in the refresh condition than in the repeat condition for younger adults and not older adults, consistent with the refreshing deficit hypothesis. This pattern may also be evident for joint source memory if refreshing impacts the bound recollection of the two sources. Finally, if refreshing does not act on existing representations in semantic memory, then there should be no reliable differences in semantic source memory between the refresh and repeat condition for either age group.

Method

Participants

In line with our pre-registration, we recruited 30 younger and 30 older participants from the university's subject pool and local community (Table 1). The sample size was determined from similar previous literature using hierarchical Bayesian MPT models (e.g., Bartsch, Loaiza, & Oberauer, 2019; Loaiza & Srokova, 2019). Bayesian inference allows continued sampling if the evidence for or against an effect is ambiguous (Rouder, 2014), but in this case the planned sample size was sufficient for drawing firm inferences. Eleven additional participants were excluded from the analysis for being non-native English speakers (one younger), currently taking medication for neurological disorders (two younger, one older), low score (i.e., < 26; 1 older) on the mini mental status examination (MMSE; Folstein et al., 1975), computer malfunction (three older), or failure to achieve 80% accuracy during the relatedness task (three older). All participants provided written informed consent and were debriefed and compensated with £9 at the end of the study. The study was approved by the Department of Psychology Ethics Committee at the University of Essex.

Materials and Procedure

results in the Supplementary Materials.

Stimuli consisted of 380 pairs of concrete singular English nouns taken from the Nelson free association norms (Nelson et al., 2004). Each pair comprised two semantically related words (forward associative strength: *M* = 0.18, *SD* = 0.07, range = 0.101 – 0.399; backward associative strength: 0; letters: *M* = 5.49, *SD* = 1.69, range = 2 – 10; log HAL frequency: *M* = 9.23, *SD* = 1.63, range = 3.09 – 13.58). For each pair, one word was taken and associated to a word from another pair in order to create an unrelated option. Thus, there was a pool of 380 triplets of words, each comprising two semantically related words and a third unrelated word. For each participant, 128-word triplets were randomly selected from the pool for the relatedness judgment task: Half the words composed the related word pairs (e.g., *pink* – *blue*) and half composed the unrelated word pairs (e.g., *pink* – *cop*), with the constraint that all words composing the pairs were unique (i.e., a word could not appear in both a related and unrelated pair). Words serving as the lures during the recognition test were drawn from the remaining unique words.

Participants were individually tested in a quiet room with a trained experimenter present. The experiment was programmed using Matlab with the Psychtoolbox extensions (Brainard, 1997; Kleiner et al., 2007). The instructions for each phase were displayed on the computer, with the experimenter clarifying any questions.

The first phase comprised two blocks of the relatedness judgment task of 64 trials each. Two practice trials preceded the first block and were repeated until participants fully understood the requirements of the task. Each trial began with two horizontally presented empty frames displayed for 1000 ms, followed by two words appearing in the respective frames for 2250 ms, which then disappeared for a 500 ms interstimulus interval (ISI). Then, participants were prompted to either refresh or repeat one of the two words for 1500 ms (500 ms ISI). In the refresh condition, an asterisk appeared in one of the frames, indicating the word that was in that frame should be refreshed aloud. In the repeat condition, one of the words from the pair reappeared in the same frame as before, and participants were instructed to read it aloud. Participants' responses were recorded and later transcribed. The contour of the boxes then turned blue, prompting participants to judge whether the words of the pair were related or unrelated by pressing a corresponding key within 10 s, and the accuracy and response times (RTs) of their decisions were recorded. After participants responded or when the time limit was reached, the frames disappeared, and the trial ended, with an interval of 2 s before a new trial. The design was fully balanced, such that, across the two blocks, there was an even representation of words in each of the design cells that fully and randomly crossed the instruction, the pair relatedness, and the location of the repeated/refreshed words.

A distraction-filled retention interval of 2 min followed each block wherein participants completed as many simple addition problems (e.g., 32 + 11 = ?) as possible. Each problem was displayed until participants entered a response by using the keyboard. Afterward, participants received instructions for the source recognition test. Each block of the test comprised 128 randomly presented words: Half were originally repeated/refreshed during the previous block (targets) and half were entirely new words to the experiment (lures). Non-repeated/non-refreshed words were not tested. Each word was presented at the center of the screen with the words "Old" and "New" displayed in two respective frames below it, and participants used the mouse to respond. If participants decided that the test word was new, another test word appeared after a 500 ms ISI. If judged as old, participants further indicated whether the test word came from a related or unrelated pair and whether it was displayed on the left or right side of the screen. These options were presented as four horizontally presented frames on the screen with each possible combination presented within each (i.e., left-related, left-unrelated, right-related, right-unrelated), and participants used the mouse to click on the relevant frame.

Finally, participants completed a demographics questionnaire and a computerized vocabulary test (Shipley et al., 2009). Additionally, older adults completed the MMSE. At the conclusion of the study, participants were offered the opportunity to view their overall performance during the source recognition test and the vocabulary test.

Data analysis

All analyses were conducted with R in RStudio. Separate 2 (age: young, older adults) x 2 (instruction: refresh, repeat) x 2 (relatedness: related, unrelated pairs) x 2 (location: left, right) Bayesian analyses of variance (BANOVAs, Rouder, Morey, Speckman, & Province, 2012; Wetzels, Grasman, & Wagenmakers, 2012) were conducted for speech RTs, and response accuracy and RTs during the relatedness judgment task using the BayesFactor package (Morey & Rouder, 2015) with the default settings. For the sake of brevity, we report on the best models, and the latter two measures are reported in the Supplementary Materials. BANOVA computes the strength of evidence in the data for different models which include a combination of main effects and interaction compared with a null model (M_{Null}) that includes only a random effect of the participant. The ratio of evidence of comparing one model (e.g., a model assuming a main effect of age, M_A) and another (e.g., the null model, M_{Null}) is referred to as the Bayes factor (BF). Thus, one can compare the relative strength of evidence for one model over another in order to draw inferences about effects and interactions of different factors. As Wetzels and colleagues (2012) advised, we compared the ratio of BFs between the model including a term of interest against the null model and the model removing this term against the null model. For example, to assess the main effect of age, we used the ratio of BFs between the model that includes age and instruction without an interaction against the null model (M_{A+I}/M_{Null}) and the model including only the instruction factor against the null model (M_I/M_{Null}); that is, BF[(M_{A+I}/M_{Null})/(M_I/M_{Null})]. BFs of greater than 10 and 100 are respectively considered strong and decisive evidence for the tested model, whereas BFs between 1 and 3 are considered ambiguous evidence for the tested model. For ease of comprehension, we will express BFs of less than 1 as their inverse (e.g., a BF of 0.286 is reported as 1/3.49).

