The effects of divided attention at encoding on specific and gist-based associative episodic memory

Nathaniel R. Greene[1] · Moshe Naveh-Benjamin[1]

Abstract

Effects of divided attention (DA) during encoding on later memory performance are widely documented. However, the precise nature of these effects on underlying memory representations and subsequent retrieval processes has not been thoroughly investigated. Here, we examined whether DA at encoding would disrupt young adults' ability to remember associations in episodic memory at highly specific levels of representation (i.e., verbatim memory), or whether the effects of DA extend also to gist memory for associations. Two groups of participants (one under full attention, one under DA) studied face–scene pairs. The DA group simultaneously completed an auditory choice reaction-time task during encoding. Following either a short or long delay, participants were tested on their ability to discriminate intact face–scene pairs from recombined pairs that were either highly similar, less similar, or completely unrelated to originally studied pairs. The DA group performed more poorly than the full attention participants at correctly classifying most types of test pairs at both delays, and results from a multinomial-processing-tree model demonstrated that participants who encoded associations under DA experienced deficits in both specific and gist memory retrieval. We also compared the DA group to full attention older adults who were tested with the same paradigm (Greene & Naveh-Benjamin, *Psychological Science, 31*[3], 316–331, 2020). The DA group had lower estimates of gist retrieval than the older adults but similar estimates of verbatim memory. These results suggest that DA at encoding disrupts episodic memories at multiple levels of representation, in contrast to age-related effects, which are restricted only to the highest levels of specificity.

Keywords Divided attention · Associative recognition · Episodic memory · Aging

Effects of divided attention (DA) with a secondary task at encoding on memory performance have been documented for numerous tasks of memory (e.g., Baddeley et al., 1984; Castel & Craik, 2003; Craik et al., 2018; Craik et al., 1996; Greene et al., in press; Kilb & Naveh-Benjamin, 2007; Murdock, 1965; Naveh-Benjamin et al., 1998; Naveh-Benjamin et al., 2003; Nieznański, 2013). However, there is scarce research examining the effects of DA at encoding on underlying memory representations. Does DA disrupt the ability to remember episodes at specific levels of representation, or do the effects of DA extend to less detailed representations? Answering this question can shed insight into the relationship between attention and memory at multiple levels

of specificity and into potential mechanisms accounting for adult age-related deficits in memory for specific details of past events (e.g., Castel, Farb, & Craik, 2007; Greene & Naveh-Benjamin, 2020; Luo & Craik, 2009; Stark et al., 2013; Tun et al., 1998), as depletion of attentional resources have been proposed to mediate age-related memory declines (e.g., Craik & Byrd, 1982).

Levels of specificity in episodic memory An episodic memory is a representation of a past event, bounded in a specific time and place (Jones, 1976; Tulving, 1983; Underwood, 1969). Associations among components of an event lie at the core of episodic memory (Tulving, 1983; Zimmer, 2006). Thus, successful remembering of a past event requires encoding and retrieving associations among event components, and failures to do so may have profound implications, for example, in eyewitness situations which require a witness to remember who committed a specific action.

An episode may be remembered at a highly specific level of representation (e.g., remembering specifically the location in which a person was previously encountered) or at less

✉ Nathaniel R. Greene
ngwfm@umsystem.edu

✉ Moshe Naveh-Benjamin
navehbenjaminm@missouri.edu

[1] Department of Psychological Sciences, University of Missouri, 9J McAlester Hall, Columbia, MO 65211, USA

specific levels of representation (e.g., remembering in general that a person had been encountered outside, but not remembering specifically *where* outside this encounter occurred; Greene & Naveh-Benjamin, 2020). This view of episodic memories as being accessible from different levels of specificity is in line with theories suggesting episodic and semantic memories exist on a continuum and that access to more specific nodes on the continuum may be affected by factors that disrupt memory, such as aging (Craik, 2002, 2006; Greene & Naveh-Benjamin, 2020).

The idea that episodic memories can be remembered on a *continuum of specificity* is somewhat unique from other popular conceptualizations of memory, such as fuzzy-trace theory, which posits that information in memory is simultaneously processed in two parallel traces—a verbatim trace, which encodes surface-level contextual details of the episode, and a gist trace, which encodes semantic details of the episode—and that with time or interference, verbatim traces are susceptible to decay, whereas gist traces remain stable (Brainerd & Reyna, 2015; Reyna & Brainerd, 1995). There are similarities between a continuum-of-specificity view and fuzzy-trace theory, including that both predict that access to the most specific information in memory is most susceptible to forgetting. However, whereas fuzzy-trace theory conceptualizes gist memory as a semantic representation of a past event, we are primarily concerned with assessing the representation of *episodic* content in memory, and whether such episodic representations are highly specific or less detailed. Nevertheless, we will use the terms specific and verbatim interchangeably, to refer to a representation of an association in memory that retains precise information about specifically which components had been paired together during encoding, and we will describe less detailed representations of associations (i.e., remembering the association at a more general level) as gist.

Effects of divided attention on different levels of specificity

To date, few studies have assessed whether DA affects the ability to remember associations that lie at the core of episodic memories across different levels of specificity. Dodson et al. (1998), using a source monitoring task, suggested that DA, manipulated at *retrieval*, impairs specific but not gist retrieval of source information associated with spoken sentences. However, effects of DA at retrieval are much less notable than effects of DA at encoding (Craik et al., 2018; Craik et al., 1996), where DA has been shown to produce marked deficits in associative memory (Craik et al., 2010; Kilb & Naveh-Benjamin, 2007; Naveh-Benjamin et al., 2003), but these studies did not examine whether DA at encoding impacted highly specific or gist representations of associations.

Some studies have investigated the effect of DA on false memory production, using the Deese–Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995), in which participants study a list of items (e.g.,

"bed," "dream," "pillow") that are closely related to an unpresented lure ("sleep"). False recognition or recall of the lure is expected to occur when individuals fail to retrieve verbatim memory traces of list items and rely only on gist memory traces (Brainerd et al., 1999; Brainerd et al., 2003; Odegard & Lampinen, 2005; for a different interpretation based on an activation-monitoring account, see Roediger & McDermott, 1995). DA at encoding has been shown to increase false recall but reduce false recognition of lures in the DRM paradigm (Dewhurst et al., 2005; Dewhurst et al., 2007; Knott & Dewhurst, 2007; Knott et al., 2018; Pérez-Mata et al., 2002). Dewhurst et al. (2007) interpreted these findings in the context of activation-monitoring theory (Roediger & McDermott, 1995), arguing that DA at encoding decreases subsequent false recognitions because the secondary tasks inhibit participants from generating semantic associates of target words during study. In contrast, they argued that higher rates of false recall in the DRM paradigm could be attributable to changes in response biases.

It is worth noting that not only was false recognition of lures reduced in these studies, but correct recognition of old items was reduced as well, and these reductions in veridical and false recognitions were only evident when "old" responses were accompanied by recollective phenomenology (Dewhurst et al., 2005; Dewhurst et al., 2007), but not by feelings of familiarity (for more on the distinction between recollection and familiarity, see Gardiner, 1988; Tulving, 1985). Thus, it is not clear whether the results from Dewhurst and colleagues reflect deficiencies in verbatim or gist memory, as both memory traces can support recollection (Brainerd et al., 2014; Brainerd et al., 1999). Also, false recognition of a related lure in the DRM paradigm may reflect that participants remember the *semantic* gist of the studied material (e.g., "bed," "dream," and "pillow" are related to "sleep"), but it does not necessarily indicate that participants retrieve a fuzzier *episodic* representation for any studied item.

Using a conjoint recognition paradigm, Odegard and Lampinen (2005) showed that effects of DA were restricted to verbatim memory (specifically, recollection rejection of related lures), and did not affect gist memory, which is in line with evidence suggesting that the gist of an item is processed before attention is deployed (Wallace et al., 2000; Wallace et al., 1998) and see Draine and Greenwald (1998) for evidence of semantic priming effects that occur before awareness of a prime's physical presence). Nevertheless, the study by Odegard and Lampinen (2005) only examined verbatim and gist memory for items. By measuring specific and gist memory for associations between components, we can more directly assess the effects of DA on different levels of specificity in episodic memory.

Recently, Greene and Naveh-Benjamin (2020) developed a paradigm, based on the simplified conjoint recognition paradigm for item memory (Stahl & Klauer, 2008), for measuring

whether associations in episodic memory can be accessed at different levels of specificity. They presented participants with face–scene pairs, such as an old man paired with a specific park (e.g., Old Man–Park 1). At test, participants were tasked with discriminating Intact pairs (Old Man–Park 1) from Recombined pairs, which included highly similar foils (e.g., Old Man–Park 2), foils that were similar at a broader level of representation (e.g., Old Man–Forest), and foils that were dissimilar (e.g., Old Man–Kitchen). Participants judged whether each test pair was "intact" (meaning identical to a studied pair), "related" (meaning highly similar to a studied pair), or "unrelated" (meaning not alike a studied pair). From the response frequencies given to the different test probes, Greene and Naveh-Benjamin (2020) estimated verbatim memory (e.g., remembering specifically that the old man had been paired with Park 1), episodic gist memory (e.g., remembering that the old man had been paired with a park, but not whether it had been Park 1 or Park 2), and even fuzzier memory (e.g., remembering encountering the old man outside, but not at a category-specific level to distinguish whether the old man had been paired with a park or a forest), with a multinomial-processing-tree (MPT) model (Stahl & Klauer, 2008). The main aim of the present study was to couple this paradigm with a DA manipulation at encoding, to see whether the effects of DA on associative memory (e.g., Naveh-Benjamin et al., 2003) are restricted to specific memory or extend to gist memory as well.

In another recent study, Greene et al. (2020) asked participants to make "old/new" judgments to face–scene pairs, including Intact (old pairs) and three different types of Recombined (new pairs), which varied in how similar they were to old pairs. In their Experiment 2, Greene et al. (2020) found that young adults who encoded associations under DA were more prone to incorrectly responding "old" to all types of Recombined pairs, regardless of how similar those pairs were to old pairs. These results suggest that DA may disrupt not only specific associative memory, but also gist-based associative memory. However, Greene et al. (2020) were unable to separate the contributions of specific and gist memory on their task, whereas in the present study, using the original paradigm from Greene and Naveh-Benjamin (2020), we can more definitively measure the contributions of specific and gist memory using a well-validated paradigm and associated measurement model.

