Memory & Cognition (2018) 46:1058–1075
https://doi.org/10.3758/s13421-018-0821-9

Multinomial models reveal deficits of two distinct controlled retrieval processes in aging and very mild Alzheimer disease

Peter R. Millar[1] · David A. Balota[1,2] · Anthony J. Bishara[3] · Larry L. Jacoby[1]

Published online: 23 May 2018

Abstract
Dual-process models of episodic retrieval reveal consistent deficits of controlled recollection in aging and Alzheimer disease (AD). In contrast, automatic familiarity is relatively spared. We extend standard dual-process models by showing the importance of a third *capture* process. Capture produces a failure to attempt recollection, which might reflect a distinct error from an inability to recollect when attempted (Jacoby et al. *Journal of Experimental Psychology: General*, *134*(2), 131–148, 2005a). We used multinomial process tree (MPT) modeling to estimate controlled recollection and capture processes, as well as automatic retrieval processes, in a large group of middle-aged to older adults who were cognitively normal (N = 519) or diagnosed with the earliest detectable stage of AD (N = 107). Participants incidentally encoded word pairs (e.g., knee bone). At retrieval, participants completed cued word fragments (e.g., knee b_n_) with primes that were congruent (e.g., bone), incongruent (e.g., bend), or neutral (i.e., &&&) to the target (e.g., bone). MPT models estimated retrieval processes both at the group and the individual levels. A capture parameter was necessary to fit MPT models to the observed data, suggesting that dual-process models of this task can be contaminated by a capture process. In both group- and individual-level analyses, aging and very mild AD were associated with increased susceptibility to capture, decreased recollection, and no differences in automatic influences. These results suggest that it is important to consider two distinct modes of attentional control when modeling retrieval processes. Both forms of control (recollection and avoiding capture) are particularly sensitive to cognitive decline in aging and early-stage AD.

Keywords Memory models · Memory · Attention · Recollection · Aging

Although Alzheimer disease (AD) is commonly associated with general deficits in episodic memory (for reviews, see Carlesimo & Oscar-Berman, 1992; R. G. Morris & Kopelman, 1986), there is evidence to suggest that these memory deficits might be specific to attentionally-controlled retrieval processes. This hypothesis is motivated by consistent observations that executive function and attentional control processes also decline in early stages of the disease (for reviews, see Balota & Duchek, 2015; Perry & Hodges, 1999). For example, very mild AD samples demonstrate difficulties in exerting control over incongruent word responses in Stroop color naming (Spieler, Balota, & Faust, 1996), inappropriate meanings of ambiguous words in sentence judgments (Faust, Balota, Duchek, Gernsbacher, & Smith, 1997), false retrieval of non-presented, but semantically-converging, lures (Balota et al., 1999), and highly associated foils in semantic categorization (Aschenbrenner et al., 2015). Together, these findings suggest that AD is associated with an attentional control deficit (Balota & Duchek, 2015; Faust & Balota, 2007), much like the one proposed in cognitively normal aging (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999; West, 1996). It is possible that these AD-related attentional deficits might mediate declines in episodic memory. Similar mediational effects have been observed in cognitively normal aging. Specifically, structural equation modeling has shown that age-related variability in episodic memory can be totally accounted for by differences in a latent variable reflecting executive/attentional control tasks (McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). Hence, memory

✉ Peter R. Millar
pmillar@wustl.edu

[1] Department of Psychological and Brain Sciences, Washington University in St. Louis, One Brookings Drive, Campus Box 1125, St. Louis, MO 63130, USA

[2] Department of Neurology, Washington University in St. Louis, St. Louis, MO, USA

[3] Department of Psychology, College of Charleston, Charleston, SC, USA

processes that are particularly dependent upon attentional control should be most sensitive to early-stage AD.

The process dissociation procedure (PDP) has been shown to be a useful method to distinguish between attentionally-controlled retrieval processes, i.e., *recollection*, and more automatic influences on memory (Jacoby, 1991; for review, see Yonelinas & Jacoby, 2012). In general, results from PDP studies provide support for dual-process models of episodic retrieval. Moreover, the PDP has been particularly informative in characterizing memory deficits that occur in healthy aging and Alzheimer disease (AD; for meta-analysis, see Koen & Yonelinas, 2014; for review, see Yonelinas & Jacoby, 2012). We consider these special populations to inform an expanded process modeling approach that distinguishes between multiple modes of attentional control during retrieval.

Millar et al. (2017) recently applied the PDP to a large sample of 519 cognitively normal middle-aged to older adults and 64 individuals with very early-stage AD. Using a task similar to one developed by Jacoby (1999a), participants incidentally encoded word pairs (e.g., knee-bone). Memory for these pairs was tested using a primed, cued fragment completion task (e.g., knee-b_n_). Critically, for the utilization of the PDP, before each retrieval trial, participants viewed a prime that was either congruent with the correct response (e.g., bone), incongruent (e.g., bend), or neutral (i.e., &&&), see Fig. 1. We refer to these conditions as congruent and incongruent because the prime dimension supports a response that is either consistent or inconsistent with the goals of the task, i.e., retrieval of the initial word pair. Hence, this manipulation functions in a similar manner to that of the word dimension in a Stroop (1935) color-naming task, in which the goal-irrelevant word is either congruent or incongruent with the goal-relevant color dimension. Incongruent primes in the misleading-prime paradigm, displayed in Fig. 1, produce opposition by rendering a plausible competing response that must be overcome by recollection to avoid producing an error. In contrast, congruent primes produce facilitation by rendering the same response produced by successful recollection, resulting in a correct response just as would recollection. As predicted by the attentional deficit account, process-specific deficits in recollection, but not automatic influences, were observed in healthy aging as well as preclinical (biomarker-positive but non-demented) and very mildly demented AD individuals (Millar et al., 2017). Estimates of recollection significantly improved discrimination between cognitively normal individuals and those with early-stage AD above and beyond a large battery of psychometric tests, including measures of memory, attention, and processing speed. These findings are consistent with results from other methods of retrieval process estimation across a variety of tasks. Briefly, age- and AD-related deficits in controlled recollection are consistently observed, while deficits in automatic familiarity are smaller and less consistent (for meta-analysis, see Koen & Yonelinas, 2014).

Millar et al. (2017) were primarily interested in the sensitivity of recollection estimates to aging and AD status from a dual-process perspective, and so they did not consider other, more complex process models that might underlie performance. It is possible that the dual-process model may be an over-simplification of retrieval processes involved in the task. In support of this possibility, Jacoby, Bishara, Hessels, and Toth (2005a) modeled data from a similar misleading-prime task with multinomial process tree (MPT) models, as shown in Fig. 2. In dual-process models, recollection and accessibility bias serve as alternative bases for cued retrieval (Fig. 2a and b). When unable to recollect a past event, participants guess with a highly accessible, plausible response. However, in order to fit data from the misleading-prime task, Jacoby et al. (2005a) had to model an additional *capture* parameter (Fig. 2c), reflecting a controlled mode of retrieval constraint. The functional consequence of capture is that a participant accepts a prime word as a correct response, without attempting to recollect. As shown in Fig. 2c, when capture is successful, the participant responds with the prime word, resulting in a correct response for congruent primes or an intrusion error for incongruent primes. In this model, the probability of capture limits the possibility of attempting to recollect, whereas the probability of recollection measures the success of recollection when attempted, i.e., when capture has failed. Thus, although the capture process might not contribute directly to the recovery of memory *per se*, this model treats capture as a mode of cognitive control that should determine whether one *attempts* recollective retrieval.

Importantly, the capture model expands upon the traditional dual-process model in that it distinguishes between two possible accounts of intrusion errors on the misleading-prime task. Specifically, on incongruent trials, participants might commit either (1) a capture error, in which the participant accepts the incongruent prime as correct, pre-empting an attempt to recollect the initial word pair, or (2) a recollection error, in which the participant is not captured by the prime, allowing for a recollection attempt, but that attempt is unsuccessful, resulting in reliance on accessibility, as in a standard dual-process model. In this way, recollection is *conditional* on the failure of capture. The distinction between capture and recollection in this model is analogous to the theoretical distinction drawn by Flavell (1970) between production and mediation in characterizing the development of verbally-mediated mnemonic strategies in children. According to Flavell, failure to engage a strategy might result from two distinct pathways: (1) a production deficiency, in which the appropriate word or phrase cannot be generated, thus negating any chance to correctly apply a mediation strategy, or (2) a mediational deficiency, in which the appropriate word or