To analyze source recognition performance, we used a MPT model for multidimensional joint source memory using TreeBUGS (Heck et al., 2018).[2] MPT models are a class of measurement models that aim to explain observable categorical data (e.g., frequency of responses during a source recognition task) in terms of underlying latent cognitive states (Batchelder & Riefer, 1999; Erdfelder et al., 2009). Rather than typical measures of observed performance (e.g., hits, false alarms) that conflate memory and decision processes (e.g., guessing), MPT modeling formally distinguishes between these processes and those of interest (e.g., source memory). Figure 1 illustrates the model parameters for item memory, source memory, and guessing processes that lead to an observed response (right-hand side of Figure 1) from a target with dimensions i (relatedness of the pair), $i \in$ {related, unrelated}, and j (location on the screen), $j \in$ {left, right}, or from a lure (left-hand side of Figure 1). Parameter D_{ij} denotes the probability of recognizing a target from source ij (e.g., a word from a related pair presented on the left) as old, and parameter D_{New} denotes the probability of identifying a lure as new. Conversely, participants may not recognize an old or new item as such with the probabilities $1 - D_{ij}$ and $1 - D_{New}$, respectively. Given a correctly identified old item (D_{ij}), participants may retrieve the correct source combination i,j (relatedness and location) in a bound fashion with probability d_{ij}. If participants cannot retrieve the correct bound source combination ($1 - d_{ij}$), they may still retrieve each source dimension independently with probabilities e_{ij}^{Rel} (relatedness) and e_{ij}^{Loc} (location). Participants may also recognize targets but not recollect any source information, and thus respond correctly based on lucky guesses or guess incorrectly. Such guesses are reflected by parameters a^{Rel}, $a_{|i}^{Loc}$, $a_{|\mathrm{A}}^{Rel}$ *and* $a_{|\mathrm{B}}^{Loc}$. If an item is not recognized

[2] To ensure identifiability of the model, some parameters were constrained to be identical (Meiser, 2014) and in line with our preregistration. More specifically, detection parameters D_{ij} were set to be equal across both locations. Assuming that the detection of a lure is determined by the lower boundary of target detection, we constrained D_{New} to be equal to the probability of detecting a target studied from an unrelated pair that was repeated. In addition, guessing parameters were equated between recognized and unrecognized targets. Finally, we imposed equality constraints on the source memory parameters d_{ij}, e_{ij}^{Rel}, and e_{ij}^{Loc} across source combinations. Note that an alternative model (detailed on the OSF and the Supplementary Materials) omitting the joint source memory parameter (Meiser & Bröder, 2002) yielded similar results.

as old or new ($1 - D_{ij}$ or $1 - D_{New}$, respectively), participants may guess that this unrecognized item is old with probability *b*. Thereafter, participants assign items to relatedness and location sources with the guessing probabilities g^{Rel}, $g_{|A}^{Loc}$, $g_{|B}^{Loc}$. For the sake of brevity, details of the modeling procedure and its fit, as well as item memory parameter estimates are reported in the Supplementary Materials.

Results

Relatedness judgment task

Speech response time. Participants spoke the correct item on 98% of the trials.[3] The BANOVA performed on mean log-transformed speech RTs showed that the best model included main effects of age and instruction (BF[M_{A+I} *vs.* M_{Null}] = 2.02*e*+7). There was strong evidence that older adults were slower (*M* = 1.06 s, *SD* = 0.17) than younger adults (*M* = 0.80 s, *SD* = 0.13; BF[M_{A+I} *vs.* M_I] = 2.82*e*+6). There was also moderate evidence that participants were slower to repeat (*M* = 0.94 s, *SD* = 0.21) than to refresh words (*M* = 0.92 s, *SD* = 0.19; BF[M_{A+I} *vs.* M_A] = 7.38). Contrary to our hypothesis and previous work (e.g., Johnson et al., 2002), there was moderate evidence against the interaction between age and instruction (BF[M_{A*I} *vs.* M_{A+I}] = 1/4.05; Figure 2).

Source recognition memory test

We fitted a Bayesian hierarchical latent-trait MPT model (Klauer, 2010) for each age group, the results of which are presented in Table 2 and Figure 3. To infer differences between experimental conditions, we inspected the 95% Bayesian-credible interval (CI) of the sampled parameter differences between conditions. Specifically, we computed the difference between refreshed and repeated targets to consider the effect of refreshing for each parameter in each age group. To consider age differences in the parameter estimates, we also computed the difference between younger and older adults' posterior distributions. A 95% CI that does not

[3] Fitting the MPT models using only data where participants spoke correctly and made correct semantic relatedness decisions yielded the same pattern of results as reported here with the full dataset.

contain 0 would suggest a difference between conditions for the corresponding parameter estimates.[4]

Our principal analyses concerned source memory: The joint recollection (d_{ij}) of semantic (i.e., the relatedness of the pair) and non-semantic (i.e., the location of the target) source-specific information, as well as the specific recollection of the semantic information (i.e., relatedness, e_{ij}^{Rel}) and of the non-semantic information (i.e., location, e_{ij}^{Loc}). For both age groups the 95% CI of the estimated mean difference in joint source memory between the refreshed and repeated conditions contained 0 (i.e., the first row labeled *refreshing effect* in Table 2). This suggests that refreshing did not benefit the joint source memory for either age group. There was also no credible age difference in the parameter for either the repeat or refresh conditions, as the respective 95% CI of the estimated mean difference between the younger and older adults both contained 0 (i.e., the column labeled *age effect* in Table 2).