Can depleted attentional resources account for older adults' specificity deficits? Another aim of the present study was to assess whether DA in young adults would produce comparable memory deficits for associations at highly specific levels that have been shown in older adults (Greene & Naveh-Benjamin, 2020), thereby assessing whether a depletion of attentional resources could be one mechanism accounting for older adults' deficits in specific associative episodic memory. Greene and Naveh-Benjamin (2020) found that older

adults' differential deficits in associative memory, which have been widely reported in the literature (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008), may be restricted to specific, verbatim levels, because older adults were as capable as younger adults at remembering gist and even fuzzier details for studied face–scene pairs. What underlying mechanisms may account for older adults' deficits in *specific* associative episodic memory? One appealing possibility is that older adults have more severely limited attentional resources than younger adults, and thus are less able to allocate attention to the encoding of specific associations into memory. DA paradigms are one useful method of testing this hypothesis, by simulating diminished attentional resources at encoding in *young adults*, who do not normally show as pronounced of a deficit in specific associative memory (e.g., Greene & Naveh-Benjamin, 2020).

The depleted attentional resources hypothesis attributes older adults' cognitive deficits (e.g., in memory) to general age-related deficits in attentional mechanisms (Craik & Byrd, 1982). Some support for this hypothesis has come from studies showing that DA in young adults produces comparable performance to older adults (e.g., Castel & Craik, 2003). However, other studies have found that DA produces a more "general" deficit than that associated with aging. For example, Naveh-Benjamin et al. (2003) found that DA at encoding in young adults resulted in deficits in both item and associative memory, whereas age-related deficits were largely restricted specifically to associative memory (see also, Craik et al., 2010; Kilb & Naveh-Benjamin, 2007). Therefore, it is not clear whether a depleted attentional resources hypothesis can adequately account for all age-related memory deficits. On the one hand, if depleted attentional resources can explain older adults' memory deficits, then we should expect that simulating depleted attention in young adults, using a DA manipulation at encoding, would produce comparable deficits in *specific* associative memory but result in preserved gist memory for associations, much like aging (Greene & Naveh-Benjamin, 2020). Alternatively, if depleting attention at encoding results in a more general deficit in the quality of memories, then we would expect that DA in young adults will result in deficits in specific and gist memory for associations.

The present study

The primary aim of the present study was to assess whether DA at encoding in young adults disrupts associative episodic memory at specific and/or gist levels of representation, using a recently developed paradigm from Greene and Naveh-Benjamin (2020). The present study will therefore provide important insight into how disruptive the effects of DA are on episodic memory. That is, are these effects observable only for highly specific associative information in memory, or do the detrimental effects of DA extend to less detailed levels of

representation (i.e., the gist of an episode)? According to fuzzy-trace theory (Reyna & Brainerd, 1995), gist memory traces are less susceptible to interference, so we may expect that DA would not affect gist memory. However, fuzzy-trace theory conceives of gist memory as representing the semantic aspects of an episode (Brainerd & Reyna, 2015). Here, we are focusing on *episodic gist*—that is, memory for an association at a less detailed level of representation. While this may rely on meaning-based aspects of an episode to some extent (e.g., "this old man was paired with a park scene," which reflects the general meaning of the pairing as "old man–park," even if the specific instantiation of the park is not retrievable), it is still the case that some episodic details must be remembered, as the example above illustrates with respect to remembering that the old man had been paired with *some* park scene, which requires remembering some episodic details of the encoding context. According to the hierarchical representation model (Craik, 2002), a "gist" representation in this sense still contains some contextual information about the encoding context and would therefore exist at a more specific level of representation than a semantic gist representation. As such, it is conceivable that the disruptive effects of DA may manifest with deficits in memory for the gist of an association. Indeed, in a recent study, Greene et al. (2020) suggested that DA at encoding may have a more widespread effect on associative memory, across various levels of specificity, including for the gist of an association, though they were not able to measure this with their paradigm.

We used the same paradigm as Experiment 2 of Greene and Naveh-Benjamin (2020) to provide the cleanest comparison of our results with the age-related effects that were observed in that study, thereby allowing us to test whether DA results in the same deficits in specific associative episodic memory that were observed with older adults, or if such effects extend to gist associative memory as well. An additional feature of this design is that there were two different delays between the encoding and retrieval phases of the experiment. Half of the blocks featured a short (5 second) delay between the study and test phases, whereas the other half included a long (5 minute) delay. Thus, we also assessed whether DA effects on task performance and specific and gist memory for associations depend on the delay between the encoding and retrieval phases. We made no specific hypotheses about whether DA effects on specific and/or gist memory would interact with delay, so we consider any findings concerning delay effects to be exploratory in nature.

Method

Participants

We recruited 107 participants from introductory psychology classes, who were randomly assigned to the full attention (FA)

condition ($n = 53$) or the DA condition ($n = 54$) and participated in exchange for research credits. One participant in the FA condition was dropped for giving only one type of response to all test probes. Five participants in the DA condition were dropped for failing to complete the secondary tone task. This resulted in a final sample size of 52 participants in the FA condition (age: $M = 19.02$ years, $SD = 1.99$) and 49 participants in the DA condition (age: $M = 19.16$ years, $SD = 2.64$). Most participants in each group self-identified as female (73.1% in the FA condition; 67.3% in the DA condition), and both groups were matched on years of education (FA: $M = 12.56$, $SD = 0.99$; DA: $M = 12.66$, $SD = 0.98$). We also compared the two groups in the present study with the older adult sample from Experiment 2 of Greene and Naveh-Benjamin (2020). Forty older adults (age: $M = 73.05$ years, $SD = 3.99$; 75% female) with no known cognitive impairments participated under FA conditions that were identical to the procedures described below, with the exception that, whereas most of the participants in the present study completed the experiment online, all older adults in Experiment 2 of Greene and Naveh-Benjamin (2020) participated in the laboratory.

We based our sample sizes on those used in Experiment 2 of Greene and Naveh-Benjamin (2020), which included 40 young and 40 older adult participants and was found to be well-powered by a Bayesian prior sensitivity analysis. Our sample sizes were slightly larger because we switched to online data collection about a third of the way through, due to the COVID-19 pandemic. This resulted in slightly more participants in each condition. Seventeen participants in the FA condition and 15 in the DA condition completed the study in the laboratory, and the remainder participated online. We examined whether there were major performance differences on the task between the laboratory and online samples and didn't find any (see Fig. S1 in the supplement).

Materials

We paired 84 faces from the FACES database (Ebner et al., 2010) with 84 scenes from a categorized scene pool (Konkle et al., 2010). The faces were all White faces with neutral expressions, appearing on a gray background, and consisted of an even number of young and older faces and of male and female faces. Twelve face–scene pairs appeared per block, across one practice and six experimental blocks. For the concurrent DA task, we used three tones varying in pitch—one low, one medium, and one high pitch—which have been used in previous DA paradigms (Naveh-Benjamin et al., 1998; Naveh-Benjamin et al., 2003). Stimuli were presented using E-Prime 2.0 software (Schneider et al., 2012) for participants tested in the laboratory and via PsyToolkit (Stoet, 2010, 2017) for participants tested online.

Procedure

All procedures were approved by the University of Missouri Institutional Review Board. Participants completed one practice and six experimental blocks, like the one depicted in Fig. 1. Each block began with a study phase, during which participants studied 12 unique face–scene pairs, one at a time, for 4 seconds each. Although no two pairs were identical, each block featured two scenes from the same category (e.g., two parks, two malls), each paired with a different face (e.g., Old Man–Park 1, Young Woman–Park 2). In addition, six of the scenes in a block came from a broader category (e.g., six nature scenes, such as two parks, two forests, two fields) and the other six scenes came from a different broader category (e.g., six indoor scenes, such as two dens, two kitchens, two bedrooms). Participants were told to study the pairs for a later memory test. Participants in the DA condition simultaneously completed an auditory choice reaction time (CRT) task, in which they attended to a series of tones and indicated whether a given tone was low-, medium-, or high-pitched by pressing the "v," "b," or "n" key, respectively. Tones were presented every 2 seconds during the study phase. Participants were instructed to respond as quickly and accurately as possible to the tones while also studying the pairs for the forthcoming memory test. Participants in the DA condition also completed a baseline phase of the auditory CRT task before the first block and after the last block of the experiment.

There was either a short or long delay between the study and test phases of each block. There were three short and three long delay blocks, which were randomly intermixed for each participant. In the short delay blocks, after the last pair was presented during the study phase, the word "Wait" appeared, centered on screen for 5 seconds, and was followed by a test prompt informing participants that the test phase was about to begin in 3 seconds. In the long delay blocks, participants were shown the name of cities, one at a time for 5 seconds each, and were instructed to provide either the state or country in which each city was located. The correct answer was then shown for 2 seconds. In total, this task spanned 5 minutes and was followed with a prompt, for 3 seconds, informing participants that the test phase was about to begin. Verbal materials were used for the geography task to avoid the creation of similarity-based interference on the visually presented materials (faces and scenes) in the main memory task.

During the test phase of each block, participants were shown 12 pairs at random, one at a time, and evenly distributed into Intact, Related, and Unrelated probes. Intact test pairs featured a face–scene pair that had previously been presented during the study phase. For example, the Intact probe in Fig. 1 shows the young man paired with the same lobby scene from the study phase. Related test pairs were recombined face–scene pairs in which the face from one pair was recombined with a similar scene from a different pair. For

example, the old woman in the Related probe in Fig. 1 was originally paired with a similar, but *different*, garden scene. There were two types of Unrelated pairs, which we termed Unrelated-Within and Unrelated-Opposite pairs, and both types of Unrelated pairs were also recombined face–scene pairs. Unrelated-Within pairs featured a scene switch from within the same *broader* category (e.g., the old man paired with an indoor dining room, when he had been paired with an indoor kitchen at study), whereas Unrelated-Opposite pairs featured a scene switch from the other broader category (e.g., the young woman paired with an indoor kitchen, when she had been paired with an outdoor garden at study). For each test pair, participants were instructed to indicate whether the pair was "intact," "related," or "unrelated" by selecting one of these labeled responses on the keyboard (for participants completing the task in the laboratory) or on the computer screen (for participants completing the task online). Participants were told to respond "intact" to any pair they thought was exactly the same as a pair from the study phase; to respond "related" to any pair that was highly similar, involving a scene switch within the same specific category (e.g., participants were given an example showing an old man appearing first with one airport and then with another airport and were told that this was a Related pair); and to respond "unrelated" to any pair that was not the same as or similar at a category-specific level to originally-studied pairs. Thus, even for Unrelated-Within pairs, the correct response was "unrelated," but higher rates of erroneous "intact" or "related" responses could be possible if participants retrieved only fuzzier representations of original pairs (e.g., remembering that the old man had been paired with an indoor scene, but not remembering whether it was a kitchen or a dining room).