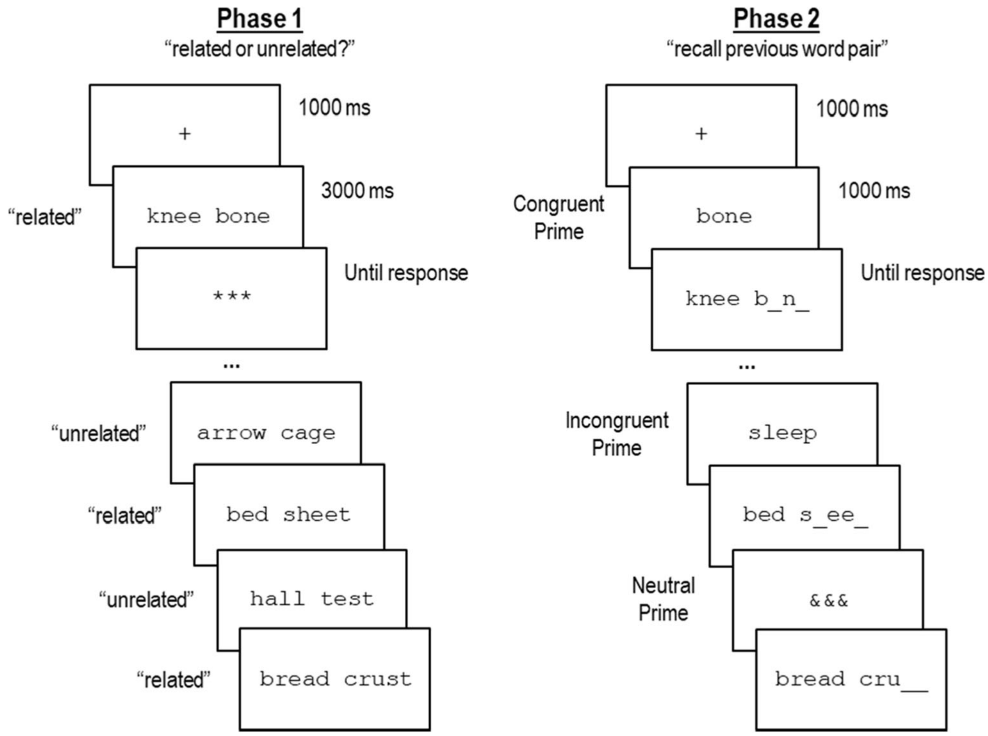

Fig. 1 Schematic representation of behavioral task with example trials. Adapted from Millar et al. (2017)

phrase is successfully generated, but the attempt to apply a mediation strategy is unsuccessful. Hence, mediation is conditional upon the success of production. As successful production affords the opportunity for mediation in Flavell's (1970) model, so does a failure of capture afford an attempt at recollection in the current MPT model (Fig. 2). Although the two models describe very disparate basic phenomena, both models assume a conditionally-nested relationship between distinct cognitive processes.

In terms of a theoretical mechanism, capture may be driven by a failure to inhibit misleading information, i.e., the prime word (cf. Hasher & Zacks, 1988). Alternatively, the process may be a form of goal neglect, i.e., a failure to maintain the task goal of recalling the previous word pairs (e.g., Duncan,

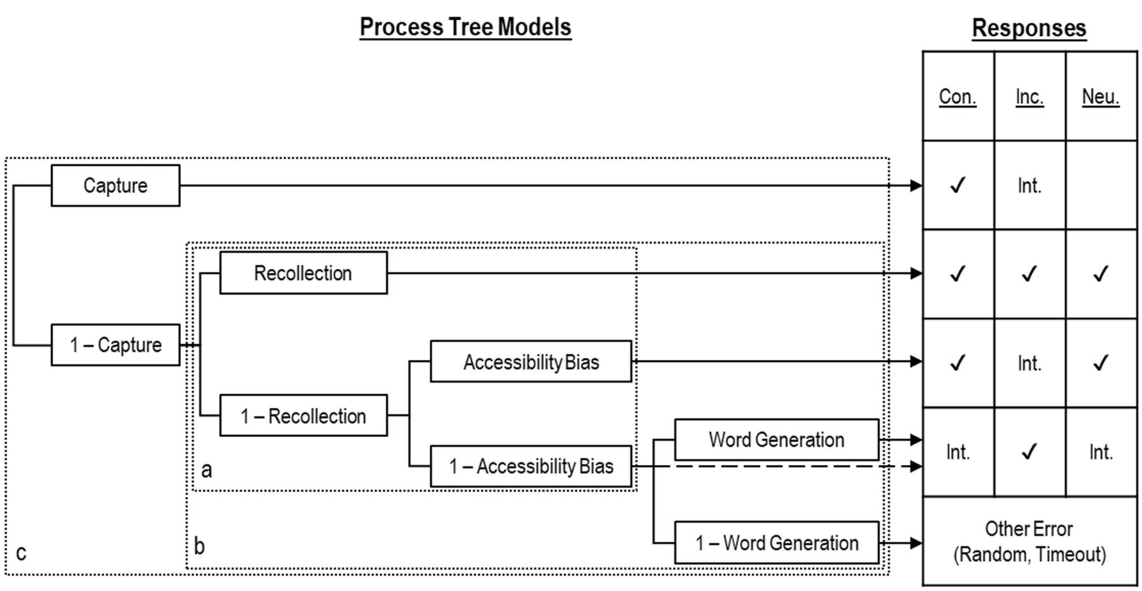

Fig. 2 Multinomial process tree models of episodic retrieval. Branches lead to a correct response (✓), an intrusion error (Int.), or other error, including a random word response or timeout, for the congruent (Con.), incongruent (Inc.), or neutral (Neu.) trials. Dotted boxes indicate the standard dual-process model with recollection and accessibility bias (**a**), the addition of word generation (**b**), and the addition of capture (**c**)

Emslie, Williams, Johnson, & Freer, 1996). Either of these possibilities implicates a failure of cognitive control process.

The role of capture is also informed by models of memory designed to account for patterns of false memory or "confabulation" that occur in neurological patients (Kopelman, 1999). In some cases, confabulating patients will "accept as veridical whatever the ecphoric process delivers to consciousness" (Moscovitch, 1989, p. 155), sometimes resulting in the intrusion of input stimuli into memory reports that are "sensible but untrue" (Burgess & Shallice, 1996, p. 360). These confabulations and the capture-based intrusions in the current paradigm might reflect similar types of memory errors. Importantly, attempts to explain confabulation implicate a relatively early process, in which the conditions for successful retrieval are specified. Terms such as "retrieval specification" (D. A. Norman & Bobrow, 1979), "descriptor" processes (Burgess & Shallice, 1996), or the "focusing" process (K. A. Norman & Schacter, 1996; Schacter, Norman, & Koutstaal, 1998) have been used to describe this stage. These processes are said to limit the elaboration of retrieval cues to specify information associated with the initial encoding context, thus constraining out the influence of possible interfering information. Hence, in the current model, capture might reflect a failure of these specification processes. These processes are also distinct from later "evaluation" processes, which compare output from memory storage against the specified retrieval cue.

The dissociation between specification (pre-access) and evaluation (post-access) processes is supported by neuropsychological findings. For example, Dab, Claes, Morais, and Shallice (1999) studied a confabulating frontal lesion patient, PAD. The authors concluded that PAD exhibited intact evaluation processes, as evidenced by recognition memory performance in the normal range. This finding contradicted the predictions of simpler models, under which a retrieval evaluation deficit was necessary for confabulation (Hanley, Davies, Downes, & Mayes, 1994). Instead, Dab et al. (1999) argued that PAD's confabulation was driven by a deficit in specification, as evidenced by free recall deficits, particularly a profile of high semantic intrusions in a multiple-list learning paradigm. Thus, considering an additional retrieval specification process proved useful in interpreting the case. Similarly, considering a capture process, distinct from recollection and automatic processes, might be useful in interpreting memory deficits observed in aging and AD.

At a more general level, it is quite possible that a capture process might sometimes contaminate recollection estimates derived from simpler two-process models. Although Millar et al. (2017) demonstrated that PDP estimates of recollection were robust in capturing age- and AD-related memory differences, the MPT modeling by Jacoby et al. (2005a) suggests that task performance might also be driven by a distinct capture process. Hence, a more complex retrieval model (including two distinct modes of cognitive control) might better reflect the controlled aspects of memory retrieval by providing not only additional estimates of previously-unmodeled capture processes, but also less contaminated estimates of recollection.

Jacoby et al. (2005a) demonstrated that the extended capture model might be useful for describing age differences in controlled retrieval processes. In two variations of the misleading-prime task that used a cued fragment completion task, a capture model was necessary to fit age-related task performance. In a third experiment, which used a recognition test instead of cued fragment completion, performance was also fit by the capture model, as well as a simpler dual-process model, although these model fits were not directly compared. Importantly, older adults reliably exhibited higher capture parameters than younger adults, consistent with an attentional deficit hypothesis (e.g., Hasher & Zacks, 1988). However, we are currently unaware of any other studies that have used this modeling approach to replicate or extend the previous findings. Hence, in the present study, we provide several novel extensions beyond the previous modeling efforts. First, we eliminate more complex aspects of the Jacoby et al. (2005a) paradigm, including subjective reports of memory or manipulations to withhold a response at retrieval, to test the model in a shorter, simpler task. These previous manipulations had been used to model an additional attribution threshold parameter, which is not of interest to the present study. Second, we extend beyond group-level MPT modeling, which assumes homogeneity in process estimates across individuals, by modeling retrieval processes at the individual participant level. Third, we test whether these process estimates are sensitive to more subtle age differences by comparing community-dwelling middle-aged and older adults, rather than a more extreme groups design, i.e., comparing young college students to older community-dwelling participants as done by Jacoby et al. (2005a). Finally, we extend this model to early-stage AD, which, as mentioned, is another population that has shown deficits in attentional control (Balota & Duchek, 2015). Together these extensions allow for a more generalizable assessment of the utility of the capture model, as compared to traditional dual-process models, using data from Millar et al. (2017).

To anticipate, we find that a capture parameter is indeed necessary for an adequate fit to the misleading-prime task data, suggesting that a capture process might contaminate estimates derived from a dual-process model of the task. Importantly, we also test for age- and AD-related process differences in this more complex model. Recollection deficits are consistently observed in early-stage AD (Koen & Yonelinas, 2014), but have not been tested in a model that also accounts for capture. Like recollection, capture is also thought to be an attention-dependent process, but has not been examined in the context of AD. Under the hypothesis that aging and AD are associated with declines of attentional control (Balota &

Duchek, 2015; Faust & Balota, 2007; Perry & Hodges, 1999), we predict process-specific group differences in both recollection and capture. In contrast, since other automatic processes in the model (by definition) should not be demanding of attention, we expect these processes to be relatively unaffected across these groups. We test this hypothesis by examining MPT model parameters derived from performance on the misleading-prime task at both the group and the individual participant levels.