Next, we considered specific source recollection. For semantic source memory (e_{ij}^{Rel}), inspection of the 95% CI of the estimated mean difference between the refreshed and repeated conditions (i.e., the second row labeled *refreshing effect* in Table 2) showed that there was no credible difference between the repeat and refresh conditions for either age group. There was also no credible difference between younger and older adults for either repeat or refresh conditions (i.e., the column labeled *age effect* in Table 2), suggesting that semantic source memory was age invariant. For non-semantic source memory (e_{ij}^{Loc}), the 95% CI of the estimated mean difference between the repeat and refresh conditions included 0 for both age groups (i.e., the third row labeled *refreshing effect* in Table 2). Thus, as was the case for joint

[4] In the preregistration of this study, we planned to inspect the 95% CI of the posterior mean group-level estimates and to base our inference on the presence of an overlap of the CI. However, other recent studies using MPT modelling approach used the 95% CI of the mean difference (e.g., Filevich et al., 2019; Schaper et al., 2019). Thus, we chose to use the same hypothesis testing method as in the current literature. It should be noted that these two different methods led to the same conclusions except in two occasions. More precisely, for the parameter D_{ij} in the Related-Repeat condition and parameter e_{ij}^{Loc} in the Repeat condition, there was an overlap between the 95% CI of the posterior mean group-level estimates between younger and older adults, suggesting no age difference, although the 95% CI of the estimated mean difference between both age groups excluded 0, suggesting an age difference.

source memory (d_{ij}), there was no evidence for a refreshing benefit in either age group. For both the repeat and refresh conditions, the 95% CI of the estimated mean difference between younger and older adults did not include 0 (i.e., the column labeled *age effect* in Table 2). Thus, younger adults had a credibly higher probability of recollecting the location of the targets than older adults, regardless of whether the targets were previously repeated or refreshed. In sum, as was the case for joint source memory (d_{ij}), there was no evidence for a refreshing benefit in either age group for the specific recollection of the semantic information (e_{ij}^{Rel}) and of the non-semantic information (e_{ij}^{Loc}).

Discussion

The present study investigated whether an age difference in refreshing, or the act of reflectively directing attention to information in WM (Camos et al., 2018), contributes to the well-documented age-related deficit in source memory (Mitchell & Johnson, 2009), especially for novel content-context bindings that have no existing representation in semantic memory. The premise for such an important role of refreshing originates from theoretical proposals that refreshing promotes content-context bindings (Johnson et al., 2002; Loaiza & McCabe, 2012; Souza et al., 2015), but may have little impact on extant semantic memory representations (Loaiza et al., 2015; Loaiza & Camos, 2018). Given work suggesting an age-related refreshing deficit (e.g., Jarjat et al., 2019; Johnson et al., 2002; Loaiza & McCabe, 2013) but preserved retention of semantically meaningful information (Mohanty et al., 2016; Naveh-Benjamin, 2000), this study makes the connection between refreshing and different types of source memory more explicit than in past work. As such, the study represents an important step in delineating the impact of aging and refreshing for different types of source memory.

The current study provides two novel results: (a) Refreshing did not impact source memory for either age group, and (b) a strong asymmetry in the effects of age for different types of source information: Unlike the typical profound deficits for non-semantic source memory, older adults were just as capable as younger adults to retrieve semantically meaningful source information. We discuss each of these findings in the next respective sections.

Refreshing and Source Memory in Older Age

In contrast to our primary predictions, refreshing had no effect on joint or non-semantic source memory beyond the mere beneficial effect of repetition for either age group. Older adults were also not slower to refresh memoranda aloud compared to younger adults, as has been suggested in prior work (e.g., Johnson et al., 2002). These results contrast with previous studies suggesting that younger adults exhibit a refreshing advantage in EM that older adults do not (Johnson et al., 2002, 2004) and the notion that this refreshing deficit may dissociate between semantic and non-semantic source memory (Loaiza et al., 2015). The results are more in line with the view that refreshing may not contribute at all to memory performance for either age group (Bartsch, Loaiza, Jäncke, et al., 2019).

At first glance, the lack of refreshing benefit in younger adults is peculiar given that the current paradigm was modeled after Johnson and colleagues' original refreshing paradigm. Specifically, the sequence of presentation of stimuli, timing parameters, and implementation of repeating/refreshing the words were the same as in prior work (e.g., Raye et al., 2008). The primary methodological differences were the relatedness judgments following each pair and that the final recognition test was known in advance, concerned memory for source, and was unspeeded. One may first consider the speech RTs as a possible indicator that refreshing was not similarly engaged as in prior work. Although the speech RTs were generally slower than in some work (e.g., Johnson et al., 2002), they were still in range of other similar studies (Bartsch, Loaiza, Jäncke, et al., 2019; Raye et al., 2008). Perhaps more important was the lack of slowed speech RTs to refresh rather than repeat the items, but it is not immediately clear why this should translate to a null impact of refreshing on source memory. It could indicate that participants engaged in other processes besides refreshing in anticipation of the relatedness judgment task, but it is also just as likely that participants simply withheld a quick response compared to the classic paradigm that only emphasizes quick responding with no other requirements of the task. Another issue could be that making the relatedness judgment in itself required refreshing because the words had disappeared from the screen. However, prior work has shown a beneficial effect of refreshing frequency in visual WM paradigms wherein participants reconstruct the original color of the items to recall them, and in so doing, perhaps refresh the color again (Souza et al., 2015). The fact that a refreshing effect is still observed in such conditions suggests that it is unlikely that the relatedness judgment in the current study diluted any potential advantage of refreshing in younger adults. The substantial variability within age groups in terms of the individual parameter estimates (Figure 3) suggests that there may be individual differences in the refreshing benefit, which may be a fruitful area of future research. Finally, and most importantly, the original paradigm has not always yielded a refreshing benefit beyond mere repetition in younger adults even when indicators like slowed speech RTs are evident (Grillon et al., 2008; Johnson et al., 2004; Raye et al., 2008). Thus, the lack of refreshing effect we observed in the current study cannot be so easily dismissed on methodological grounds given that it has not been consistently observed in the original paradigm.