Analysis

All analyses were implemented in a Bayesian statistical framework. Data and analysis scripts are publicly available (https://osf.io/cdx9z/). We analyzed performance differences between FA and DA young adult groups on the task at both short and long delays using hierarchical Bayesian logistic regression analyses, a more powerful technique than standard analyses (e.g., ANOVA applied to aggregated data), due to its ability to account for trial-level responses, nested within participants, and to model these responses with a more suitable distributional form (e.g., a binomial distribution to reflect the possibility of a correct or incorrect response on each trial) than the normal distribution assumed by ANOVA (e.g., Dixon, 2008). To better understand the effects of DA on specific/verbatim and gist memory, which are not observable by nature, requires a more sophisticated modeling technique, so accordingly, we also employed MPT modeling to estimate the contributions of these cognitive processes to task performance. Below, we describe these analyses in more detail.
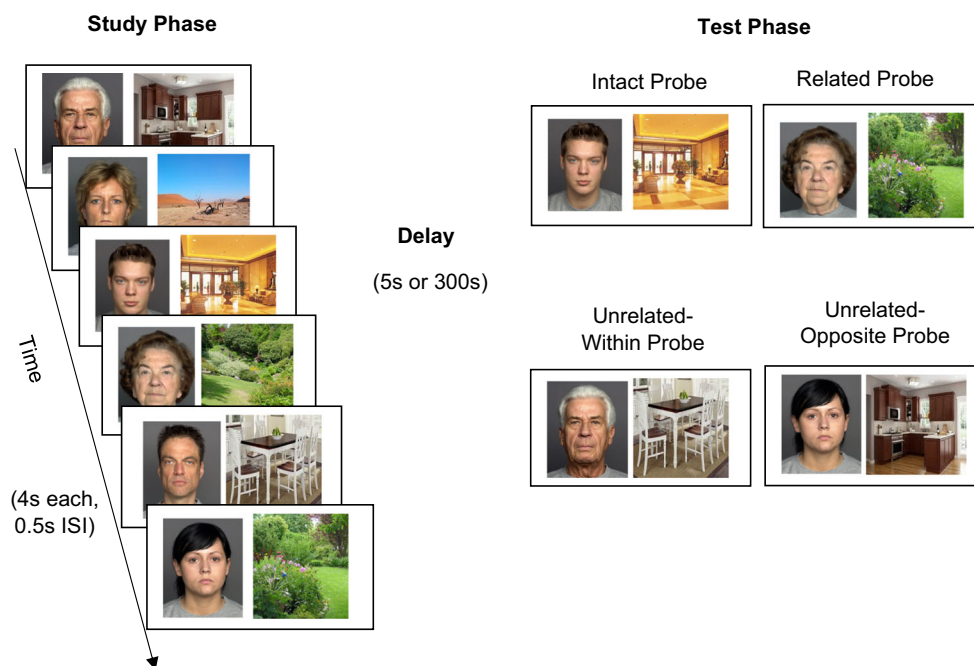
Study Phase

Test Phase

Intact Probe   Related Probe

Delay

(5s or 300s)

Unrelated-   Unrelated-
Within Probe   Opposite Probe

Time

(4s each,
0.5s ISI)

Fig. 1 Example of the procedure for one block. Participants studied 12 unique face–scene pairs for 4 seconds each (study phase; only 6 shown in Figure). Participants in the divided attention group simultaneously completed an auditory choice reaction-time task (see text for details). Then there was a delay of 5 seconds or 5 minutes, followed by the test phase, which featured Intact, Related, Unrelated-Within, and Unrelated-Opposite probes. Participants indicated whether each pair was "intact," "related," or "unrelated." Faces depicted in figure are approved for display for purposes of illustrating research methodology

Hierarchical Bayesian logistic regression analyses

We tested for Attention (coded as -1 = FA, 1 = DA), Delay (-1 = short, 1 = long), and Attention × Delay differences in the number of correct responses given to each probe with a series of hierarchical Bayesian logistic regression models implemented in the brms package for R (Bürkner, 2017; R Core Team, 2020). In each model, the response on trial i for subject j was recorded as 1 if the correct response was given (i.e., responding "intact" to Intact probes, "related" to Related probes, and "unrelated" to both types of Unrelated probes) and 0 if an incorrect response was given. For the two types of Unrelated probes, we also tested whether there were differences in accuracy between Unrelated-Within and Unrelated-Opposite probes by adding an effect-coded Probe (-1 = Unrelated-Opposite, 1 = Unrelated-Within) variable, plus its interactions with Attention and Delay. In addition to the analyses reported in the main text, we also tested for Sampling Site (-1 = Online, 1 = Laboratory) differences in accuracy. Analyses including Sampling Site are reported in the supplement and demonstrate no credible evidence for effects of Sampling Site (see Fig. S1).

Next, we analyzed error responses to each probe. For Intact probes, we coded for whether participants gave more "related" (coded as 1) than "unrelated" (coded as 0) responses. For Related probes, we coded for whether participants gave more "intact" (coded as 1) than "unrelated" (coded as 0) responses.

Finally, for both types of Unrelated probes, we coded for whether participants gave more "intact" (coded as 1) than "related" (coded as 0) responses. All error response models included effects of Attention and Delay, plus their interaction.

In logistic regression, the predicted probability of a correct response, $\hat{\pi}$, is modeled through the function $\text{logit}(\hat{\pi}) = \log(\hat{\pi} / (1-\hat{\pi}))$. Following the approach to parameter interpretation advocated by Kruschke (2011, 2018), we specified a region of practical equivalence (ROPE) around 0. As a slope of 0 corresponds to a change in $\hat{\pi} = 0.50$ (i.e., equal probability between the two levels of a factor), we considered a negligible change to be 0.50 ± 0.03 (see Kruschke, 2018).[1] This corresponds to a ROPE on the log-odds scale of [-0.06, 0.06]. If the 95% highest density interval (HDI) of the estimate excludes the ROPE, we conclude there is evidence for an effect. If the 95% HDI is entirely contained within the ROPE, we conclude there is evidence for a null effect. Finally, if the 95% HDI partially overlaps with the ROPE, but partially excludes it, we remain agnostic.[2]

All models included a random intercept and a random slope for Delay for each participant. We specified Cauchy (0, 2.5) priors on the population-level (i.e., "fixed" effects) slopes,

---

[1] This decision rule is based on the assumption that some values are not meaningfully different from zero. This is a more conversative decision rule than one based simply on whether an estimate differs from zero.

[2] Remaining agnostic is similar to a nonsignificant p value (i.e., $p > .05$).

based on recommendations in the literature (Gelman et al., 2008). We retained the program's default half-t priors on the standard deviations of the random effects and used an LKJ(1) prior, which assumes a uniform prior on the random effects correlation matrix (Lewandowski et al., 2009).

MPT (multinomial-processing-tree) analyses

Next, we used the MPT model from the simplified conjoint recognition paradigm (Stahl & Klauer, 2008), which was adapted to an associative recognition paradigm by Greene and Naveh-Benjamin (2020) and is depicted in Fig. 2. MPTs attempt to explain how participants arrive at their responses to given memory probe by way of latent cognitive processes (for reviews, see Batchelder & Riefer, 1999; Erdfelder et al., 2009).

The MPT model from the simplified conjoint recognition paradigm (Stahl & Klauer, 2008) has been empirically validated as a measurement tool for estimating the contributions of specific (i.e., verbatim) and gist memory in recognition tasks. The model has parameters corresponding to verbatim memory (parameters V_i and V_r) and gist memory (parameters G_i and G_r). In gist retrieval states, participants guess whether a probe is "intact" or "related" with probabilities a or $1 - a$, respectively. Participants may sometimes give an "intact" or "related" response to a probe even when they do not access gist memory, and this is modeled via parameter b, with subsequent guessing processes modeled by parameter a_b (guessing "intact" in this cognitive state) or $1 - a_b$ (guessing "related" in this cognitive state). The model presented here includes the additional retrieval of a less detailed, or fuzzier, representation, when participants are presented with Unrelated-Within probes. This retrieval state is modeled with parameter F, and it assumes a guessing process identical to the guessing processes that are modeled to occur in gist retrieval states.

As with any model, the MPT model presented in Fig. 2 is a simplification of reality, but it has been shown to provide good approximations of the processes of interest (i.e., verbatim, gist, and fuzzy retrieval) in previous research (Greene & Naveh-Benjamin, 2020; Stahl & Klauer, 2008). One simplification of the model is that correct "unrelated" responses to Unrelated-Within and Unrelated-Opposite probes are assumed to arise only through parameter b (specifically, pathway $1 - b$), which is a guessing parameter. While parameter b does measure guessing (i.e., a tendency to respond "intact" or "related" even when there is no specific *or* gist information), this does not necessarily mean that all correct "unrelated" responses arise from guesses. In fact, a more appropriate way to think of parameter b is that it indexes the probability that participants decide to guess "intact" or "related," but that when participants proceed down the complementary pathway $(1 - b)$, they have decided *not* to guess "intact" or "related." Consequently, some proportion of correct responses to

Unrelated probes may occur through guessing "unrelated," but some proportion likely occurs through knowledge that the probe is Unrelated. Thus, parameter b indexes the probability that a participant will decide "this *could* be an Intact or Related probe, even though I do not remember this face being paired with this or a similar scene."

The model has eight parameters corresponding to eight degrees of freedom and is thus saturated. Nevertheless, the model has been shown to be a useful measurement tool for estimating the contributions of specific and gist memory (Stahl & Klauer, 2008). While model fit is typically evaluated through χ^2 goodness-of-fit tests in a frequentist framework, in a Bayesian framework, model fit is evaluated by simulating data from the posterior distribution of the model parameters and comparing the posterior-predicted values to the observed frequencies, using posterior predictive p (PPP) values (Meng, 1994). Model fit was evaluated based on the T_1 and T_2 statistics proposed by Klauer (2010), measuring the correspondence between the posterior-predicted and observed means and covariances, respectively, and was considered satisfactory if PPP >.05. PPP values for the T_1 statistic were 0.56 for the DA group and 0.29 for the FA group, and for the T_2 statistic, PPP values were 0.38 and 0.45 for the DA and FA groups, respectively, indicating satisfactory model fit. Parameters of the MPT model were estimated under a hierarchical Bayesian latent-trait model using the TreeBUGS package for R (Heck et al., 2018; R Core Team, 2020; for more information on latent-trait models, see Klauer, 2010). Further details about sampling routines are discussed in the supplement. In Table S1 of the supplement, we show that parameters of the MPT model were comparable between the short and long delays, in both the FA and DA groups, so, in the main text, we report results collapsed across delay.