Method

Participants

Participants included 617 individuals, recruited from two longitudinal studies administered by the Charles F. and Joanne Knight Alzheimer Disease Research Center at Washington University in St. Louis. An initial report of these data is provided by Millar et al. (2017). Cognitive status of the participants was evaluated by a trained clinician via the Clinical Dementia Rating scale (CDR; J. C. Morris, 1993), with CDRs of 0, .5, 1, 2, or 3, indicating, respectively, cognitive normality, very mild, mild, moderate, or severe dementia. The final sample consisted of 510 cognitively normal individuals (CDR 0), and 107 individuals with very mild AD (CDR 0.5). Clinical diagnosis of AD in these CDR 0.5 individuals was based on NINCDS-ADRDA criteria (McKhann et al., 1984). This CDR 0.5 sample included individuals with a clinical diagnosis of symptomatic AD (n = 64) or uncertain dementia (n = 43), excluding non-AD or mixed etiologies, such as frontotemporal dementia, Lewy body dementia, or contributing depression. Demographic characteristics of the sample are reported in Table 1. The Washington University Human Research Protection office approved all research methods. All participants gave written informed consent.

Materials

Word stimuli were selected according to previously-described norms (Jacoby, 1996, 1999b; Millar et al., 2017). To summarize, related word triads were constructed to include two potential target words that were each semantically related to a common target. Further, the two target words were constrained to have the same length and at least two identical letters in the same position (e.g., knee bone, knee bend). For each triad, a word fragment was constructed that might be completed by either of the target words (e.g., b_n_). A total of 34 such triads were produced and were assigned to serve as congruent, incongruent, neutral, or buffer trials. The frequency of target words was equated across the congruent, incongruent, and neutral conditions, based on log-transformed Subtitle Frequency (Log SUBTL-WF; Brysbaert & New, 2009), $F(2,27) < 1.00$, $p = .395$.

Procedure

Participants completed the short, 10-min task in two phases (see Fig. 1). Phase 1 was an incidental encoding task, in which participants made a judgment of relatedness for each of 30 related word pairs (e.g., knee bone) and ten unrelated pairs (e.g., arrow cage). Immediately afterward, participants completed Phase 2, in which the same 30 related word pairs served as targets for a primed, cued fragment completion task with explicit retrieval instructions. For each trial, participants were instructed to retrieve a word pair from Phase 1 in order to complete a cued word fragment (e.g., knee b_n_). Before the onset of each cued fragment, a prime stimulus appeared for 1,000 ms. Ten primes were congruent with the correct response (e.g., bone); ten were incongruent with the correct response (e.g., bend); and ten were neutral symbols (i.e., &&&). Incongruent primes were all valid completions of the fragment and were semantically related to the cue, but had not been presented in Phase 1. Participants were instructed to silently read each prime before completing the cued fragment. They were also informed that the primes might be congruent or incongruent with the correct answer to be retrieved from Phase 1. The full script used to run this task is available for download at https://osf.io/pwhfd/?view_only=5feaa8c64c1f4660a3b02070f25938ec.

Results

Overview of analyses

In order to simplify group-level MPT models, we tested the effects of aging by dividing the sample of CDR 0 individuals into two groups by a median split at age 70 years. We tested the effects of early-stage AD by comparing the CDR 0.5 sample to a subsample of age-matched CDR 0s (aged 68 years or older). This age-matched CDR 0 control group included all of the individuals from the older CDR 0 group, as well as some individuals (N = 38) from the younger CDR 0 group.

For both age- and AD-related comparisons, we first report differences in task performance for descriptive purposes, which is a brief recapitulation of results reported in Millar et al. (2017). We then report our primary analyses of interest, testing for differences in retrieval process estimates derived from MPT models at both the group and the individual levels.

Table 1 Demographic measures, mean (SD), grouped by age and CDR status

Variable (units)	CDR 0 Age <70	CDR 0 Age 70+	*Age F(df)*	*Age p*	CDR 0 Controls	CDR 0.5	*AD F(df)*	*AD p*
N	254	265			303	107		
% Female	65%	61%			61%	41%		
Age (years)	61.1 (6.2)	76.4 (5.1)	944.3 (1,517)	<.001	75.4 (5.5)	75.5 (7.4)	0.01 (1,408)	.91
Education (years)	15.9 (2.5)	15.4 (2.7)	4.3 (1,512)	.04	15.4 (2.7)	15.2 (2.7)	0.6 (1,400)	.45
MMSE	29.3 (1.0)	28.7 (1.5)	30.02 (1,492)	<.001	28.7 (1.5)	26.9 (2.9)	68.6 (1,394)	<.001

Note: F(*df*)s report the univariate *F* statistics (and degrees of freedom) for the effects of Age (CDR 0 Age < 70 vs. Age 70+) and AD (CDR 0.5 vs. Age-matched CDR 0 Controls)

Memory task performance as a function of aging

Figure 3 displays the mean proportion of each response type (correct, intrusion error, or other error) as a function of age (3a) and CDR status (3b). We tested the effects of aging on memory task performance in a 2 × 3 mixed-model analysis of variance (ANOVA), with the proportion of correct responses as the dependent variable, age group as a between-subjects factor (< 70 years or 70+ years), and condition as a within-subjects factor (congruent, incongruent, or neutral). Only CDR 0s were included in this analysis. As expected, this analysis revealed a main effect of age, $F(1,517) = 32.56$, $p < .001$, $\eta_p^2 = .06$, a main effect of condition, $F(2,1034) = 505.77$, $p < .001$, $\eta_p^2 = .50$, and an interaction between age and condition, $F(2,1034) = 13.90$, $p < .001$, $\eta_p^2 = .03$. As shown in Fig. 3a, older age was associated with a greater decrease in correct responses for incongruent trials compared with neutral or congruent trials.

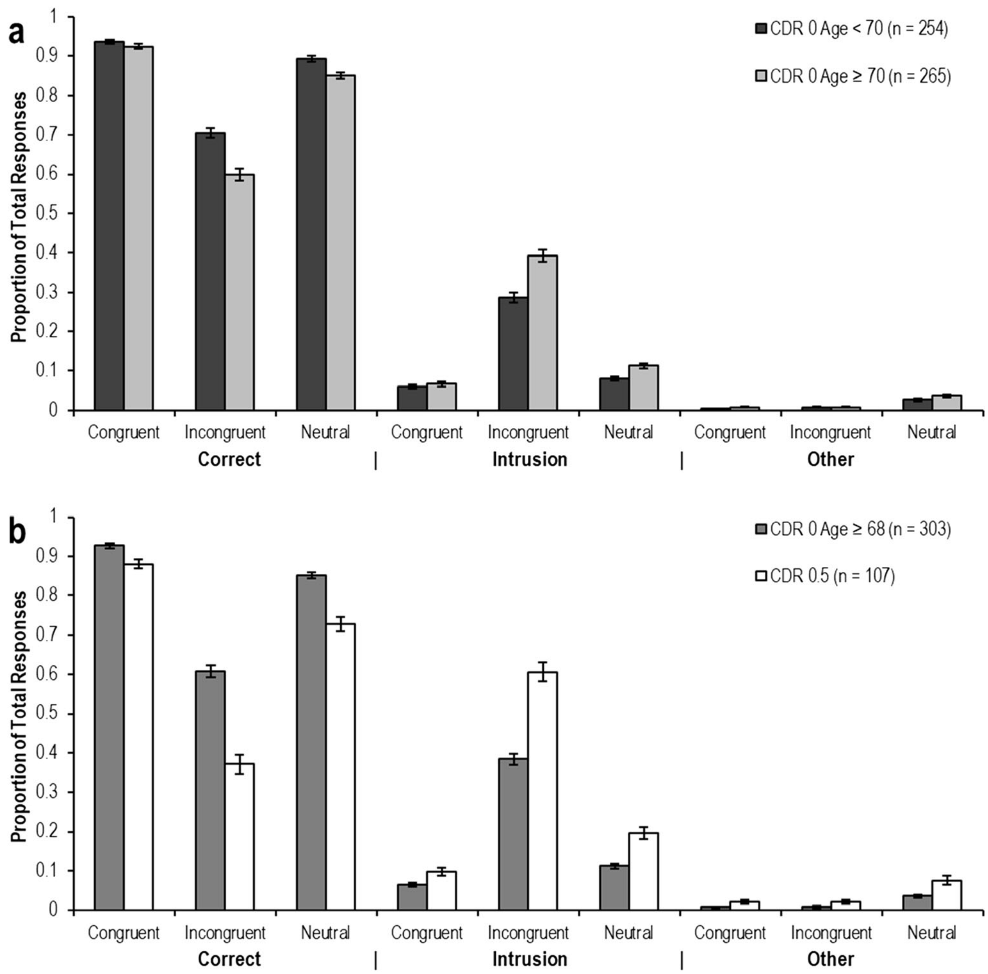

Fig. 3 Mean proportion (+/- standard error of the mean) of correct responses, intrusions, and other errors as a function of trial type (congruent, incongruent, or neutral), age (**a**), and clinical diagnosis (**b**)

The age effects on correct memory responses might be driven by differences in the proportion of intrusions of the critical/incongruent prime word or other errors. Notably, intrusion errors occurred in all three prime conditions, including congruent trials, for both younger and older adults.[1] Analysis of the intrusion errors revealed a main effect of age, $F(1,517) = 31.63$, $p < .001$, $\eta_p^2 = .06$, a main effect of condition, $F(2,1034) = 549.91$, $p < .001$, $\eta_p^2 = .52$, and a reliable interaction between age and condition, $F(2,1034) = 16.04$, $p < .001$, $\eta_p^2 = .03$. As shown in Fig. 3a, older age was associated with a greater increase in intrusion responses for incongruent trials compared with neutral or congruent trials.