Even if the methodological differences between the current and original paradigms nullified the refreshing advantage in younger adults, this begs the question of the role of refreshing. The concept of refreshing has played a central role in theorizing concerning developmental and individual differences in how people maintain and manipulate information in WM and the consequent long-term retention of that information in EM (Camos et al., 2018). As such, the concept has been pushed beyond the original simple procedure of briefly thinking back to previously presented stimuli, with no further required tasks or direct knowledge of a later memory test. If refreshing plays a direct and purposeful role in EM formation, where the demands of ongoing cognition are much greater than simple repetition, then it is necessary to test the concept of refreshing in more demanding and meaningful settings. If a refreshing effect is limited to situations where it is not relevant to keeping the contents of mind active, then it necessarily cannot be a useful concept for understanding the nature of age differences in memory ability. Note that we do not believe this to be the case, but this reasoning underlines the importance of pushing the methodology of refreshing further as we have tried here.

So far, these enumerated possibilities center on the manipulation of refreshing itself. Still another possibility is that the nature of the non-semantic source information, i.e., the original location of the memoranda on the screen, reduced the possibility to observe an impact of refreshing. Previous work has suggested that non-semantic sources such as the location of the memoranda or the list in which they are presented are rather immutable during encoding (Malmberg & Shiffrin, 2005). According to the one-shot context hypothesis (Malmberg & Shiffrin, 2005), there is a relatively fixed amount of context/source storage that is not very susceptible to factors that might otherwise strengthen memory representations, such as deep levels of processing, increased encoding time or, perhaps in this current case, instructions to refresh information. Thus, it could be the case that merely refreshing memoranda may not result in a dramatic change in non-semantic source memory simply because such information is encoded in a relatively fixed way.

Other related work has investigated whether the age-related refreshing deficit could manifest beyond a relative inability to focus attention to include further subcomponents of refreshing, such as switching attention between active memoranda (Loaiza & Souza, 2018) and preserving the benefits of focused attention even after distraction (Loaiza & Souza, 2019). The results of these studies suggested that older adults show a similar refreshing benefit to younger adults when the task requires simply focusing attention and switching attention to other relevant memoranda, whereas the refreshing benefit is disproportionately susceptible to distractor interference in older age. This indicates that specific subcomponents of refreshing (i.e., preserving the benefits of refreshing against distraction) are deficient, whereas other subcomponents (i.e., focusing and switching attention) are relatively intact in older age (see also Plancher et al., 2017). Such findings are relevant here because the current refreshing manipulation only required participants to focus their attention, and thus it is possible that refreshing effects may be more strongly evident under conditions with distraction. Future work could explore whether a parallel finding is evident in this paradigm.

In summary, the null effect of refreshing on non-semantic source memory conflicted with our predictions and the refreshing deficit hypothesis more generally. Notwithstanding, there are potential explanations that must be explored in future research in order to reinforce these findings. The most theoretically significant include whether the results are similar in conditions wherein participants must retain the effects of refreshing after their attention has become distracted (Loaiza & Souza, 2019) and whether refreshing may impact non-semantic source information other than spatial-temporal contexts (Malmberg & Shiffrin, 2005).

An Age Dissociation of Semantic and Non-Semantic Source Memory

A striking finding from the present study is that of the strong asymmetry in terms of how age impacted the retention of different source information, irrespective of the null impact of refreshing discussed previously. Whereas there was evidence for an age difference in source memory for non-semantic information (i.e., the original locations of the memoranda), we found no evidence for an age difference in source memory for semantic information (i.e., the relatedness of the memoranda). This finding suggests that the nature of the association is most important to EM performance in older age, such that source memory is unimpaired for semantically meaningful information.

This finding is in line with other studies showing that age differences are reduced when source memory relies on established prior knowledge and meaningful information (Mohanty et al., 2016; Naveh-Benjamin, 2000). For example, Loaiza and Srokova (2019) demonstrated that adapting the semantic relatedness of presented word pairs according to participants' ongoing performance improved the retention of the bindings between the pairs to such an extent that older adults' performance matched that of younger adults. Critically, older adults required a greater proportion of the pairs to be semantically related in order to achieve matched binding memory to younger adults, thus emphasizing the notion that older adults are especially sensitive to semantically meaningful information that is rooted in their prior knowledge (Umanath & Marsh, 2014). The current study extends this existing literature by suggesting that age differences in source memory depend on the type of source information tested.

Interestingly, there was also no age difference in the joint source memory parameters. It should be noted, however, that the posterior-mean estimates for all the joint source memory parameters were very low overall, and thus it is possible that any potential age deficits or indeed impact of refreshing are obfuscated by floor-level estimates. Thus, we refrain from interpreting the findings regarding joint source memory, as it appears that independent source memory was much more likely overall in the current study.

In summary, the second critical finding of the current study paints a very different picture than what is otherwise typical of aging and source memory studies. Unlike the commonly demonstrated profound deficit for non-semantic source information, source memory for semantically meaningful information, such as the relatedness of the original pair, appears to be much more intact in normal aging.