Results

Logistic regression results

Accuracy results

The proportion of responses given to each probe by participants in the FA and DA conditions in each delay is depicted in Fig. 3. Difference scores obtained by subtracting the posterior distribution, on the accuracy scale, of the DA from the FA groups in each delay are depicted in Fig. 4. As depicted in Fig. 4, most of the 95% HDIs of the difference scores were positive, indicating that accuracy was higher in the FA than DA group. Supporting this, the results of the hierarchical Bayesian logistic regression analyses provided credible evidence for an effect of Attention on response accuracy to each probe, except for Intact probes, for which the 95% HDI partially overlapped with the ROPE, such that we remained
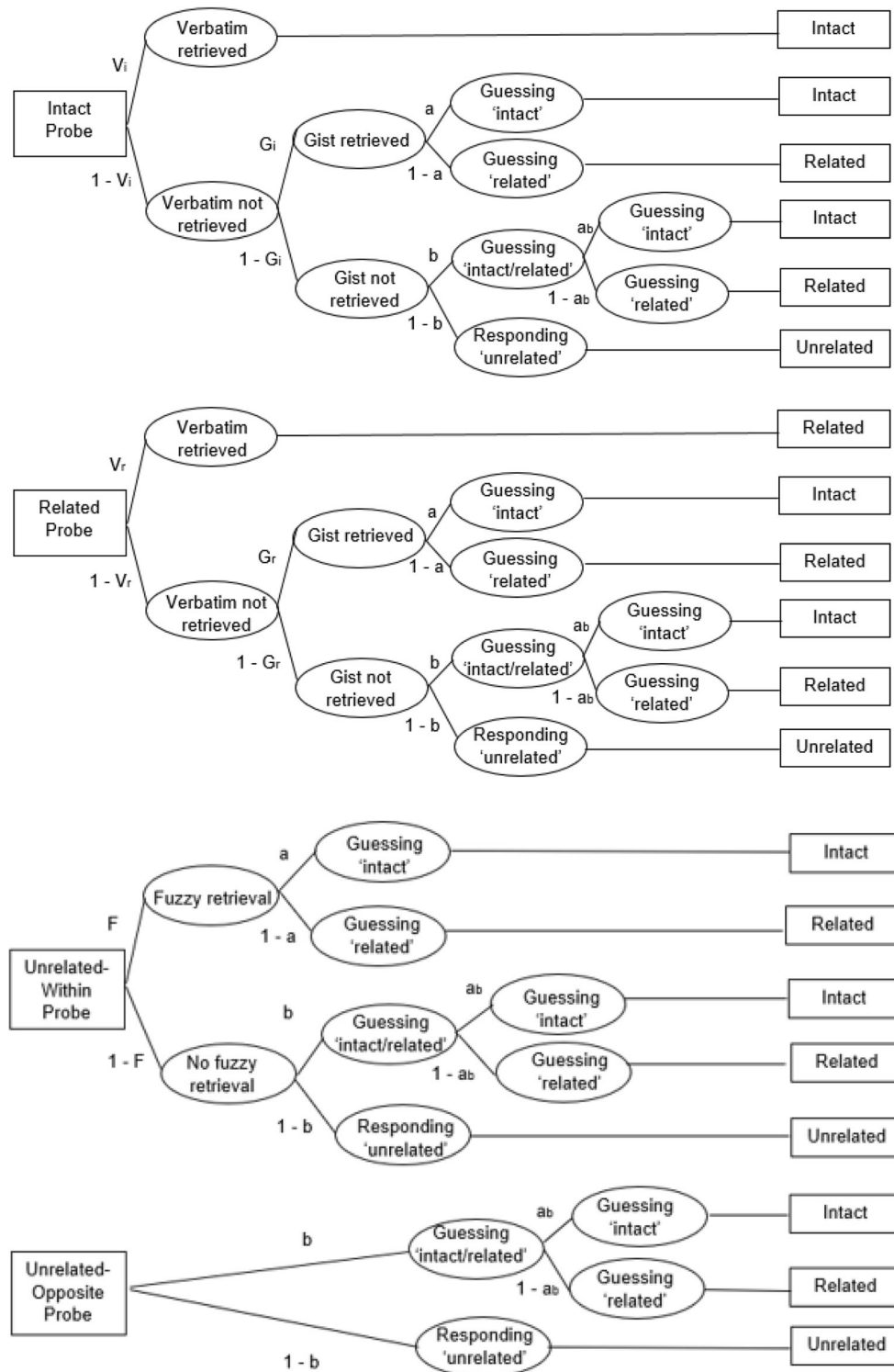
agnostic, even though the slope was in the direction of an effect (see Analyses section). The population-level (i.e., "fixed effects") slopes of Attention on the logit-scale were: for Intact probes, $\beta_{\text{Attention}} = -0.20$, 95% HDI [−0.38, −0.04]; for Related probes, $\beta_{\text{Attention}} = -0.26$, 95% HDI [−0.39, −0.12]; for Unrelated-Within probes, $\beta_{\text{Attention}} = -0.38$, 95% HDI [−0.61, −0.12]; and for Unrelated-Opposite probes, $\beta_{\text{Attention}} = -0.44$, 95% HDI [−0.71, −0.18].

As depicted in Fig. 4, the differences in response accuracy between the FA and DA groups were similar in each delay. For all probes, the 95% HDI for the slope of the Attention × Delay interaction partially overlapped with the ROPE, such

◀ **Fig. 2** Expanded version of the multinomial processing tree model from the simplified conjoint recognition paradigm (Stahl & Klauer, 2008), here including separate trees for different types of Unrelated probes, from Greene and Naveh-Benjamin (2020). Boxes on the left represent memory probes, which are connected to participants' responses (boxes on the right) by way of different cognitive processes (the branches in the middle, where the ovals describe to what each parameter corresponds). The two V parameters correspond to the probability that participants retrieve the verbatim representation of an association given either an Intact probe (V_i) or a Related probe (V_r). The two G parameters correspond to the conditional probabilities that participants retrieve the gist of an association for Intact probes (G_i) or Related probes (G_r), given that they have not retrieved more specific representations. Parameter F corresponds to the probability that participants retrieve a fuzzier representation given an Unrelated-Within probe. If participants retrieve the gist or a fuzzier representation, they then guess whether the probe is "intact" (with probability a) or "related" (with probability $1 - a$). If a probe elicits no verbatim or gist information for a participant, then the participant can still guess that the probe is "intact" or "related" with probability b, followed by guessing "intact" (probability a_b) or "related" (probability $1 - a_b$). Otherwise, participants respond "unrelated" with probability $1 - b$

that we remained agnostic as to whether the Attention differences in response accuracy reported above depended on whether tests occurred after a short or long delay. The population-level slopes for the Attention × Delay interaction were: for Intact probes, $\beta_{\text{Attention} \times \text{Delay}} = -0.04$, 95% HDI [−0.13, 0.05]; for Related probes, $\beta_{\text{Attention} \times \text{Delay}} = -0.09$, 95% HDI [−0.18, −0.01], which is in the direction of a more pronounced effect of DA on these probes following a longer delay; for Unrelated-Within probes, $\beta_{\text{Attention} \times \text{Delay}} = 0.11$, 95% HDI [−0.02, 0.24]; and for Unrelated-Opposite probes, $\beta_{\text{Attention} \times \text{Delay}} = 0.06$, 95% HDI [−0.09, 0.20].

There was, however, credible evidence for an effect of Delay on response accuracy to Intact probes, $\beta_{\text{Delay}} = -0.27$, 95% HDI [−0.36, −0.17]. As shown in Fig. 3, participants in both the FA and DA groups were more accurate at classifying these probes in the short than long delay blocks. For all other probes, we remained agnostic about effects of Delay, as the 95% HDI partially overlapped with the ROPE. Estimated
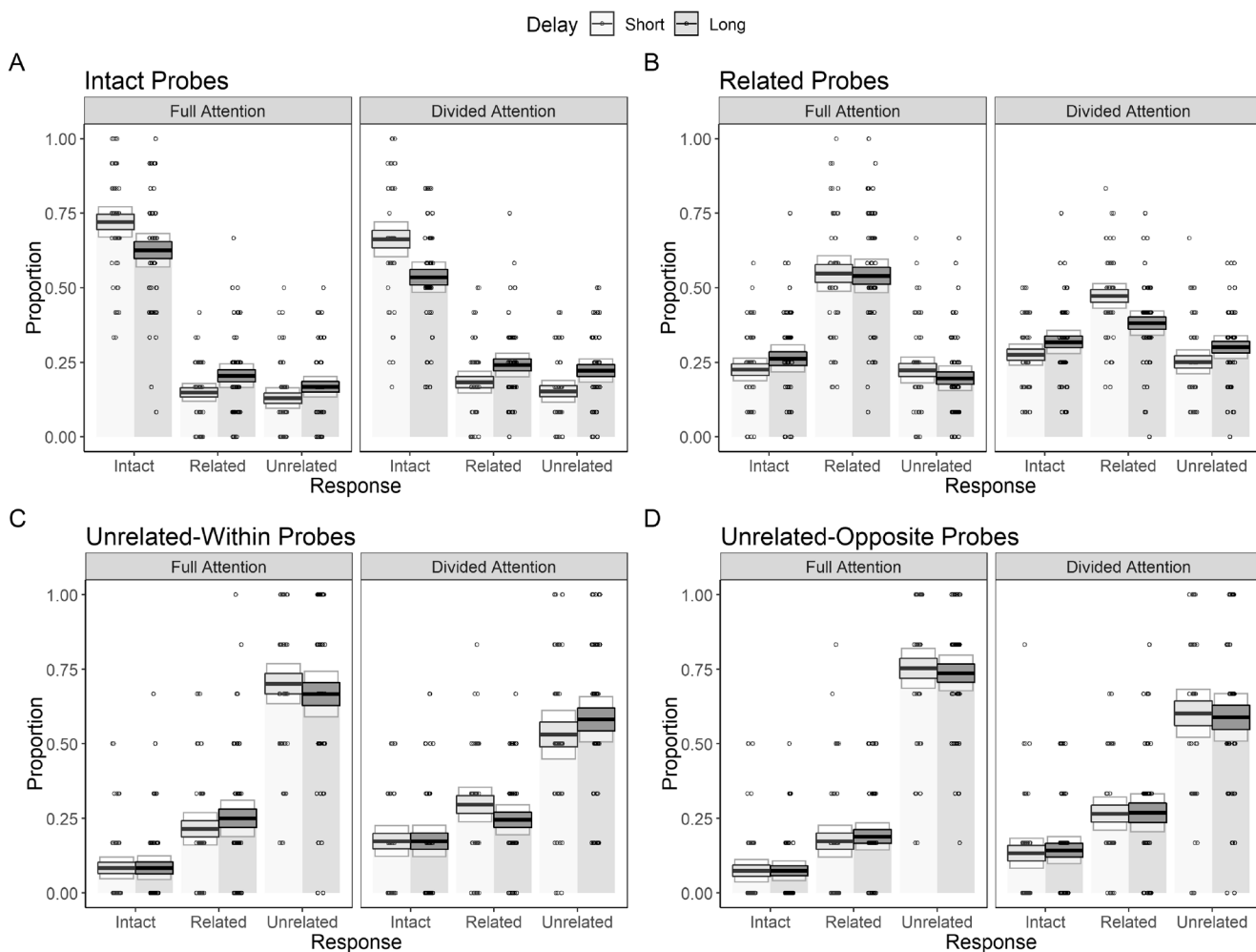