Other errors included trials in which the participant responded with a word that was neither the correct response nor the critical lure, as well as trials in which the participant did not respond within 20 s. For these errors, there was a marginal main effect of age, $F(1,517) = 3.76$, $p = .053$, $\eta_p^2 = .007$, and a main effect of condition, $F(2,1034) = 62.41$, $p < .001$, $\eta_p^2 = .11$, but the interaction between age and condition was not significant, $F(2,1034) = 2.11$, $p = .12$, $\eta_p^2 < .01$.

In summary, as reported in Millar et al. (2017), the results indicate that age-related differences in task performance are strongly driven by intrusions of the primes in the incongruent condition, and to a much smaller extent, by other errors.

Memory task performance as a function of CDR status

A 2 (CDR: 0 or 0.5) × 3 (condition: congruent, incongruent, or neutral) mixed-model ANOVA was conducted on mean correct performance. This analysis revealed a main effect of CDR status, $F(1,408) = 106.34$, $p < .001$, $\eta_p^2 = .21$, and an interaction between CDR status and condition, $F(2,816) = 27.52$, $p < .001$, $\eta_p^2 = .06$. As shown in Fig. 3b, the group difference in performance between CDR 0s and 0.5s was greater for incongruent trials compared with neutral trials or congruent trials.

The effects of CDR status on intrusion and other errors were tested in separate ANOVA models, using the factor structure described above for correct responses. For intrusion errors, there was a significant main effect of CDR status, $F(1,408) = 87.55$, $p < .001$, $\eta_p^2 = .18$, and an interaction between CDR status and condition, $F(2,816) = 30.69$, $p < .001$, $\eta_p^2 = .07$. As shown in Fig. 3b, the group difference in intrusion responses between CDR 0s and 0.5s was greater for incongruent trials compared with neutral trials or congruent trials.

For other errors, there was a significant main effect of CDR status, $F(1,408) = 28.76$, $p < .001$, $\eta_p^2 = .07$, and an interaction between CDR status and condition, $F(2,816) = 6.85$, $p = .001$, $\eta_p^2 = .02$. As shown in Fig. 3b, the group difference in other errors between CDR 0s and 0.5s was greater for neutral trials compared with congruent trials or incongruent trials.

In summary, the present results indicate that, in an age-matched comparison, AD-related differences in task performance are strongly driven by intrusions of the primes in the incongruent condition, and to a much smaller extent, by other errors.

MPT capture model

The full 4-parameter capture model is shown in Fig. 2. The first parameter, capture (C), represents the probability that a participant would respond with the cue word, overriding a recollection attempt. Successful capture would produce a correct response for congruent trials, and an intrusion for incongruent trials, and would have no effect on neutral trials, since the neutral prime (i.e., &&&) is not a valid response. Given a failure of capture, a participant may attempt to recollect (R), i.e., actively reconstruct the previous encoding phase in search of the correct response. Successful recollection would produce a correct response on all trials. Given a failure of recollection, a participant might be subject to accessibility bias (A), i.e., a tendency to produce the response that comes to mind most fluently. Successful accessibility bias would produce a correct response for congruent and neutral trials and an intrusion for incongruent trials. Given a failure of accessibility bias, a participant may generate a word (W), i.e., produce an unprimed response that is consistent with the constraints of the task, including the correct number and placement of letters and semantic relation to the cue. That response would be an intrusion error for congruent or neutral trials, but would be correct for incongruent trials. Given a failure of word generation, a participant may finally produce an "other" error, including invalid completions of the fragment or failures to respond.

All MPT models were constructed and tested using the R package, "MPTinR" (Singmann & Kellen, 2013). We began

[1] All participant groups committed intrusion errors in all three conditions, including congruent and neutral trials, in which the intrusion response was never presented. For example, on some congruent trials, participants might have studied the word pair, "knee bone," then at retrieval, received the congruent prime, "bone," for the trial, "knee b_n_." On average, participants made intrusion errors on about 5–10% of such congruent trials, responding with "bend" even though it was never presented as a target or a prime. Our MPT model (Fig. 2) accounts for these errors as cases where capture, recollection, and accessibility processes fail and word generation is successful. Thus, participants produce an unprimed response that is consistent with the constraints of the trial.

It is possible that in situations of an intrusion error, participants may falsely believe that primes on congruent trials are actually incongruent, leading them to construct an alternative response to fit the fragment and avoid using the prime word. This interpretation is possible, since participants are indeed instructed to expect some incongruent trials. However, we cannot account for participants' beliefs about the prime type or the intent of their responses in the current task or models, and so we will not attempt to interpret these errors any further.

by comparing the fits of three competing models – the RA and RAW dual-process models, as well as the CRAW capture model (see Fig. 2a, b, and c) – within groups based on age and CDR. After identifying the best-fitting model, we then proceeded to test group differences within each of the model's parameters.

Comparison of model fits within groups

Due to the very large number of observations, our MPT modeling analyses might be over-powered using a conventional significance criterion of α = .05 or even α = .001. As a consequence, MPT models might identify trivially small parameter differences as significant or reject appropriate models due to very small deviations from observed data. Thus, we performed compromise power analyses to set a critical α for each significance test, using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009). For each group, we calculated the critical G^2 for a minimum effect size of ω = 0.10. Since the critical value was based on effect size, the value of α varied for each test from ~1 × 10^{-6} to ~.008 as a function of the total sample size and degrees of freedom on the model, although we fixed the β/α ratio to 1.0. For example, the critical G^2 for the test of the RA model (*df* = 1), in the younger CDR 0 group (*N* = 254 participants * 30 trials = 7,620 observations) was 19.72, corresponding to $\alpha = \beta = 9 \times 10^{-6}$.

For the younger CDR 0 group, the data were poorly fit by both the RA ($G^2(1)$ = 125.39, critical = 19.72) and RAW models ($G^2(3)$ = 91.33, critical = 23.66). Adding the capture parameter significantly improved fit over the RAW model, $\Delta G^2(1)$ = 64.26, critical = 19.72, but the full CRAW model still significantly deviated from the data, $G^2(2)$ = 27.08, critical = 21.85. Since the RA model was not nested within either of the other models, we also compared fit between models using the Bayesian information criteria (BIC). This comparison suggested that the capture model (BIC = 62.83) provided a better fit than either the RA (BIC = 143.24) or RAW models (BIC = 118.15).

For the older CDR 0 group, the data were also poorly fit by the RA ($G^2(1)$ = 159.30, critical = 20.54) and RAW models ($G^2(3)$ = 102.50, critical = 24.52). Again, the capture parameter significantly improved model fit, $\Delta G^2(1)$ = 78.26, critical = 20.54. Although the CRAW model significantly deviated from the data, $G^2(2)$ = 24.24, critical = 22.70, it still provided a better fit (BIC = 60.16) than either the RA (BIC = 177.23) or RAW models (BIC = 129.44).

The CDR 0.5 group was also poorly fit by the RA ($G^2(1)$ = 76.25, critical = 8.66) and RAW models ($G^2(3)$ = 33.00, critical 11.97). However, the CRAW model provided an adequate fit to the data, $G^2(2)$ = 5.57, critical = 10.45. This model produced a significant improvement in fit over the RAW model, $\Delta G^2(1)$ = 27.43, critical = 8.66, and again, the CRAW model had a lower BIC (13.57) than either the RA (BIC = 80.25) or RAW models (BIC = 39.00).

Finally, the age-matched CDR 0 control group followed a similar pattern. This group was poorly fit by the RA ($G^2(1)$ = 170.46, critical = 23.40) and RAW models ($G^2(3)$ = 107.98, critical 27.48), but adequately fit by the CRAW model ($G^2(2)$ = 25.55, critical = 25.61). The CRAW model produced a significant improvement in model fit, $\Delta G^2(1)$ = 82.43, critical = 23.40, and also had the lowest BIC (RA = 188.65, RAW = 135.32, CRAW = 62.01).

Within each group, the addition of a capture parameter consistently led to a significant improvement in model fit. Although the full capture model still significantly deviated from the data in the case of the younger and older CDR 0 groups, comparison of the BIC consistently identified it as the best-fitting model among those considered. We interpret these observations to suggest that the capture parameter is necessary to model the current dataset.

MPT estimates as a function of age

We tested for age differences in each parameter with nested model comparisons of the older and younger CDR 0 groups. To summarize, we first built a model in which both groups shared a complete set of MPT parameter estimates. We then compared the fit of that model to a nested model with separate, group-specific estimates for a given parameter. Thus, significant group differences in a parameter are inferred when the group-specific parameter model yields a significant improvement over the shared-parameter model (Riefer & Batchelder, 1988). Nested model comparisons of single-parameter effects revealed significant age differences in both capture, $\Delta G^2(1)$ = 58.19, and recollection, $\Delta G^2(1)$ = 67.77, but no differences in accessibility bias $\Delta G^2(1)$ = 4.48, or word generation, $\Delta G^2(1)$ = 1.34, critical = 39.60.