Conclusions

In summary, the current study provides novel evidence that age differences in source memory were much smaller, even null, when the memoranda were semantically meaningful, and regardless of whether they were refreshed in WM or not. Although we did not observe a refreshing advantage in younger adults, this study provides insight into the nature of age-related deficits in source memory, such that novel, episodic-based associations that have no basis in existing semantic memory are likely to exhibit typical deficits, whereas associations that are strongly represented in semantic memory do not indicate any such decline. Such findings imply that, rather than any deficit to refresh memoranda in WM, older adults' EM performance is increasingly sensitive to the scaffolding their relatively superior semantic memory can provide.

Funding

This work was supported in part by an Experimental Psychology Society (EPS) grant to G.J.

Acknowledgments

We acknowledge Hiu "Lily" Cheung and Nathanael Knight for their assistance with data collection.

References

Bartsch, L. M., Loaiza, V. M., Jäncke, L., Oberauer, K., & Lewis-Peacock, J. A. (2019). Dissociating refreshing and elaboration and their impacts on memory. *NeuroImage, 199*, 585–597. https://doi.org/10.1016/j.neuroimage.2019.06.028

Bartsch, L. M., Loaiza, V. M., & Oberauer, K. (2019). Does limited working memory capacity underlie age differences in associative long-term memory? *Psychology and Aging*, *34*, 282–293. https://doi.org/10.1037/pag0000317

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*(1), 57–86.

Baumans, C., Adam, S., & Seron, X. (2012). Effect of Cognitive Load on Working Memory Forgetting in Aging. *Experimental Psychology*, *59*(6), 311–321. https://doi.org/10.1027/1618-3169/a000158

Boywitt, C. D., Kuhlmann, B. G., & Meiser, T. (2012). The role of source memory in older adults' recollective experience. *Psychology and Aging*, *27*(2), 484–497. https://doi.org/10.1037/a0024729

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 19–32. https://doi.org/10.1111/nyas.13616

Chalfonte, B. l., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*(4), 403–416. https://doi.org/10.3758/BF03200930

Duarte, A., Hearons, P., Jiang, Y., Delvin, M. C., Newsome, R. N., & Verhaeghen, P. (2013). Retrospective attention enhances visual working memory in the young but not the old:

An ERP study: Retrospective attention and aging. *Psychophysiology*, *50*(5), 465–476. https://doi.org/10.1111/psyp.12034

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial Processing Tree Models: A Review of the Literature. *Zeitschrift Für Psychologie / Journal of Psychology*, *217*(3), 108–124. https://doi.org/10.1027/0044-3409.217.3.108

Filevich, E., Horn, S. S., & Kühn, S. (2019). Within-person adaptivity in frugal judgments from memory. *Psychological Research*, *83*(3), 613–630. https://doi.org/10.1007/s00426-017-0962-7

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Grillon, M.-L., Johnson, M. K., Krebs, M.-O., & Huron, C. (2008). Comparing effects of perceptual and reflective repetition on subjective experience during later recognition memory. *Consciousness and Cognition*, *17*(3), 753–764. https://doi.org/10.1016/j.concog.2007.09.004

Heck, D. W., Arnold, N. R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, *50*(1), 264–284. https://doi.org/10.3758/s13428-017-0869-7

Jarjat, G., Portrat, S., & Hot, P. (2019). Aging Influences the Efficiency of Attentional Maintenance in Verbal Working Memory. *The Journals of Gerontology: Series B*, *74*(4), 600–608. https://doi.org/10.1093/geronb/gby067

Johnson, M. K., Mitchell, K. J., Raye, C. L., & Greene, E. J. (2004). An age-related deficit in prefrontal cortical function associated with refreshing information. *Psychological Science*, *15*(2), 127–132.

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: Prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective & Behavioral Neuroscience*, *5*(3), 339–361. https://doi.org/10.3758/CABN.5.3.339

Johnson, M. K., Reeder, J. A., Raye, C. L., & Mitchell, K. J. (2002). Second thoughts versus second looks: An age-related deficit in reflectively refreshing just-activated information. *Psychological Science*, *13*(1), 64–67. https://doi.org/10.1111/1467-9280.00411

Klauer, K. C. (2010). Hierarchical Multinomial Processing Tree Models: A Latent-Trait Approach. *Psychometrika*, *75*(1), 70–98. https://doi.org/10.1007/s11336-009-9141-0

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Kuhlmann, B. G., & Boywitt, C. D. (2016). Aging, source memory, and the experience of "remembering." *Aging, Neuropsychology, and Cognition*, *23*(4), 477–498. https://doi.org/10.1080/13825585.2015.1120270

Loaiza, V. M., & Camos, V. (2018). The Role of Semantic Representations in Verbal Working Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(6), 863–881. https://doi.org/10.1037/xlm0000475

Loaiza, V. M., & McCabe, D. P. (2012). Temporal–contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition*, *40*(2), 191–203. https://doi.org/10.3758/s13421-011-0148-2

Loaiza, V. M., & McCabe, D. P. (2013). The influence of aging on attentional refreshing and articulatory rehearsal during working memory on later episodic memory performance. *Aging, Neuropsychology, and Cognition*, *20*(4), 471–493.

Loaiza, V. M., Rhodes, M. G., & Anglin, J. (2015). The Influence of Age-Related Differences in Prior Knowledge and Attentional Refreshing Opportunities on Episodic Memory. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *70*(5), 729–736. https://doi.org/10.1093/geronb/gbt119

Loaiza, V. M., & Souza, A. S. (2018). Is refreshing in working memory impaired in older age? Evidence from the retro-cue paradigm. *Annals of the New York Academy of Sciences*, *1424*(1), 175–189. https://doi.org/10.1111/nyas.13623

Loaiza, V. M., & Souza, A. S. (2019). An Age-Related Deficit in Preserving the Benefits of Attention in Working Memory. *Psychology and Aging*, *34*, 268–281.