Fig. 3 Proportion of "intact," "related," and "unrelated" responses to Intact probes (**a**), Related probes (**b**), Unrelated-Within probes (**c**), and Unrelated-Opposite probes (**d**) for the full and divided attention groups in the short and long delays. Lines at the top of the vertical bars denote group means. Shaded box around the means denote ±1 standard error. White box around the standard error represents the 95% confidence interval. Jittered points denote individual participants' data
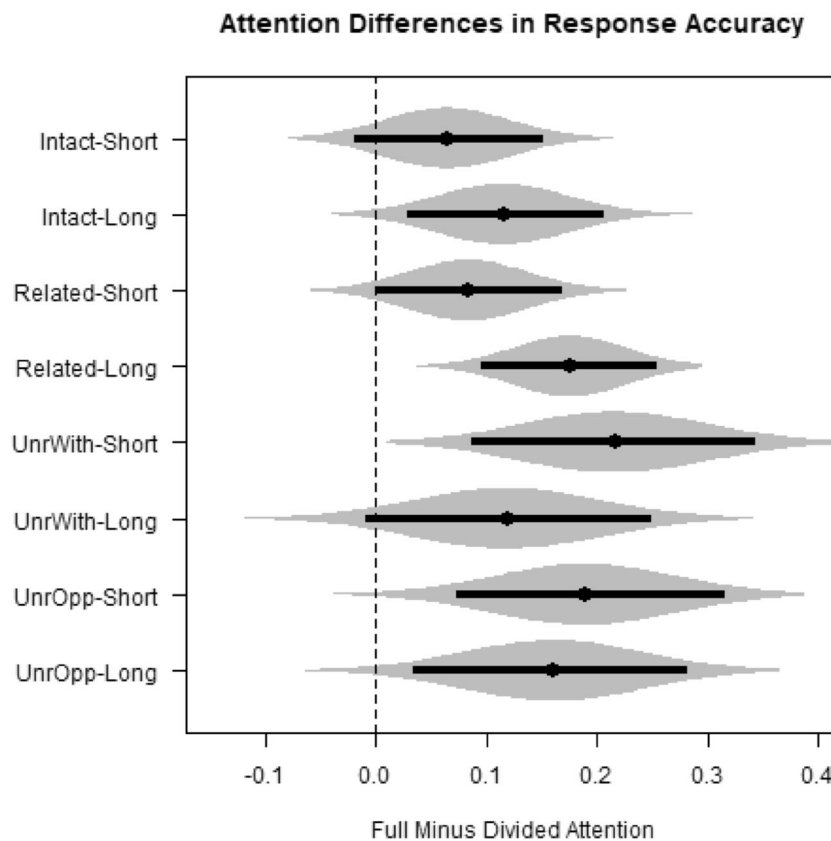
Attention Differences in Response Accuracy

Fig. 4 Violin plots of posterior density, transformed to the accuracy scale, depicting the difference in response accuracy for each probe in each delay between the full and divided attention groups. Points corresponds to the posterior mean; solid black lines correspond to the 95% highest density interval. UnrWith = Unrelated-Within probes; UnrOpp = Unrelated-Opposite probes

slopes for Delay were: for Related probes, $\beta_{\text{Delay}} = -0.10$, 95% HDI $[-0.19, -0.02]$; for Unrelated-Within probes, $\beta_{\text{Delay}} = 0.03$, 95% HDI $[-0.11, 0.15]$; and for Unrelated-Opposite probes, $\beta_{\text{Delay}} = -0.11$, 95% HDI $[-0.26, 0.04]$.

Error Responses

The proportion of error responses to each probe are also shown in Fig. 3. Regression coefficients from the logistic regression analyses examining differences in error responses are given in Table 1. The Intercept corresponds to the grand mean. All effects of Attention, Delay, and the interaction of Attention x Delay overlapped with the ROPE, such that we remained agnostic as to whether there were any differences in error response tendency between different levels of these factors. For Intact probes, we assessed whether participants were more inclined to respond "related" rather than "unrelated," and the intercept suggests a somewhat greater tendency to do so, with an odds ratio (*OR*) of 1.19, 95% HDI [1.03, 1.38] in favor of responding "related" rather than "unrelated." For Related probes, we analyzed whether participants were more inclined to respond "intact" rather than "unrelated," but the 95% HDI of the intercept was partially positive and partially negative, providing an inconclusive *OR* of 1.11, 95%

HDI [0.95, 1.27] in favor of responding "intact" rather than "unrelated." For both types of Unrelated probes, we analyzed whether participants were more likely to mistakenly endorse these probes as "intact" rather than "related," but there was actually credible evidence that participants were more likely to respond "related" (see Table 1), with *OR*s in favor of responding "related" rather than "intact" of 2.20, 95% HDI [1.65, 2.92], for the Unrelated-Within probes and 2.32, 95% HDI [1.75, 3.16], for the Unrelated-Opposite probes.

Interim summary

To summarize, the DA group performed worse than the FA group for all probes with the exception of Intact probes, though the evidence was marginally in favor of an effect of Attention on these probes, as well. Attention differences in response accuracy did not meaningfully depend on Delay, and the only credible evidence for a Delay effect was for response accuracy to Intact probes, which was lower in the long- than short-delay blocks in both attention groups. An analysis of errors revealed that participants in both the FA and DA groups were somewhat more likely to mistakenly call an Intact probe "related" rather than "unrelated," and to call Unrelated-Within and Unrelated-Opposite probes "related"

Table 1 Population-level (fixed effects) intercepts and slopes [95% highest density intervals] for error response models

Probe	Intercept	Attention	Delay	Attention x Delay
Intact	β = 0.17 [0.03, 0.32]	β = −0.04 [−0.13, 0.12]	β = 0.01 [−0.13, 0.14]	β = −0.05 [−0.19, 0.09]
Related	β = 0.10 [−0.05, 0.24]	β = −0.02 [−0.16, 0.12]	β = 0.04 [−0.08, 0.16]	β = −0.07 [−0.18, 0.05]
Unrelated-Within	**β = −0.78 [−1.07, −0.50]**	β = 0.24 [0.00, 0.52]	β = 0.04 [−0.18, 0.28]	β = 0.05 [−0.16, 0.27]
Unrelated-Opposite	**β = −0.84 [−1.15, −0.56]**	β = 0.10 [−0.17, 0.37]	β = −0.03 [−0.27, 0.21]	β = 0.10 [−0.15, 0.33]

Note. Intercept corresponds to the grand mean. For Intact probes, we coded for whether participants gave more "related" than "unrelated" responses. For Related probes, we coded for whether participants gave more "intact" than "unrelated" responses. For both types of Unrelated probes, we coded for whether participants gave more "intact" than "related" responses. Bolded values correspond to credible evidence in favor of one response over the other as these values exclude the region-of-practical equivalence

rather than "intact." However, for Related probes, the evidence was inconclusive as to whether participants were more prone to mistakenly calling these probes "intact" rather than "unrelated."

MPT results

Parameter estimates collapsed across delay conditions, for the FA and DA groups are reported in Table 2.

Figure 5 shows the posterior distributions of each parameter, transformed to the probability scale. Parameters whose distributions mostly overlap with each other do not meaningfully differ between groups. It appears that parameters V_i (verbatim retrieval for Intact probes) and G_r (gist retrieval for Related probes) were smaller in the DA than FA groups, while all three guessing parameters appeared to be mostly larger in the DA groups. To confirm these visual trends, we computed difference scores by subtracting the posterior distributions of

each parameter of the DA group from the FA group. Difference scores are depicted in Fig. 6. Parameters whose 95% Bayesian credible interval (CI) of the difference score excludes 0 credibly differ between the DA and FA groups (Smith & Batchelder, 2010). There were group differences in parameter V_i, $\Delta V_i = 0.26$, 95% CI [0.02, 0.56], indicating that the DA group had lower estimates of verbatim retrieval for Intact probes. The DA group also had lower estimates of gist retrieval for Related probes, $\Delta G_r = 0.25$, 95% CI [0.08, 0.40]. For guessing parameters, there was clear evidence for a group difference in parameter b, $\Delta b = -0.17$, 95% CI [−0.26, −0.07], indicating that the DA group had a higher tendency to guess "intact" or "related" when no verbatim or gist information was present or retrieved. The difference scores for parameters a, $\Delta a = -0.23$, 95% CI [−0.49, 0.05], and a_b, $\Delta a_b = -0.12$, 95% CI [−0.23, 0.00], overlapped with and straddled 0, respectively. Therefore, we cannot definitively rule out the possibility that there was no difference in these guessing parameters between the FA and DA groups, but most of the 95% CI for the difference scores for these parameters was negative, suggesting the DA group had a greater tendency to guess "intact." All other parameters did not credibly differ between groups.

Full and divided attention young adults versus older adults

We compared the MPT parameter estimates of the DA and FA young adult groups in the present study to those obtained from the older adults, who completed the study phase under FA, in Experiment 2 of Greene and Naveh-Benjamin (2020). We computed difference scores by subtracting the posterior samples of the older adults from the DA group and the FA group, separately (see Fig. 7). Whereas older adults were deficient in verbatim memory retrieval for Intact probes relative to FA young adults, in line with the results of Greene and Naveh-Benjamin (2020), estimates of verbatim retrieval were not different between the older adults and the young adults under DA from the present study. However, older adults had higher estimates of gist retrieval given Related probes (parameter G_r) than the DA young adults, $\Delta G_r = -0.16$, 95% CI [−0.30,

Table 2 Population-level parameter estimates [95% credible intervals] of the MPT model

Parameter	Full Attention	Divided Attention
Vi	0.55 [0.38, 0.65]	0.29 [0.02, 0.48]
Vr	0.09 [0.00, 0.31]	0.13 [0.01, 0.25]
Gi	0.53 [0.43, 0.65]	0.48 [0.29, 0.67]
Gr	0.67 [0.53, 0.75]	0.42 [0.31, 0.54]
F	0.05 [0.01, 0.12]	0.06 [0.01, 0.14]
a	0.40 [0.30, 0.63]	0.64 [0.42, 0.85]
ab	0.18 [0.10, 0.27]	0.30 [0.22, 0.37]
b	0.22 [0.16, 0.28]	0.39 [0.32, 0.45]