Based on the results of these single-parameter effects, we then considered a model with age differences in capture and recollection, but shared parameters for accessibility bias and word generation, i.e., CCRRAW model. This model, however, marginally deviated from the data, $G^2(6)$ = 52.38, critical = 49.53. Nested model comparisons revealed that the fit could not be significantly improved by adding age differences in accessibility bias and/or word generation, ΔG^2s ≤ 1.07, critical values ≥ 39.60. Moreover, we also compared the fit of the CCRRAW model to all possible non-nested combinations of age differences in all four parameters. Comparison of the BIC, as well as Akaike information criteria (AIC) and Fisher information approximation (FIA) indices,

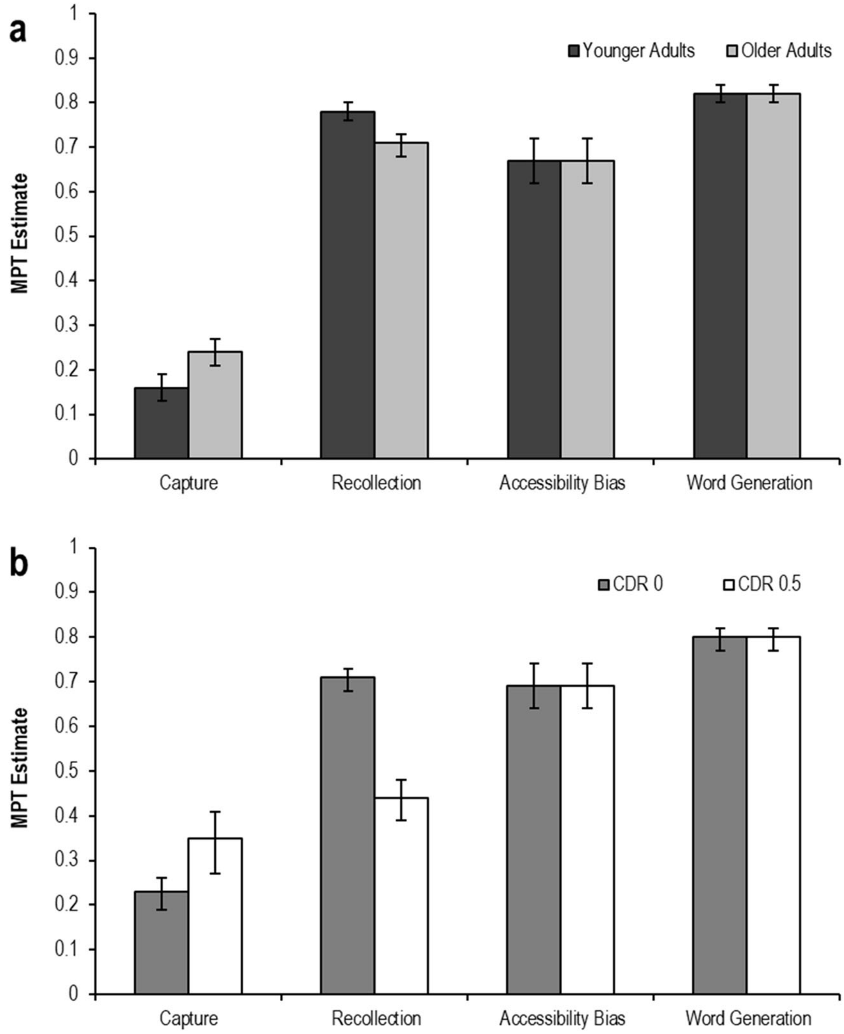

Fig. 4 Parameter estimates (and bootstrapped 95% confidence intervals) of best-fitting MPT models of aging differences (**a**) and very mild AD differences (**b**) in retrieval processes

converged on this model as the best fit (see Appendix Table 2). Thus, despite failing to satisfy the critical G^2 value, we concluded that the CCRRAW model provided the best fit to the age group differences.

In order to provide converging evidence, we then estimated the variability of parameter estimates in this model by performing a parametric bootstrap (Efron & Tibshirani, 1994; Singmann & Kellen, 2013). After fitting the data to the CCRRAW model, we generated 10,000 bootstrap samples based on the parameter estimates of the original model. We then fit each bootstrap sample to the same CCRRAW model and derived the empirical 95% confidence interval (CI) from the distribution of the bootstrapped parameter estimates. As shown in Fig. 4a, older participants were more likely to be captured by the prime (C_o = .24, 95% CI = [.21, .27]) than younger participants (C_y = .16, 95% CI = [.13–.19]). Older participants were less likely to recollect (R_o = .71, 95% CI = [.68–.73]) than younger participants (R_y = .78, 95% CI = [.76–.80]). Since the accessibility bias (A = .67, 95% CI = [.62–.72]) and word generation (W = .82–95% CI = [.80–.84]) parameters were shared, both age groups were equally likely to engage those processes.

MPT estimates as a function of CDR status

Nested model tests of single-parameter effects revealed significant AD group differences in capture, $\Delta G^2(1)$ = 116.74, and recollection, $\Delta G^2(1)$ = 263.17, but no differences in accessibility bias, $G^2(1)$ = 0.11, or word generation, $G^2(1)$ = 16.20, critical = 31.43. Based on these results, we then considered the same CCRRAW model, described in the age analyses above, to model the AD-related group differences in MPT parameters. This model provided an adequate fit to the data, $G^2(6)$ = 38.51, critical = 40.90. Nested model comparisons revealed that the fit could not be significantly improved by adding AD group differences in accessibility bias and/or word generation,

ΔG^2s ≤ 7.39, critical values ≥ 31.43. We also compared the fit to non-nested models including all possible combinations of AD group differences in the four parameters. Comparison of the BIC across models identified this CCRAW model as the best fit (see Appendix Table 3). In contrast to the BIC, the AIC and FIA statistics identified a more complex model, which included group differences in not only capture and recollection, but also word generation, as the best fit. As already mentioned, however, the nested model comparison revealed that adding the group difference in word generation did not significantly improve the fit over the CCRRAW model, $\Delta G^2(1)$ = 6.91, critical = 31.43. Thus, we concluded that the CCRRAW model provided the best fit to the AD group differences.

We then estimated variability of parameter estimates in this model, using the same parametric bootstrap procedures described above. As shown in Fig. 4b, CDR 0.5s were more likely to be captured by the prime ($C_{0.5}$ = .35, 95% CI = [.27–.41]) than CDR 0s (C_0 = .23, 95% CI = [.19–.26]). CDR 0.5s were less likely to recollect ($R_{0.5}$ = .44, 95% CI = [.39–.48]) than CDR 0s (R_0 = .71, 95% CI = [.68–.73]). Since the accessibility bias (A = .69, 95% CI = [.64–.74]) and word generation (W = .80, 95% CI = [.77–.82]) processes were shared, both CDR groups were equally likely to engage those processes.

Individual participant parameter estimation

Modeling process differences at the group level is often a common practice for MPT modeling. However, this method makes a strong assumption that parameter estimates are homogeneous across participants within groups. In order to seek converging evidence for the group differences observed these models, we also estimated parameters for

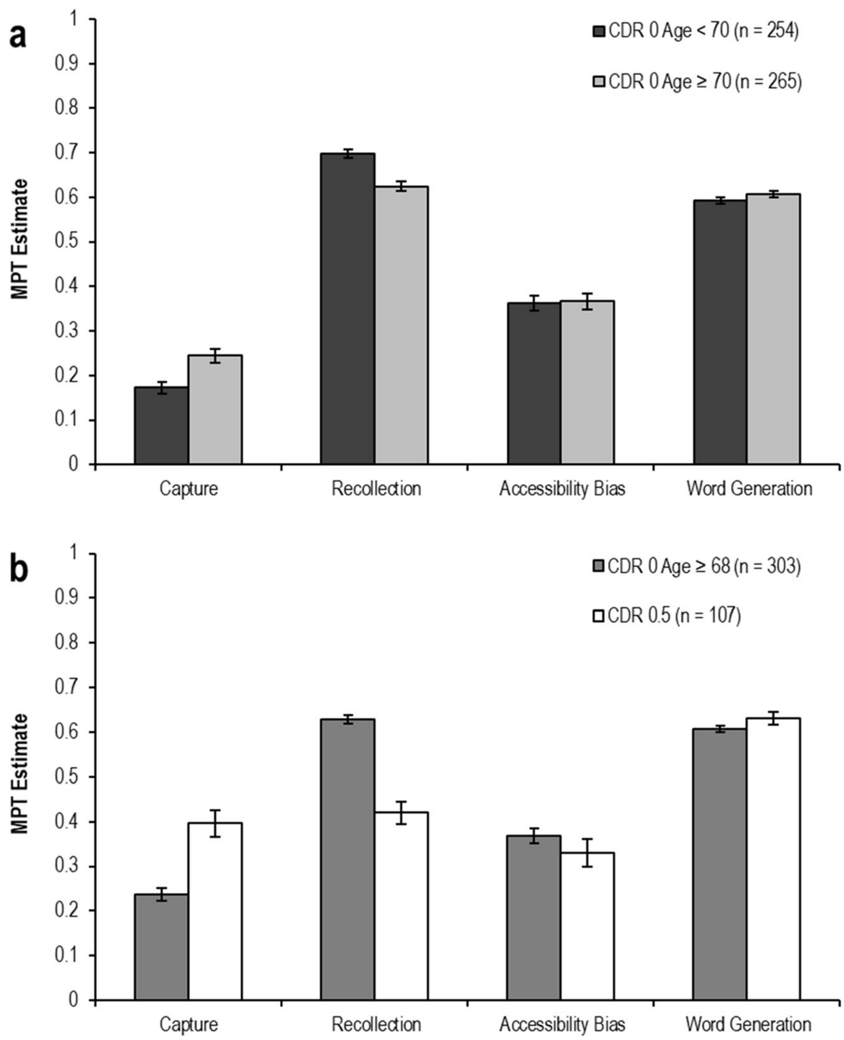

Fig. 5 Mean individual-level process estimates (+/- standard error of the mean) of capture, recollection, accessibility bias, and word generation as a function of age (**a**) and clinical diagnosis (**b**)