Loaiza, V. M., & Srokova, S. (2019). Semantic Relatedness Corrects the Age-Related Binding Deficit in Working Memory and Episodic Memory. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*. https://doi.org/10.1093/geronb/gbz055

Malmberg, K. J., & Shiffrin, R. M. (2005). The “one-shot” hypothesis for context storage. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(2), 322–336. https://doi.org/10.1037/0278-7393.31.2.322

Meiser, T. (2014). Analyzing stochastic dependence of cognitive processes in multidimensional source recognition. *Experimental Psychology*, *61*(5), 402–415. https://doi.org/10.1027/1618-3169/a000261

Meiser, T., & Bröder, A. (2002). Memory for Multidimensional Source Information. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(1), 116–137. https://doi.org/10.1037/0278-7393.28.1.116

Mitchell, K. J., & Johnson, M. K. (2009). Source Monitoring 15 Years Later: What Have We Learned From fMRI About the Neural Mechanisms of Source Memory? *Psychological Bulletin*, *135*(4), 638–677. https://doi.org/10.1037/a0015849

Mohanty, P., Naveh-Benjamin, M., & Ratneshwar, S. (2016). Beneficial effects of semantic memory support on older adults' episodic memory: Differential patterns of support of item and associative information. *Psychology and Aging*, *31*(1), 25–36. https://doi.org/10.1037/pag0000059

Morey, R. D., & Rouder, J. N. (2015). *BayesFactor: Computation of Bayes factors for common designs.* (Version 0.9.12-2) [Computer software]. http://CRAN.R-project.org/package=BayesFactor

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(5), 1170–1187. https://doi.org/10.1037//0278-7393.26.5.1170

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

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

Plancher, G., Boyer, H., Lemaire, B., & Portrat, S. (2017). Under Which Conditions Can Older Participants Maintain Information In Working Memory? *Experimental Aging Research*, *43*(5), 409–429. https://doi.org/10.1080/0361073X.2017.1369730

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a Single Thought: Dorsolateral PFC Activity Associated with Refreshing Just-Activated Information. *NeuroImage*, *15*(2), 447–453. https://doi.org/10.1006/nimg.2001.0983

Raye, C. L., Mitchell, K. J., Reeder, J. A., Greene, E. J., & Johnson, M. K. (2008). Refreshing one of several active representations: Behavioral and functional magnetic resonance imaging differences between young and older adults. *Journal of Cognitive Neuroscience*, *20*(5), 852–862. https://doi.org/10.1162/jocn.2008.20508

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301–308. https://doi.org/10.3758/s13423-014-0595-4

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Schaper, M. L., Kuhlmann, B. G., & Bayen, U. J. (2019). Metamemory expectancy illusion and schema-consistent guessing in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 470–496. https://doi.org/10.1037/xlm0000602

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009). *Shipley-2*. CA: Western Psychological Services.

Souza, A. S. (2016). No age deficits in the ability to use attention to improve visual working memory. *Psychology and Aging*, *31*(5), 456–470. https://doi.org/10.1037/pag0000107

Souza, A. S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 20–31. https://doi.org/10.1111/nyas.12603

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

Umanath, S., & Marsh, E. J. (2014). Understanding How Prior Knowledge Influences Memory in Older Adults. *Perspectives on Psychological Science, 9*(4), 408–426. https://doi.org/10.1177/1745691614535933

Wetzels, R., Grasman, R. P. P. P., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for ANOVA designs. *American Statistician*, *66*(2), 104–111. https://doi.org/10.1080/00031305.2012.695956

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human Memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 293–357). Mahwah, N.

Table 1

Sample characteristics.

Measure	Younger adults	Older adults	Age Comparison
Final *N* (*n* exclusions)	30 (3)	30 (8)	-
Age in years - *M* (*SD*)	20.40 (3.11)	68.47 (5.19)	-
Age range	18 – 31	60 – 79	-
Sex (male/female)	9/21	9/21	BF_{10} = 1/3.49
MMSE	-	29.10 (1.09)	-
Years of education	15.12 (2.16)	14.83 (3.92)	BF_{10} = 1/3.62
Shipley vocabulary (proportion correct)	0.75 (0.09)	0.90 (0.06)	BF_{10} = 3.27e+7
Number of medications	0.47 (0.82)	1.83 (2.23)	BF_{10} = 28.15
Rated current health (1 - 5, 1 = very good)	1.83 (0.75)	1.83 (0.75)	BF_{10} = 1/12.74
Rated general health (1 - 5, 1 = very good)	1.80 (0.66)	1.83 (0.70)	BF_{10} = 1/19.36
Rated restrictions of health (1 - 4, 1 = no restrictions)	1.30 (0.53)	1.47 (0.82)	BF_{10} = 1/21.13

Note. MMSE = mini mental status examination; BF = Bayes factor.

Table 2

Posterior-mean estimates for joint source memory (d_{ij}), semantic source memory (e_{ij}^{Rel}), and non-semantic source memory (e_{ij}^{Loc}), and 95% Bayesian credible intervals (CI) as a function of experimental conditions.

		Age group				Age effect	
		Younger adults		Older adults			
Parameter	Condition	Mean	95% CI	Mean	95% CI	Mean	95% CI
d_{ij}	Repeat	.21	[.03, .39]	.05	[.00, .14]	-.16	[-.36, .04]
	Refresh	.09	[.01, .26]	.05	[.00, .16]	-.04	[-.23, .11]
	Refreshing effect	-.12	[-.34, .13]	<.01	[-.11, .12]		
e_{ij}^{Rel}	Repeat	.69	[.58, .78]	.67	[.59, .74]	-.02	[-.15, .11]
	Refresh	.75	[.67, .80]	.67	[.59, .74]	-.08	[-.17, .02]
	Refreshing effect	.06	[-.05, .18]	<.01	[-.09, .10]		
e_{ij}^{Loc}	Repeat	.40	[.19, .58]	.09	[.01, .22]	-.31	[-.52, -.07]
	Refresh	.55	[.40, .67]	.25	[.08, .40]	-.31	[-.51, -.10]
	Refreshing effect	.15	[-.05, .37]	.15	[-.04, .32]		