Note. Vi = probability of verbatim retrieval given an Intact probe; Vr = probability of verbatim retrieval given a Related probe; Gi = probability of gist retrieval given an Intact probe; Gr = probability of gist retrieval given a Related probe; F = probability of fuzzier retrieval given an Unrelated-Within probe; a = probability of guessing "intact" when gist is retrieved; ab = probability of guessing "intact" when no verbatim or gist information is retrieved; b = probability of guessing "intact/related" when there is no verbatim or gist information

Population-Level Parameter Estimates

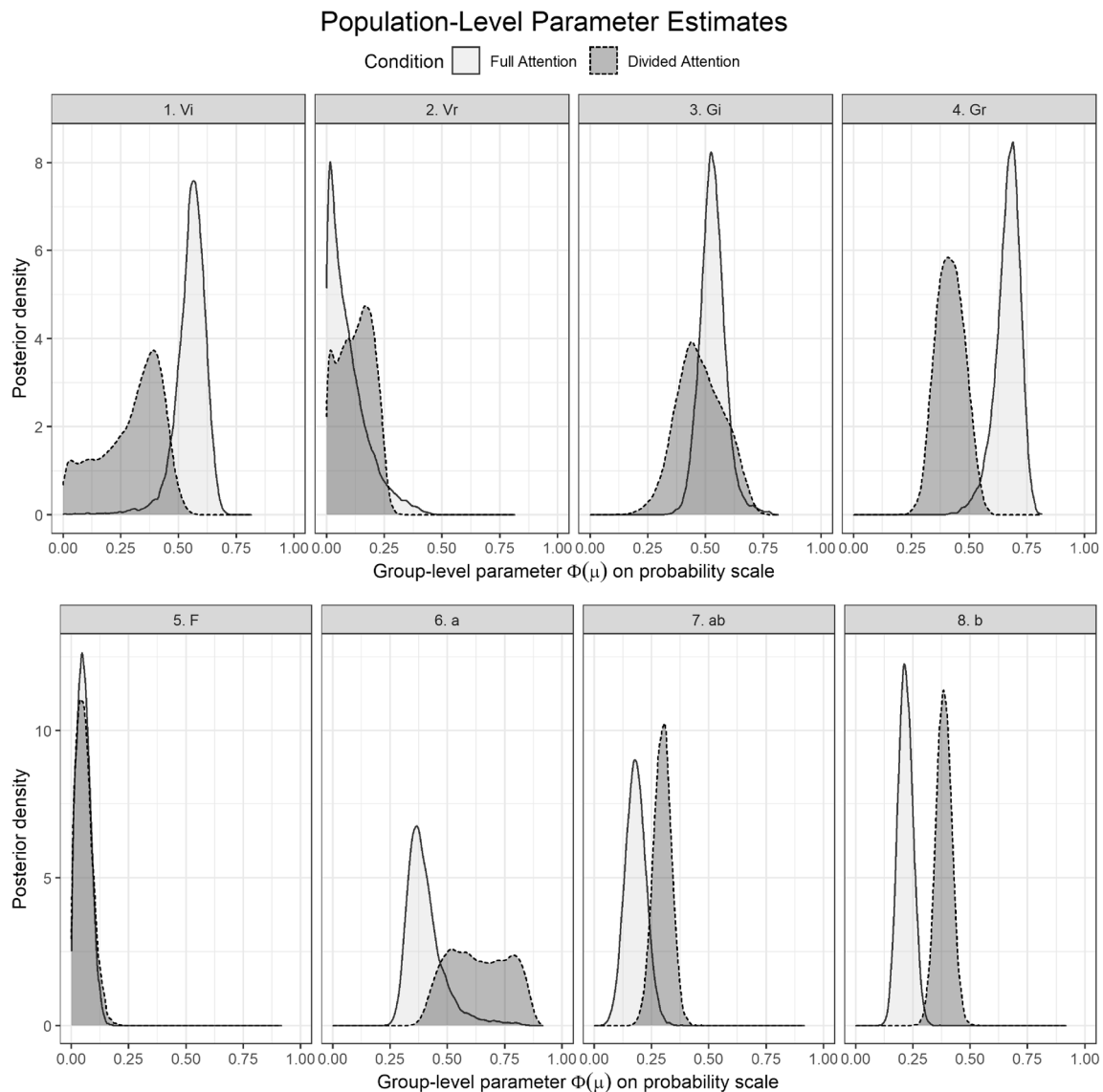Condition ☐ Full Attention   ▨ Divided Attention

Fig. 5 Posterior distributions of the inverse-probit transformed group-level parameters on the probability scale. Parameter descriptions are given in Fig. 2. Parameters 1–5 (Vi, Vr, Gi, Gr, and F) are memory parameters. Parameters 6–8 (a, ab, and b) are guessing parameters

−0.02]. Finally, the only guessing parameter that definitively differed between the older adults and the DA young adults was parameter b, $\Delta b = 0.11$, 95% CI [0.01, 0.21], indicating that the DA young adults had a greater tendency to respond "intact/related" in states in which verbatim or gist information was not retrieved.

Interim summary

To summarize the MPT results, DA at encoding resulted in lower estimates of verbatim retrieval for Intact probes and gist retrieval for Related probes, compared with FA at encoding. Comparing these results with those of older adults from Experiment 2 of Greene and Naveh-Benjamin (2020) also revealed that the DA young adults had lower estimates of gist retrieval for Related probes than did older adults.

Secondary task performance

Finally, we examined performance on the auditory CRT task. For each participant in the DA group, we computed their average accuracy and reaction time (RT) on the task, at both the baseline periods and during the study phases. We used Bayesian paired-samples t tests, implemented using the BayesFactor package for R (Morey & Rouder, 2015; R Core Team, 2020). The resulting Bayes factor (BF_{10}) provides the strength of evidence in favor of the alternative hypothesis (corresponding to a difference between the baseline and study phases) to the null hypothesis. Accuracy was high and did not meaningfully differ between the baseline ($M = 0.91$) and Study ($M = 0.89$) phases of the experiment, $BF_{10} = 0.75$, but as usually shown in DA experiments (e.g., Craik et al., 1996), participants were faster to respond during the baseline ($M =$
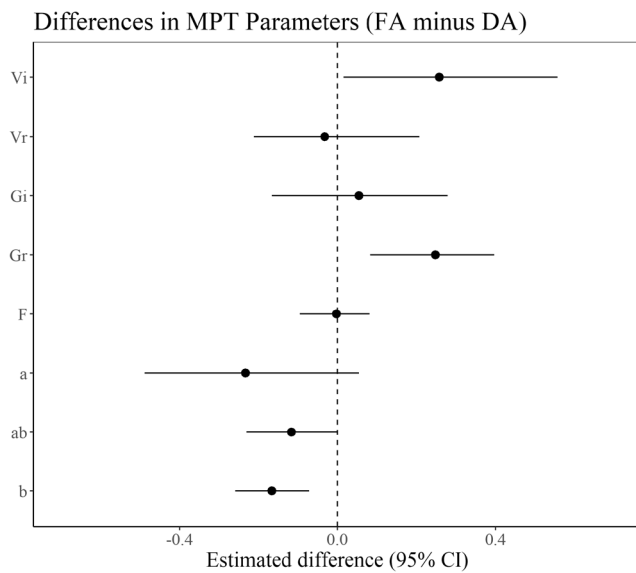
Differences in MPT Parameters (FA minus DA)

Fig. 6 Forest plot of difference scores for each parameter obtained by subtracting the posterior samples of the DA group from the FA group. Points correspond to the Bayesian posterior mean and lines denote the 95% Bayesian credible interval. Dashed line at 0.0 corresponds to no difference. Parameters whose difference scores overlap with zero do not meaningfully differ between groups. Parameter descriptions given in Fig. 2

779.45ms) than the Study ($M = 910.79$ms) phases, $BF_{10} = 8.58 \times 10^{10}$.

Discussion

We examined whether DA at encoding would disrupt specific and gist memory for associations in episodic memory. Young adults who encoded face–scene pairs under DA performed more poorly than young adults who encoded those pairs under FA on all types of test probes in an associative recognition task assessing memory for both highly specific and gist representations. DA effects were observed following both short and long delays between the study and test phases of our experiment, showing that the effects of DA may emerge early on and endure across a delay of up to 5 minutes. In addition, fits of an MPT model revealed that DA young adults had lower estimates of verbatim and gist memory than FA young adults, being less likely to remember the specific association when shown an Intact probe or to retrieve the gist of an association when shown a Related probe. These results provide perhaps the most concrete evidence to date that the effects of DA at encoding extend to multiple levels of specificity for episodic memories.

Reexamining the effects of divided attention at different levels of specificity

Numerous studies have demonstrated that DA at encoding results in deficits in memory performance (e.g., Baddeley
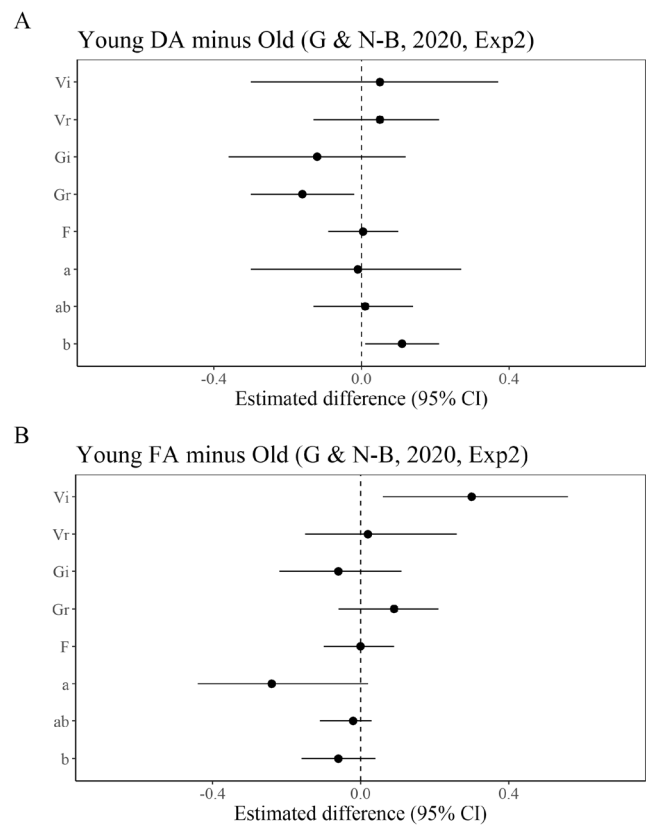