each individual participant, based on their conditional proportions of response types. This approach affords an important novel extension over earlier demonstrations of modeling a capture parameter, in which the MPT model was fit to group-level means (i.e., Jacoby et al., 2005a). In contrast to group-level models, individual-level analyses allow for improved model accuracy by allowing for variability in process estimates between individuals within groups. Of course, one limitation of the individual participant method was that many participants produced no responses for several critical bins, preventing the model-fitting algorithm from converging on a set of parameters. To correct for this floor effect, we added $1/(2N)$ to the cells with a value of 0, where N was the total number of trials per condition, i.e., 10. This correction is similar to others commonly performed to estimate d' in signal detection models (Hautus, 1995; Macmillan & Kaplan, 1985).

We then fit each participant's corrected conditional response proportions with the four-parameter capture model using "MPTinR" (Singmann & Kellen, 2013) and recorded the four parameter estimates. Mean parameter estimates are presented in Fig. 5, broken down by the age and AD comparisons tested below.

Distributions of parameter estimates were highly non-normal, including very high degrees of skew. Additionally, the test of AD effects included highly unequal sample sizes (i.e., 303 vs. 107). Independent *t*-tests are often robust to either non-normality or unequal sample sizes alone, but in combination, these tests might suffer from inflated Type I error (Bradley, 1978). Thus, to avoid violating the assumptions of parametric tests and preserve statistical power, we tested age- and AD-related group differences in parameter estimates by performing a series of nonparametric randomization tests.

We tested age differences in individual parameter estimates by randomly reassigning group labels to the full sample of CDR 0 individuals 10,000 times. Each reassignment produced groups of 254 and 265 individuals to match the sample sizes of the original younger and older groups. For each reassignment, we calculated the mean group difference for a given parameter estimate and stored that difference value. After resampling was complete, we compared the original mean difference between age groups to the empirical sampling distribution built over the 10,000 reassignments. A power analysis conducted using G*Power 3.1 (Faul et al., 2009) indicated that, with this sample size, a significance test with a two-tailed α of .05 could detect even relatively small effects (d = 0.30) with adequate statistical power ($1 - \beta = .93$). Age groups significantly differed in the capture parameter. The observed age difference in capture was more extreme than 99.9% of the empirical sampling distribution, corresponding to an empirical two-tailed p value equal to .001, Cohen's $d = 0.30$. Age groups also significantly differed in recollection, empirical $p < .001$, Cohen's d = -0.46, but did not differ in accessibility bias, empirical p = .854, Cohen's d = 0.02, or in word generation, empirical p = .171, Cohen's d = 0.12. Importantly, these participant-level analyses strongly converged with the group-level MPT analyses of age-related differences.

We also tested for AD-related differences in individual parameter estimates by repeating the same randomization test procedure described above, this time randomly reassigning group labels to the sample of CDR 0.5 individuals and age-matched CDR 0s. A similar power analysis indicated that a test in this sample could detect small effects (d = 0.30) with statistical power of .76. AD groups significantly differed in capture, empirical p < .001, Cohen's d = 0.58, and recollection parameters, empirical p < .001, Cohen's d = -1.03, but not in accessibility bias, empirical p = .250, Cohen's d = -0.13, or word generation parameters, empirical p = .114, Cohen's d = 0.18. Again, these results were consistent with the group-level MPT analyses of AD-related differences. Finally, we also provide a brief description of an alternative hierarchical Bayesian MPT modeling approach.[2]

[2] We also attempted to model the data with a latent-trait hierarchical Bayesian MPT approach (Klauer, 2010), using the 'TreeBUGS' package in R (Heck, Arnold, & Arnold, 2017), as suggested by a reviewer. To briefly summarize the results, we found that only the full capture model (but not simpler dual-process models) was able to adequately fit the CDR 0.5 and older CDR 0 samples. However, we did not achieve an adequate fit for the younger CDR 0s using any of the tested models. Further, a final hierarchical model, with AD and age as predictors of latent parameter estimates, also failed to adequately fit the data. Despite the misspecification, we performed preliminary analyses on the age and AD effects in the interest of examining consistency with the group- and individual-level analyses. We found that the results were entirely consistent with those reported above, with one exception in a unique negative relationship between age and word generation, which was marginally significant.

It appears that the final hierarchical latent-trait model is likely misspecified, possibly due to violations of the assumptions of this approach, particularly multivariate normality of the probit-transformed parameters. Granted, the CDR 0 data were also poorly fit by the group-level MPT models. It is possible that, like the G^2 tests for the group-level models, the posterior-predictive tests of hierarchical Bayesian model fit might also be overpowered to detect small deviations as significant, due to the large sample size. Hence, the fits of these models might be relatively satisfactory after adjusting the significance threshold in a compromise power analysis. However, instead of focusing on the hierarchical analyses, our reported analyses reveal highly consistent results in both group-level models (which make the strong assumption of completely identical parameters across participants) and individual participant models (which make no assumptions about the homogeneity of participant parameters or the distribution of parameter estimates). Although the hierarchical Bayesian results largely do not contradict the group- and individual-level models (with the exception of the unique, marginal relationship between age and word generation), they require additional assumptions to reach a compromise between the two extremes, which we are likely to violate.

General discussion

The present study provides a fine-grained analysis of an episodic memory task designed to dissociate controlled and automatic retrieval processes. This task was relatively brief (10 min; 30 test trials) compared to similar tasks previously employed to estimate the same processes (i.e., over 40 min, 60–90 test trials in Jacoby et al., 2005a). From this task, we derived group- and individual-level process estimates that were sensitive to age and AD. Specifically, we tested the utility of modeling two distinct controlled retrieval processes (i.e., capture and recollection), compared to the traditional dual-process framework.

Utility of modeling a capture process

Dual-process retrieval models are typically effective in fitting a variety of memory paradigms, specifically tasks in which retrieval processes are often put in opposition by manipulating the presentation frequency of distractor and/or target items (e.g., Jacoby, 1999b). Indeed, simple dual-process (RA) MPT models can fit data from such tasks quite well. However, consistent with Jacoby et al. (2005a), MPT analyses indicate that a more complex capture model is required to accommodate data from the misleading-prime task. Dual-process PDP estimates of recollection from this task might sometimes be contaminated by an unmodeled capture process. The misleading-prime task renders an interfering prepotent response at the time of retrieval, akin to the interference produced by incongruent trials in the Stroop (1935) color-naming task. It is noteworthy that the simpler dual-process recollection estimate from this task is still a reliable and sensitive cognitive marker of aging, very mild, and preclinical AD (Millar et al., 2017). However, the present modeling work clearly indicates that it is also important to model a distinct controlled capture process.

The capture model distinguishes between two distinct modes of cognitive control during retrieval. The capture process might reflect a type of specification mode, in which retrieval processes are constrained at the front end to focus on an appropriate description of the retrieval target and to define the conditions for verifying correct retrieval (D. A. Norman & Bobrow, 1979). As noted in the introduction, such a distinction has previously proven useful in interpreting memory deficits in neuropsychological patients, particularly those that exhibit confabulation (Burgess & Shallice, 1996).

One unique benefit of MPT modeling is its ability to test conditionally-dependent relationships among distinct processes, such as how attempts at recollection might be conditional upon capture in the current model. Recently, researchers in other cognitive domains have demonstrated that MPT models might be useful in revealing other similar distinctions between conditional process models. For instance, Cooper, Greve, and Henson (2017) recently applied MPT models to item and source memory judgments in healthy younger and older adults, as well as individuals with hippocampal lesions and age-related memory problems. Critically, they found that not only were MPT parameters more consistent than analyses of raw data in these groups, but also that the interpretation of group differences differed in a model that treated source memory as dependent on item memory, as opposed to a model with the reverse conditional dependency. A similar MPT modeling approach has been used to reveal distinctions between conditional processes in illusory truth task data (Fazio, Brashier, Payne, & Marsh, 2015). MPT modeling in this task favored a model in which a memory search process is conditional upon the reliance on fluency in assessing the truth of a statement. The authors argued from this conditionally-nested relationship that in some cases, individuals may not attempt to retrieve prior knowledge to verify a statement if that statement is sufficiently fluent (e.g., through repetition priming). From a modeling perspective, this situation resembles capture-driven errors in the current model in that these errors override an attempt at controlled recollection. These reports are similar to the present study in that they examined the conditional dependencies of putative cognitive processes in a way that cannot be assessed from raw task performance data alone.