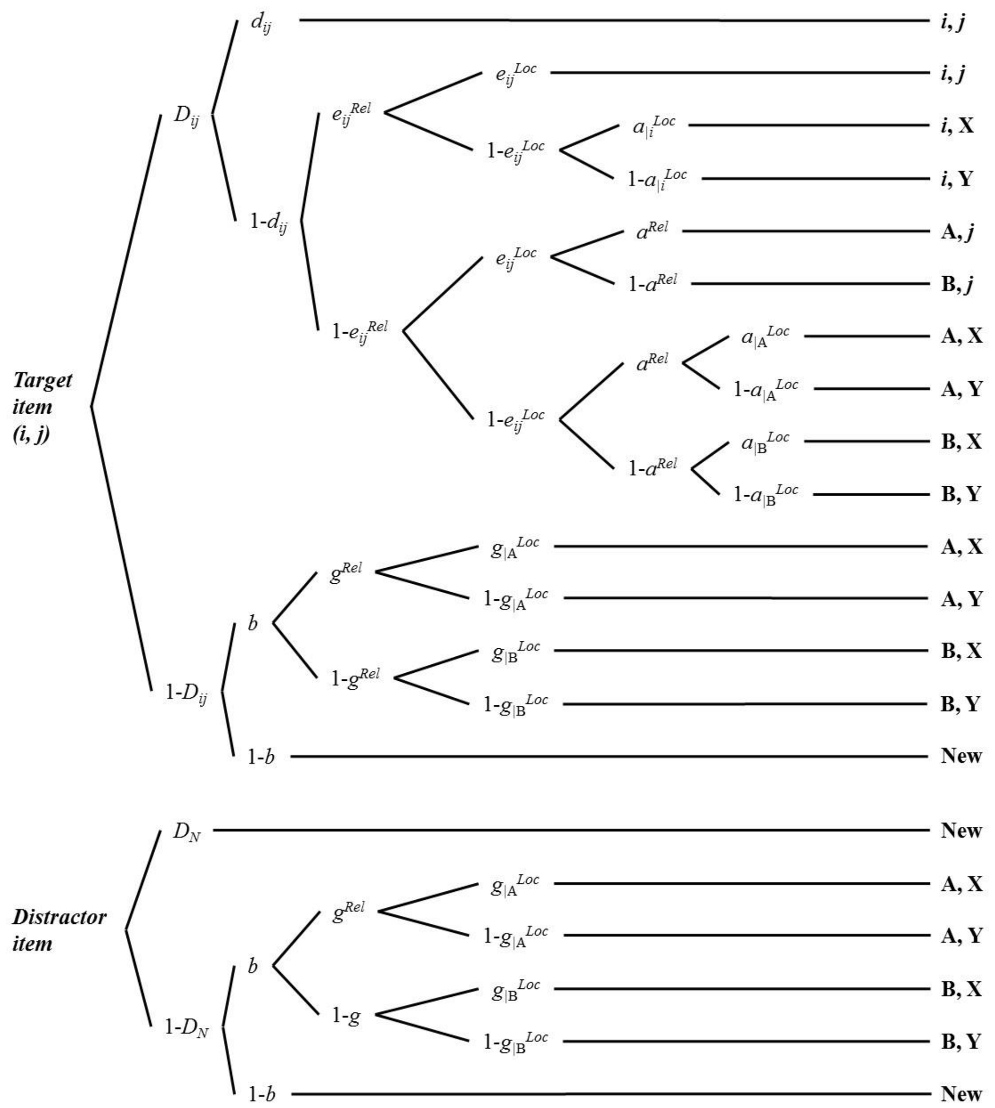

Figure 1. Processing-tree diagram of the multinomial model of multidimensional joint source memory. Attribute $i \in \{A; B\}$ on the first dimension (Relatedness), A denoting “Related” and B ”Unrelated”. Attribute $j \in \{X; Y\}$ on the second dimension (Location), X denoting “Right” and Y “Left”. D_{ij} = probability of recognizing studied items from the sources i and j; D_{New}= probability of recognizing distractor items as new; d_{ij} = probability of joint retrieval of source i and source j; e_{ij}^{Rel} = probability of independent retrieval of source i; $\text{e}_{ij}^{\text{Loc}}$ = probability of independent retrieval of source j; a^{Rel} = probability of guessing ‘‘A’’ on the first dimension for recognized studied items; $a_{|\text{A}}^{\text{Loc}}$, $a_{|\text{B}}^{\text{Loc}}$ = probability of guessing ‘‘X’’ on the second dimension for recognized studied items assigned to A or B, respectively; b = probability of guessing ‘‘old’’; g^{Rel} = probability of guessing ‘‘A’’ on the first dimension for unrecognized target or lure items; $g_{|\text{A}}^{\text{Loc}}$, $g_{|\text{B}}^{\text{Loc}}$ = probability of guessing ‘‘X’’ on the second dimension for unrecognized target or lure items assigned to A or B, respectively.