Fig. 7 **a** Forest plot of difference scores for each parameter obtained by subtracting the posterior samples of the older adults in Experiment 2 of Greene and Naveh-Benjamin (2020) from the divided attention (DA) young adults in the present experiment. **b** Difference scores subtracting older adults in Experiment 2 of Greene and Naveh-Benjamin (2020) from full attention (FA) young adults in the present experiment. Points correspond to the Bayesian posterior mean and lines denote the 95% Bayesian credible interval. Dashed line at 0.0 corresponds to no difference. Parameters whose difference scores overlap with 0 do not meaningfully differ between groups. Parameter descriptions given in Fig. 2

et al., 1984; Castel & Craik, 2003; Craik et al., 2018; Craik et al., 1996; Kilb & Naveh-Benjamin, 2007; Murdock, 1965; Naveh-Benjamin et al., 1998; Naveh-Benjamin et al., 2003; Nieznański, 2013). Some research on the effects of DA on false memory production, using DRM procedures (Deese, 1959; Roediger & McDermott, 1995), has found that DA results in increased false recall but decreased false recognition of unpresented lures (Dewhurst et al., 2005; Dewhurst et al., 2007; Knott & Dewhurst, 2007; Knott et al., 2018; Pérez-Mata et al., 2002). These findings have been taken to suggest that DA at encoding prevents participants from generating semantic associates of target words during study. However, these studies have consistently observed reduced rates of "old" responses to both old items and lures, and these reductions in "old" responses only occurred for judgments accompanied by recollection of the encoding context (Dewhurst et al., 2005; Dewhurst et al., 2007). According to fuzzy-trace theory, recollection can be obtained from either verbatim or

gist retrieval (Brainerd et al., 2014; Brainerd et al., 1999), so it is unclear from these earlier studies whether DA disrupted verbatim or gist memory. In another study, Odegard and Lampinen (2005) measured item memory using a conjoint recognition procedure (Brainerd et al., 1999) and found that DA at encoding impaired verbatim memory (specifically, recollection rejection of related lures) but not gist memory.

Our results go beyond these earlier studies by probing the effects of DA on specific and gist memory for associations, which lie at the core of episodic memory (e.g., Tulving, 1983). Results from the present study shed important insight into the effects of DA at multiple levels of specificity by showing that DA at encoding impairs participants' ability to later remember not just highly specific details but also gist details. We found that participants who encoded face–scene pairs under DA were less capable of remembering specific, verbatim information about these pairs when re-presented the same pairs (as Intact pairs) during the test phases and were also less capable of remembering gist information about these pairs when presented with highly similar foils (Related pairs) at retrieval. This finding is compatible with the *encoding-specificity hypothesis* (Tulving & Thompson, 1973), as DA effects were most noticeable on memory representations that are most easily accessed by a given probe (verbatim representations for Intact probes, and gist representations for Related probes).

Dodson et al. (1998) investigated the effects of DA at *retrieval* on specific- and partial-source memory and found that participants who completed the retrieval phase under DA were deficient at remembering specifically which voice spoke a given word (e.g., "was it this female or that female?") but had preserved partial-source memory (i.e., remembering whether a word had been spoken by a male or female voice), which suggests that DA at retrieval may affect specific, but not gist-based information of complex episodic memories. However, DA at encoding has been shown to produce more pronounced effects on memory than DA at retrieval does (Craik et al., 1996), including for item and context information (Greene et al., in press; Nieznański, 2013). Our findings that DA at encoding impaired both specific and gist associative memory provide further support that the effects of DA at encoding are more pronounced than those of DA at retrieval.

Our findings are also compatible with recent research on individual differences in verbatim and gist memory and inhibition by Nieznański and Obidziński (2019), who showed that working memory capacity was positively associated with both verbatim and gist memory. DA at encoding should result in reduced capacity to encode information, as participants must allocate their limited working memory resources to process the face–scene pairs while simultaneously attending to auditorily presented stimuli.

Potential underlying mechanisms

What encoding mechanisms may have been disrupted to lead to these specific and gist memory deficits? One possibility is that DA at encoding results in less elaborative processing of associations, resulting in a shallower episodic representation. However, this interpretation is incompatible with studies showing that DA at encoding impaired both item and associative memory performance to the same degree under intentional and incidental learning instructions (Naveh-Benjamin & Brubaker, 2019; Naveh-Benjamin et al., 2014). That is, if DA disrupted elaborative processing (such as effortful strategy use), then effects of DA should be more pronounced when DA and FA groups of participants are given intentional, explicit instructions to encode the material, rather than incidental instructions in which participants are not made aware of a forthcoming memory test. However, Naveh-Benjamin and Brubaker (2019) found that, although incidental learning resulted in lower recognition performance than intentional learning and that DA disrupted memory performance more than FA, the DA effect was no larger in the intentional than incidental conditions.

An alternative possibility is that DA disrupts the initial registration of the stimuli, and in particular the association between two stimuli (such as a between a face and a scene in our paradigm), as this phase of encoding has been shown to be especially vulnerable to interference from a concurrent task (Naveh-Benjamin et al., 2007). Interestingly, Odegard and Lampinen (2005) found no effects of DA on gist memory for *items*, which could reflect that the gist of an item forms even before attention is fully deployed to registering that item (e.g., Wallace et al., 2000; Wallace et al., 1998). In contrast, we found that DA did disrupt gist memory for associations. Thus, it could be that forming the gist of an association (such as that between a face and a scene) does require more initial registration-based attentional resources than those needed for encoding the gist of an item, and as such would not be registered as well under DA, also causing the interruption of the verbatim information.

Reexamining the depleted attentional resources hypothesis of cognitive aging

Another motivation for the present study was to test whether depleted attentional resources could be one mechanism accounting for older adults' deficits in specific associative memory (Greene & Naveh-Benjamin, 2020). There is a long history of research comparing performance of young adults under DA with older adults, including on studies of item and associative recognition (Castel & Craik, 2003; Craik et al., 2010; Kilb & Naveh-Benjamin, 2007; Naveh-Benjamin et al., 2003), and this research has resulted in a somewhat mixed set of findings. Whereas Castel and Craik (2003) found

that DA in young adults impaired associative memory performance, but not item recognition, resulting in a similar associative deficit observed in older adults (Naveh-Benjamin, 2000), others have found that DA produces a more "general" deficit, affecting item and associative information (e.g., Naveh-Benjamin et al., 2003). Our findings are more in line with this latter set of studies, as we found that DA in young adults resulted in impairments in verbatim and gist memory, whereas aging was associated only with deficits in verbatim memory (Greene & Naveh-Benjamin, 2020).

Nevertheless, the present set of findings cannot definitively rule out a depleted attentional resources hypothesis of age-related cognitive decline. Indeed, the young adults under DA had similar estimates of verbatim retrieval as older adults. It is plausible that older adults may have diminished attentional resources that lead to their deficits in associative episodic memory at the highest levels of specificity. However, older adults may have enough preserved attentional resources that enable them to encode enough sufficient information about associations to later remember the gist of these associations. This idea is compatible with research suggesting that older adults rely on a gist-based processing strategy during encoding (Tun et al., 1998; for a recent review, see Devitt & Schacter, 2016).

Limitations

The present study is not without its limitations. First, due to the COVID-19 pandemic, we switched to online data collection about a third of the way through our participant recruitment. Thus, participants were not all tested in the same, standardized environment. Nevertheless, there was no credible evidence for differences in response accuracy between the lab-based and online samples (see Fig. S1 in the supplement).

Second, the discrete-state assumption underlying MPT models has been the subject of some controversy (Batchelder & Alexander, 2013; Bröder & Schütz, 2009; Dube & Rotello, 2012; Klauer & Kellen, 2011a, 2011b; Pazzaglia et al., 2013; Province & Rouder, 2012). Particularly in studies of item recognition, analyses based on receiver operating characteristic (ROC) curves tend to favor models based on signal detection theory more than MPT models (Dube & Rotello, 2012; Pazzaglia et al., 2013), although this is not always the case (Klauer & Kellen, 2011a). However, for associative memory, discrete-state models often capture the ROC form better (Rotello, 2017; Yonelinas, 1997). Nevertheless, the specific modeling approaches employed can affect the interpretation and conclusions that can be drawn (e.g., Pazzaglia et al., 2013).

Third, we focused on memory for face–scene pairs to make our task ecologically valid by attempting to simulate remembering where someone was encountered (see Gruppuso et al.,

2007). However, similarities exist at both a conceptual level (e.g., two park scenes are both nature scenes) and at a perceptual level (e.g., two park scenes look physically similar), and as such, it is unclear whether DA effects on the episodic gist of an association (e.g., remembering whether an old man had been paired with a park) result from deficits in memory for the perceptual or conceptual representation of the association, or both.

Conclusions

In conclusion, results from the present study show that divided attention at encoding impairs specific and gist memory for associations that lie at the core of episodic memory. These findings suggest that the effects of divided attention are more general than those of aging, which is associated with deficits in verbatim, but not gist-based, memory for associations.

Declarations

Ethics approval This study was approved by the University of Missouri Institutional Review Board, IRB Project Number 1141837, "Binding of memories: Adult age differences and the effects of divided attention in young on episodic memory."

Consent to participate All participants provided informed consent prior to participating.

Consent for publication All authors are listed and have approved this manuscript for submission.

Conflicts of Interest We have no conflicts of interest to report.

References

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

Batchelder, W. H., & Alexander, G. E. (2013). Discrete-state models: Comment on Pazzaglia, Dube, and Rotello (2013). *Psychological Bulletin, 139*(6), 1204–1212. https://doi.org/10.1037/a0033894

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*, 57–86. https://doi.org/10.3758/BF03210812

Brainerd, C. J., Gomes, C. F. A., & Moran, R. (2014). The two recollections. *Psychological Review, 121*(4), 563–599. https://doi.org/10.1037/a0037668

Brainerd, C. J., & Reyna, V. F. (2015). Fuzzy-trace theory and lifespan cognitive development. *Developmental Review, 38*, 89–121. https://doi.org/10.1016/j.dr.2015.07.006

Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review, 106*(1), 160–179. https://doi.org/10.1037/0033-295X.106.1.160

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110*(4), 762–784. https://doi.org/10.1037/0033-295X.110.4.762

Bröder, A., & Schütz, J. (2009). Recognition ROCs are curvilinear—Or are they? On premature arguments against the two-high-threshold model of recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(3), 587–606. https://doi.org/10.1037/a0015279

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging, 18*(4), 873-885. https://doi.org/10.1037/0882-7974.18.4.873

Castel, A. D., Farb, N. A. S., & Craik, F. I. M. (2007). Memory for general and specific value information in younger and older adults: Measuring the limits of strategic control. *Memory & Cognition, 35*(4), 689–700. https://doi.org/10.3758/BF03193307

Craik, F. I. M. (2002). Levels of processing: Past, present … and future? *Memory, 10*(5-6), 305-318. https://doi.org/10.1080/09658210244000135

Craik, F. I. M. (2006). Remembering items and their contexts: Effects of ageing and divided attention. In H. D. Zimmer, A. Meclinger, & U. Lindenberger (Eds.), *Handbook of binding and memory: Perspectives from cognitive neuroscience* (pp. 571-594). Oxford University Press.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp. 191–211). Plenum Press.