Although the current MPT models are informative as to conditional dependencies among putative processes, we cannot speak directly to the time course of capture, recollection, or automatic processes. Coane, Balota, Dolan, and Jacoby (2011) previously used rhythmic cueing and response deadlining procedures to examine the time course of controlled and automatic processes in a memory exclusion task. Importantly, they found that, under fast deadlines, participants made more exclusion errors to low-frequency, as compared to high-frequency, words. This pattern could not be explained by the *absolute,* or baseline, familiarity of the words, under which activation reflects overall lifetime exposures (i.e., frequency). Hence, these low-frequency exclusion errors were interpreted as evidence for a distinct, fast-acting *relative* familiarity process, which tracks the *increase* in mnemonic activation through repeated study exposures. This finding parallels the current report in that it expands the traditional dual-process model of recognition to a third process. However, Coane et al. (2011) argued that their results distinguished between two fast-acting automatic influences, whereas our current distinction is between two conditionally-dependent modes of attentional control. Future work combining MPT and deadlining would be useful to more fully understand the nature of these processes.

Retrieval processes in aging and AD

As predicted, aging and very mild AD were both associated with deficits in recollection. These results agree with the highly consistent evidence for recollection deficits in these populations (for meta-analysis, see Koen & Yonelinas, 2014, but see also Footnote 3[3]). More specifically, these results are consistent with the age- and AD-related deficits observed in dual-process PDP estimates of recollection derived from an overlapping dataset (Millar et al., 2017). However, the present demonstrations are unique in that the recollection deficits were observed in models that simultaneously account for capture. Hence, the recollection deficit in these populations might be distinct from group differences in capture processes.

As predicted by our more novel hypotheses, aging and AD were both associated with increases in capture. These results are in accord with previous demonstrations of increasing capture by misleading information in aging, as assessed by MPT modeling (Jacoby et al., 2005a). Analyses of false recall rates on the same misleading-prime task in traumatic brain injury are consistent with a similar capture differences in that population, as well (Dockree et al., 2006). Another approach to studying retrieval selection processes has involved testing memory for recognition foils (Jacoby, Shimizu, Daniels, & Rhodes, 2005b). This paradigm assumes that selection operates to constrain processing of potential retrieval targets in a manner consistent with the encoding context. In support of this interpretation, younger adults' memory for recognition foils is better for tests of deeply-processed blocks than for shallow-processed blocks, suggesting that at retrieval they successfully constrain processing to match the depth of encoding. Older adults, however, do not differ in memory for foils on deep or shallow blocks, suggesting less successful constraint in retrieval processing.

Previous demonstrations of age-related capture differences (Jacoby et al., 2005a) used extreme-groups designs, comparing younger (M ~ 20 years) and older adults (M ~ 75 years). The present results extend those findings to suggest that capture is sensitive to more subtle age comparisons, i.e., middle-aged (M = 61.1 years, range = 45–69) vs. older adults (M = 76.4 years, range = 70–95). These capture differences also agree with the proposals that both aging (Hasher & Zacks, 1988; Hasher et al., 1999; West, 1996) and AD (Balota & Duchek, 2015; Faust & Balota, 2007) are associated with breakdowns in attentional control processes. Failures of control might contribute to memory impairment in these populations through increased intrusions of interfering information or failures to maintain task goals, resulting in false retrieval. These memory errors might be conceptually similar to confabulation observed in neurological patients. It has been proposed that, similar to intrusion errors in the misleading-prime task, confabulation might arise due to a deficit in a phase of retrieval specification (Burgess & Shallice, 1996; D. A. Norman & Bobrow, 1979), which is similar to the role of capture in the present model.

In contrast to the controlled processes, we found no age- or AD-related differences in automatic influences, i.e., accessibility bias and word generation. We estimate that the present samples provided adequate statistical power to detect even relatively small effects (d = 0.30) in aging ($1 - \beta = .93$) and AD comparisons ($1 - \beta = .76$), but no such effects were observed. These findings are in accord with observations that age- and early-stage AD-related differences in automatic familiarity processes tend to be smaller and less consistent than differences in controlled recollection (Koen & Yonelinas, 2014).

Together, the present findings that aging and AD were both associated with (1) declines in recollection, (2) increases in capture, and (3) no changes in accessibility

[3] In one regard, the present results were inconsistent with the earlier application of the capture model to comparisons of older and younger adults. Specifically, in one experiment that applied the model to performance on a similar task, Jacoby et al. (2005a) found that older and younger adults differed in *capture only* with no difference in recollection. This version of the task was slightly different from the present task. In particular, older adults in the previous version were allowed longer encoding times during the study phase (3 s vs. 1 s) in an attempt to equate the strength of encoding between the two groups. By contrast, encoding times were fixed at 3 s for all participants in the present study. It is possible that this encoding manipulation could mitigate age-related differences in recollection, but not in capture. Such an effect would not be surprising if recollection is indeed a reconstructive process while capture reflects attentional constraint, driven by the prime presented at retrieval. Future experimental manipulations should test this hypothesis directly.

bias or word generation are consistent with the proposal that attentionally-controlled processes of episodic retrieval are uniquely disrupted in these populations.

Limitations

The capture model assumes that capture and recollection are independent retrieval processes. However, across comparisons of aging and AD, group-level increases in capture were paired with decreases in recollection. Hence, in the present study, we do not have strong evidence for the independence of these processes. In order to meet this assumption, we must rely on previous demonstrations of dissociation between these two processes, in which age differences in capture were observed even when recollection was matched between groups (Jacoby et al., 2005a). Future studies should test whether capture and recollection are independent by exploring whether these processes might be dissociable by other group comparisons or experimental manipulations, e.g., manipulations of encoding duration, retrieval response deadlines (as discussed above), or task instructions.

Additionally, we argue that capture reflects an attentionally-controlled influence on memory, but we cannot presently rule out other possibilities. For example, it is possible that the older adults and early-stage AD individuals may be less motivated to engage in demanding recollection (cf. Hess, Germain, Swaim, & Osowski, 2009) and hence more likely to be captured. We believe this outcome would still reflect changes in the extent to which the control system is fully engaged in accomplishing the goals of the task. Future studies should attempt to disentangle attentional and motivational accounts of capture, perhaps by manipulating motivation to engage recollection or by testing individual-difference relationships with other attentional or motivational measures.

There are also practical limitations associated with MPT modeling. Some researchers have argued that continuous models based on signal detection theory (SDT) are preferable to discrete MPT models of recognition task performance because MPT models predict invalid, linear ROC curves (Pazzaglia, Dube, & Rotello, 2013). However, another approach to assessing model fit uses a minimum description length framework, developed to account for differences in functional flexibility between models. In this framework, MPT models outperform SDT models, particularly for individual-level data, as used in the current study (Kellen, Klauer, & Bröder, 2013). Thus, despite their limitations in ROC fits, MPT models of recognition might be useful to the extent that they effectively compress data into theoretically informative parameters that cannot be observed in the raw data alone. However, the relevance of these arguments to the current dataset might be limited, as they are predominantly based on modeling of typical forced-choice recognition tasks, whereas the current task involves cued fragment completion. As discussed above, several recent applications of MPT have been useful in formalizing cognitive models for a variety of tasks (Cooper et al., 2017; Fazio et al., 2015).

Moreover, the group-level MPT approach also assumes homogeneity in items and individuals (Rouder, Lu, Morey, Sun, & Speckman, 2008). Model accuracy can be improved by avoiding aggregation of data at the level of items and/or individuals. Hierarchical MPT models have been developed to address this concern (e.g., Klauer, 2010), but this approach requires the further assumption that probit-transformed MPT estimates follow a multivariate normal distribution across individuals – an assumption that is likely violated in the present data (see Footnote 2). Instead, we attempted to overcome the assumption of homogeneity by estimating MPT parameters in individual participants. One further limitation of this individual-level MPT approach is that the sparse number of trials in our task might yield unreliable parameter estimates. We attempted to minimize this limitation by testing group-level differences in these individual-level estimates, rather than individual differences. Across both group- and individual-level models, our results were remarkably consistent, suggesting that the age- and AD-related differences are clear and robust.

Conclusion

In summary, the present results provide evidence of *two* distinct controlled processes involved in retrieval, hence extending standard dual-process models, which assume a single controlled process: recollection. Both capture and recollection displayed unique sensitivities to both healthy aging and early-stage AD. Similar tasks that assess controlled retrieval processes, like capture and recollection, might be particularly effective indicators of cognitive decline in aging and AD. Importantly, although dual-process models that emphasize a distinction between controlled and automatic processes have been critical in understanding memory retrieval (for review, see Yonelinas & Jacoby, 2012), the present study provides converging evidence for the importance of distinguishing multiple levels of control during retrieval.