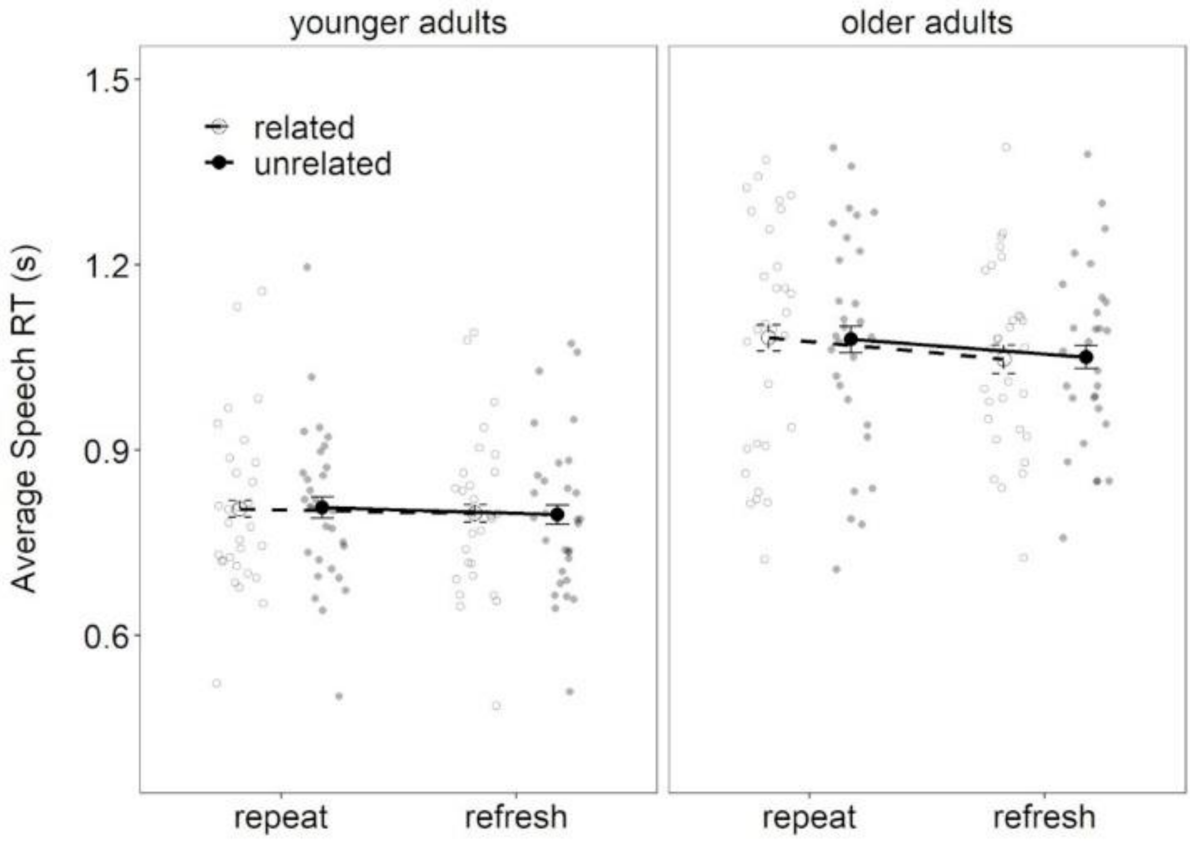

Figure 2. Measured mean speech reaction times in the relatedness judgement task as a function of Age, Relatedness, and Instruction. Error bars represent 95% within-subjects confidence intervals.

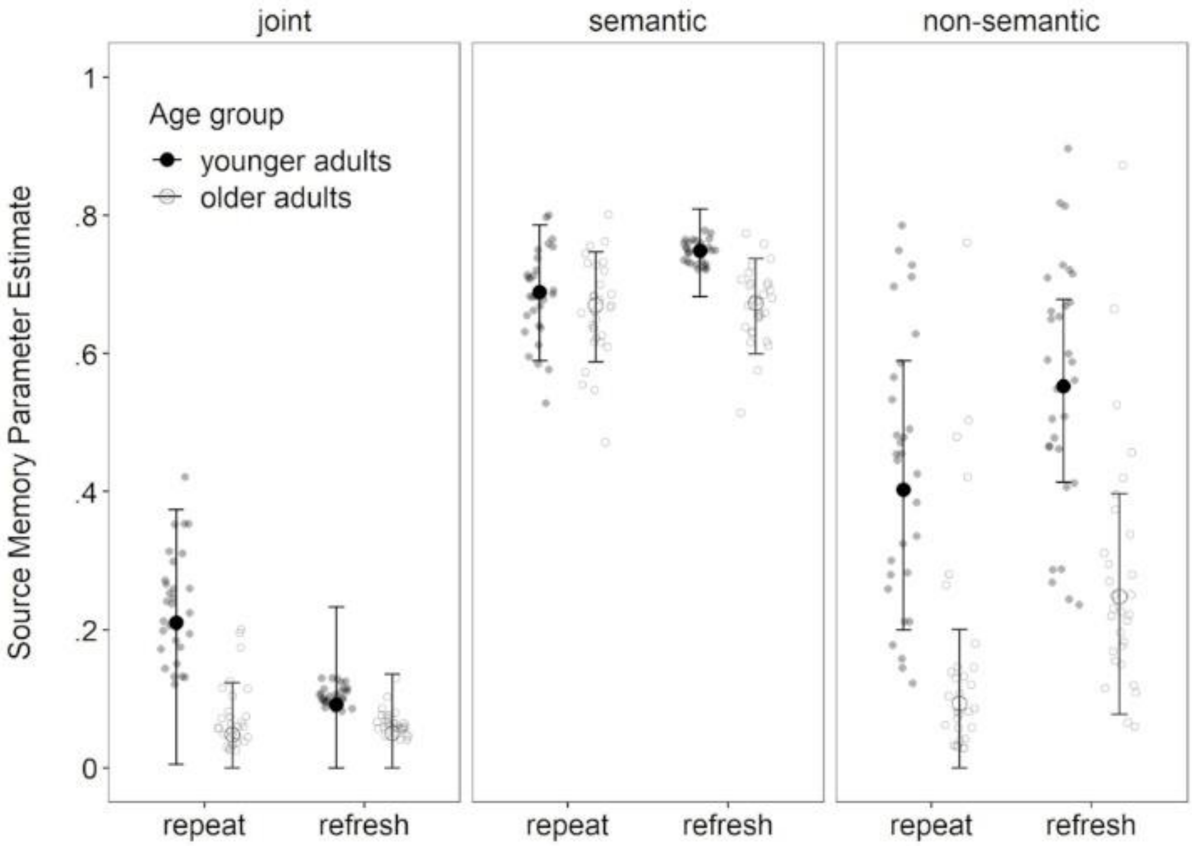

Figure 3. Posterior means of the estimated source memory parameters. Error bars reflect 95% Bayesian credibility intervals and individual points reflect the individual participants' parameter estimates.