Craik, F. I. M., Eftekhair, E., & Binns, M. A. (2018). Effects of divided attention at encoding and retrieval: Further data. *Memory & Cognition, 46*, 1263–1277. https://doi.org/10.3758/s13421-018-0835-3

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125*, 159–180. https://doi.org/10.1037//0096-3445.125.2.159

Craik, F. I. M., Luo, L., & Sakuta, Y. (2010). Effects of aging and divided attention on memory for items and their contexts. *Psychology and Aging, 25*(4), 968–979. https://doi.org/10.1037/a0020276

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22. https://doi.org/10.1037/h0046671

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia, 91*, 346–359. https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Dewhurst, S. A., Barry, C., & Holmes, S. J. (2005). Exploring the false recognition of category exemplars: Effects of divided attention and explicit generation. *The European Journal of Cognitive Psychology, 17*, 803–819. https://doi.org/10.1080/09541440540000013

Dewhurst, S. A., Barry, C., Swannell, E. R., Holmes, S. J., & Bathurst, G. L. (2007). The effect of divided attention on false memory depends on how memory is tested. *Memory & Cognition, 35*(4), 660–667. https://doi.org/10.3758/BF03193304

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59*, 447–456. https://doi.org/10.1016/j.jml.2007.11.004

Dodson, C. S., Holland. P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1121–1136. https://doi.org/10.1037/0278-7393.24.5.1121

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*(3), 286–303. https://doi.org/10.1037/0096-3445.127.3.286

Dube, C., & Rotello, C. M. (2012). Binary ROCs in perception and recognition memory are curved. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 130–151. https://doi.org/10.1037/a0024957

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavioral Research Methods, 42*, 351–362. https://doi.org/10.3758/BRM.42.1.351

Erdfelder, E. Auer, T. S., Hilbig, B. E., Aβfalg, A. Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology, 217*, 108–124. https://doi.org/10.1027/0044-3409.217.3.108

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*, 309–313. https://doi.org/10.3758/BF03197041

Gelman, A., Jakulin, A., Pittaue, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic regression. *The Annals of Applied Statistics, 2*(4), 1360–1383. https://projecteuclid.org/euclid.aoas/1231424214

Greene, N. R., Chism, S., & Naveh-Benjamin, M. (2020). *Levels of specificity in episodic memory: Insights from response accuracy and subjective confidence ratings in older adults and in younger adults under full or divided attention.* PsyArXiv Preprints. https://psyarxiv.com/k9z6m

Greene, N. R., Martin, B. A., & Naveh-Benjamin, M. (in press). The effects of divided attention at encoding and at retrieval on multidimensional source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Greene, N. R., & Naveh-Benjamin, M. (2020). A specificity principle of memory: Evidence from aging and associative memory. *Psychological Science, 31*(3), 316-331. https://doi.org/10.1177/0956797620901760

Gruppuso, V., Lindsay, D. S., & Masson, M. E. J. (2007). I'd know that face anywhere! *Psychonomic Bulletin & Review, 14*, 1085–1089. https://doi.org/10.3758/BF03193095

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

Jones, G. V. (1976). A fragmentation hypothesis of memory: Cued recall of pictures and of sequential position. *Journal of Experimental Psychology: General, 105*(3), 277–293. https://doi.org/10.1037/0096-3445.105.3.277

Kilb, A., & Naveh-Benjamin, M. (2007). Paying attention to binding: Further studies assessing the role of reduced attentional resources in the associative deficit of older adults. *Memory & Cognition, 35*(5), 1162–1174. https://doi.org/10.3758/BF03193486

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75*, 70–98. https://doi.org/10.1007/s11336-009-9141-0

Klauer, K. C., & Kellen, D. (2011a). Assessing the belief bias effect with ROCs: Reply to Dube, Rotello, and Heit (2010). *Psychological Review, 118*(1), 164–173. https://doi.org/10.1037/a0020698

Klauer, K. C., & Kellen, D. (2011b). The flexibility of models of recognition memory: An analysis by the minimum-description length

principle. *Journal of Mathematical Psychology, 55*(6), 430–450. https://doi.org/10.1016/j.jmp.2011.09.002

Knott, L. M., & Dewhurst, S. A. (2007). The effects of divided attention at study and test on false recognition: A comparison of DRM and categorized lists. *Memory & Cognition, 35*(8), 1954–1965. https://doi.org/10.3758/BF03192928

Knott, L. M., Howe, M. L., Toffalini, E., Shah, D., & Humphreys, L. (2018). The role of attention in immediate emotional false memory enhancement. *Emotion, 18*(8), 1063–1077. https://doi.org/10.1037/emo0000407

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21*(11), 1551–1556. https://doi.org/10.1177/0956797610385359

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science, 6*, 299–312. https://doi.org/10.1177/1745691611406925

Kruschke, J. K. (2018). Rejecting or accepting parameter values in Bayesian estimation. *Advances in Methods and Practices in Psychological Science, 1*(2), 270–280. https://doi.org/10.1177/2515245918771304

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis, 100*(9), 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

Luo, L., & Craik, F. I. M. (2009). Age differences in recollection: Specificity effects at retrieval. *Journal of Memory and Language, 60*(4), 421–436. https://doi.org/10.1016/j.jml.2009.01.005

Meng, X.-L. (1994). Posterior predictive *p*-values. *The Annals of Statistics, 22*(4), 1142–1160.

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs (R package version 0.9.11-1) [Computer software]. http://CRAN.R-project.org/package=BayesFactor

Murdock, B. B., Jr. (1965). Effects of a subsidiary task on short-term memory. *British Journal of Psychology, 56*, 413–419. https://doi.org/10.1111/j.2044-8295.1965.tb00983.x

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), 1170–1887. https://doi.org/10.1037/0278-7393.26.5.1170

Naveh-Benjamin, M., & Brubaker, M. S. (2019). Are the effects of divided attention on memory encoding processes due to the disruption of deep-level elaborative processes? Evidence from cued- and free-recall tasks. *Journal of Memory and Language, 106*, 108–117. https://doi.org/10.1016/j.jml.2019.02.007

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1091–1104. https://doi.org/10.1037/0278-7393.24.5.1091

Naveh-Benjamin, M., Guez, J., Hara, Y., Brubaker, M. S., & Lowenschuss-Erlich, I. (2014). The effects of divided attention on encoding processes under incidental and intentional learning instructions: Underlying mechanisms? *The Quarterly Journal of Experimental Psychology, 67*(9), 1682–1696. https://doi.org/10.1080/17470218.2013.867517

Naveh-Benjamin, M., Guez, J., & Marom, M. (2003). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition, 31*, 1021–1035. https://doi.org/10.3758/BF03196123

Naveh-Benjamin, M., Guez, J., & Sorek, S. (2007). The effects of divided attention on encoding processes in memory: Mapping the locus of interference. *Canadian Journal of Experimental Psychology, 61*(1), 1–12. https://doi.org/10.1037/cjep2007001

Nieznański, M. (2013). Effects of resource demanding processing on context memory for context-related versus context-unrelated items. *Journal of Cognitive Psychology, 25*(6), 745–758. https://doi.org/10.1080/20445911.2013.819002

Nieznański, M., & Obidziński, M. (2019). Verbatim and gist memory and individual differences in inhibition, sustained attention, and working memory capacity. *Journal of Cognitive Psychology, 31*, 16–33. https://doi.org/10.1080/20445911.2019.1567517

Odegard, T. N., & Lampinen, J. M. (2005). Recollection rejection: Gist cuing of verbatim memory. *Memory & Cognition, 33*(8), 1422–1430. https://doi.org/10.3758/BF03193375

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

Pazzaglia, A. M., Dube, C., & Rotello, C. M. (2013). A critical comparison of discrete-state and continuous models of recognition memory: Implications for recognition and beyond. *Psychological Bulletin, 139*(6), 1173–1203. https://doi.org/10.1037/a0033044

Pérez-Mata, M. N., Read, J. D., & Diges, M. (2002). Effects of divided attention and word concreteness on correct recall and false memory reports. *Memory, 10*(3), 161–177. https://doi.org/10.1080/09658210143000308

Province, J. M., & Rouder, J. N. (2012). Evidence for discrete-state processing in recognition memory. *Proceedings of the National Academy of Sciences of the United States of America, 109*(36), 14357–14362. https://doi.org/10.1073/pnas.1103880109

R Core Team. (2020). R: A language and environment for statistical computing (Version 3.6.3) [Computer software]. Retrieved from http://www.R-project.org/

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1–75. https://doi.org/10.1016/1041-6080(95)90031-4

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Rotello, C. M. (2017). Signal detection theories of recognition memory. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (2nd ed.). Academic Press.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime 2.0 reference guide manual.* Psychology Software Tools.

Smith, J. B., & Batchelder, W. H. (2010). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology, 54*, 167–183. https://doi.org/10.1016/j.jmp.2009.06.007

Stahl, C., & Klauer, K. C. (2008). A simplified conjoint recognition paradigm for the measurement of gist and verbatim memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(3), 570–586. https://doi.org/10.1037/0278-7393.34.3.570

Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia, 51*(12), 2442–2449. https://doi.org/10.1016/j.neuropsychologia.2012.12.014

Stoet, G. (2010). PsyToolkit—A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Tulving, E. (1983). *Elements of episodic memory.* Clarendon Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*(1), 1–12. https://doi.org/10.1037/h0080017

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging, 13*(2), 230–241. https://doi.org/10.1037/0882-7974.13.2.230

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76*(6), 559–573. https://doi.org/10.1037/h0028143

Wallace, W. P., Malone, C. P., & Spoo, A. D. (2000). Implicit word activation during prerecognition processing: False recognition and remember/know judgments. *Psychonomic Bulletin & Review, 7*, 149–157. https://doi.org/10.3758/BF03210734

Wallace, W. P., Stewart, M. T., Shaffer, T. R., & Wilson, J. A. (1998). Are false recognitions influenced by prerecognition processing? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24*(2), 299–315. https://doi.org/10.1037/0278-7393.24.2.299

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition, 25*, 747–763. https://doi.org/10.3758/BF03211318

Zimmer, H. D., Mecklinger, A., & Lindenberger, U. (2006). Levels of binding: Types, mechanisms, and functions of binding in remembering. In H. D. Zimmer, A. Meclinger, & U. Lindenberger (Eds.), *Handbook of binding and memory: Perspectives from cognitive neuroscience* (pp. 3–22). Oxford University Press.

Open Practices Statement

The data and analysis scripts for this study are available on the Open Science Framework (https://osf.io/vkcgh/).