Appendix

Table 2 Quality of fit measures for MPT models of aging differences. Model parameters include capture (C), recollection (R), accessibility bias (A), and word generation (W)

Core model	Age differences	*df*	G^2	Critical G^2	BIC	AIC	FIA
RA	–	4	366.89	46.04	386.16	370.89	191.90
RA	R	3	288.33	44.14	317.24	294.33	156.47
RA	A	3	360.90	44.14	389.81	366.90	192.55
RA	R,A	2	284.69	42.06	323.24	292.69	157.87
RAW	–	9	281.85	54.29	310.81	287.85	153.87
RAW	R	8	197.32	52.75	235.93	205.32	115.50
RAW	A	8	278.10	52.75	316.71	286.10	155.84
RAW	W	8	280.75	52.75	319.37	288.75	157.10
RAW	R,A	7	194.19	51.16	242.45	204.19	117.38
RAW	R,W	7	197.19	51.16	245.46	207.19	118.85
RAW	A,W	7	276.35	51.16	324.62	286.35	158.72
RAW	R,A,W	6	193.83	49.53	251.75	205.83	120.73
CRAW	–	8	140.20	52.75	178.82	148.20	84.51
CRAW	C	7	82.01	51.16	130.28	92.01	58.95
CRAW	R	7	72.43	51.16	120.70	82.43	55.13
CRAW	A	7	135.72	51.16	183.99	145.72	85.30
CRAW	W	7	138.87	51.16	187.13	148.87	87.16
CRAW	C,R	6	52.38	49.53	110.30[a]	64.38[a]	47.63[a]
CRAW	C,A	6	72.90	49.53	130.82	84.90	56.48
CRAW	C,W	6	80.68	49.53	138.60	92.68	61.54
CRAW	R,A	6	56.46	49.53	114.38	68.46	49.51
CRAW	R,W	6	72.37	49.53	130.29	84.37	58.06
CRAW	A,W	6	133.47	49.53	191.39	145.47	87.59
CRAW	C,R,A	5	51.69	47.83	119.26	65.69	48.84
CRAW	C,R,W	5	52.14	47.83	119.71	66.14	50.53
CRAW	C,A,W	5	72.53	47.83	140.10	86.53	59.50
CRAW	R,A,W	5	55.86	47.83	123.43	69.86	52.41
CRAW	C,R,A,W	4	51.31	46.04	128.54	67.31	51.69

[a] lowest Bayesian information criteria (BIC), Akaike information criteria (AIC), or Fisher information approximation (FIA)

Table 3 Quality of fit measures for MPT models of very mild AD differences. Model parameters include capture (C), recollection (R), accessibility bias (A), and word generation (W)

Core model	CDR differences	*df*	G^2	Critical G^2	BIC	AIC	FIA
RA	–	4	479.99	37.58	498.78	483.99	248.21
RA	R	3	247.40	35.76	275.59	253.40	135.49
RA	A	3	476.01	35.76	504.19	482.01	249.61
RA	R,A	2	246.71	33.77	284.28	254.71	138.13
RAW	–	9	421.89	45.45	450.15	427.89	223.54
RAW	R	8	147.64	43.98	185.31	155.64	90.06
RAW	A	8	421.89	43.98	459.56	429.89	227.13
RAW	W	8	407.49	43.98	445.16	415.49	219.86
RAW	R,A	7	147.22	42.46	194.31	157.22	93.04
RAW	R,W	7	141.09	42.46	188.18	151.09	89.94
RAW	A,W	7	407.38	42.46	454.46	417.38	223.38
RAW	R,A,W	6	140.98	40.90	197.49	152.98	93.21
CRAW	–	8	315.46	43.98	353.13	323.46	171.66
CRAW	C	7	198.71	42.46	245.80	208.71	116.57
CRAW	R	7	52.28	42.46	99.37	62.28	44.23
CRAW	A	7	315.35	42.46	362.43	325.35	174.39
CRAW	W	7	299.25	42.46	346.34	309.25	166.64
CRAW	C,R	6	38.51[a]	40.90	95.01[b]	50.51	39.72
CRAW	C,A	6	114.68	40.90	171.18	126.68	76.41
CRAW	C,W	6	182.60	40.90	239.11	194.60	111.53
CRAW	R,A	6	45.14	40.90	101.64	57.14	42.82
CRAW	R,W	6	46.60	40.90	103.11	58.60	44.14
CRAW	A,W	6	298.60	40.90	355.11	310.60	169.19
CRAW	C,R,A	5	37.40[a]	39.28	103.32	51.40	40.48
CRAW	C,R,W	5	31.60[a]	39.28	97.52	45.60[b]	39.02[b]
CRAW	C,A,W	5	108.40	39.28	174.32	122.40	76.23
CRAW	R,A,W	5	37.69[a]	39.28	103.61	51.69	42.11
CRAW	C,R,A,W	4	31.12[a]	37.58	106.46	47.12	40.14

[a] significant fit, i.e., G^2 < critical G^2

[b] lowest Bayesian information criteria (BIC), Akaike information criteria (AIC), or Fisher information approximation (FIA)

Author Note This research was supported by grants from the National Institute on Aging (P01-AG026276, P01-AG03991, T32-AG000030-32). Preparation of this article was supported by a traineeship from the National Science Foundation (IGERT-0548890).

We thank the participants for their dedication to the project and the Clinical Core of the Knight Alzheimer Disease Research Center at Washington University for providing clinical evaluations.

We have no conflict of interests to report. Aspects of this report were presented in a poster at the 2016 Annual Meeting of the Psychonomic Society, in Boston, MA. Separate independent analyses of the same dataset were reported in *Neuropsychology* (Millar et al., 2017).

References

Aschenbrenner, A. J., Balota, D. A., Tse, C.-S., Fagan, A. M., Holtzman, D. M., Benzinger, T. L. S., & Morris, J. C. (2015). Alzheimer Disease Biomarkers, Attentional Control, and Semantic Memory Retrieval: Synergistic and Mediational Effects of Biomarkers on a Sensitive Cognitive Measure in Non-Demented Older Adults. *Neuropsychology*, *29*, 368–381. https://doi.org/10.1037/neu0000133

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., McDermott, K. B., & Yerys, B. E. (1999). Veridical and False Memories in Healthy Older Adults and in Dementia of the Alzheimer â€TM S. *Cognitive Neuropsychology*, *16*(3/4/5), 361–384. https://doi.org/10.1080/026432999380834

Balota, D. A., & Duchek, J. M. (2015). Attention, Variability, and Biomarkers in Alzheimer's Disease. In *Remembering: Attributions, Processes, and Control in Human Memory* (pp. 285–303). New York: Psychology Press.

Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, *31*(2), 144–152. https://doi.org/10.1111/j.2044-8317.1978.tb00581.x

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for

American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Burgess, P. W., & Shallice, T. (1996). Confabulation and the control of recollection. *Memory*, *4*(4), 359–412. https://doi.org/10.1080/096582196388906

Carlesimo, G. a., & Oscar-Berman, M. (1992). Memory deficits in Alzheimer's patients: A comprehensive review. *Neuropsychology Review*, *3*(2), 119–170. https://doi.org/10.1007/BF01108841

Coane, J. H., Balota, D. A., Jacoby, L. L., & Dolan, P. O. (2011). Not all sources of familiarity are created equal: The case of word frequency and repetition in episodic recognition. *Memory and Cognition*, *39*(5), 791–805. https://doi.org/10.3758/s13421-010-0069-5

Cooper, E., Greve, A., & Henson, R. N. (2017). Assumptions behind scoring source versus item memory: Effects of age, hippocampal lesions and mild memory problems. *Cortex*, *91*, 297–315. https://doi.org/10.1016/j.cortex.2017.01.001

Dab, S., Claes, T., Morais, J., & Shallice, T. (1999). Confabulation With a Selective Descriptor Process Impairment. *Cognitive Neuropsychology*, *16*(3–5), 215–242. https://doi.org/10.1080/026432999380771

Dockree, P. M., O'Keeffe, F. M., Moloney, P., Bishara, A. J., Carton, S., Jacoby, L. L., & Robertson, I. H. (2006). Capture by misleading information and its false acceptance in patients with traumatic brain injury. *Brain: A Journal of Neurology*, *129*(Pt 1), 128–40. https://doi.org/10.1093/brain/awh664

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, *30*(3), 257–303. https://doi.org/10.1006/cogp.1996.0008

Efron, B., & Tibshirani, R. (1994). *Introduction to the Bootstrap*. https://doi.org/10.1214/ss/1063994971

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–60. https://doi.org/10.3758/BRM.41.4.1149

Faust, M. E., & Balota, D. A. (2007). Inhibition, Facilitation, and Attention Control in Dementia of the Alzheimer Type: The role of unifying principles in cognitive theory development. In D. S. Gorfein & C. McLeod (Eds.), *The place of inhibition in Cognition.* (pp. 213–238).

Faust, M. E., Balota, D. A., Duchek, J. M., Gernsbacher, M. A., & Smith, S. (1997). Inhibitory Control during Sentence Comprehension in Individuals with Dementia of the Alzheimer Type*1, *2, *3. *Brain and Language*, *57*(2), 225–253. https://doi.org/10.1006/brln.1997.1747

Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge Does Not Protect Against Illusory Truth. *Journal of Experimental Psychology: General*, *144*(5), 993–1002. https://doi.org/10.1037/xge0000098

Flavell, J. H. (1970). Developmental Studies of Mediated Memory. *Advances in Child Development and Behavior*, *5*, 181–211.

Hanley, R. J., Davies, A. D. M., Downes, J. J., & Mayes, A. R. (1994). Impaired recall of verbal material following rupture and repair of an anterior communicating artery aneurysm. *Cognitive Neuropsychology*, *11*(5).

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. Bower (Ed.), *The psychology of learning and motivation, 22* (pp. 193–225). New York: Academic Press.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory Control, Circadian Arousal, and Age. In D. Gopher & A. Koriat (Eds.), *Attention and performance* (XVII, pp. 653–675). Cambridge, MA: MIT Press.

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. https://doi.org/10.3758/BF03203619

Heck, D. W., Arnold, N. R., & Arnold, D. (2017). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, 1–21. https://doi.org/10.3758/s13428-017-0869-7

Hess, T. M., Germain, C. M., Swaim, E. L., & Osowski, N. L. (2009). Aging and selective engagement: The moderating impact of motivation on older adults' resource utilization. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *64*(4), 447–456. https://doi.org/10.1093/geronb/gbp020

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

Jacoby, L. L. (1996). Dissociating Automatic and Consciously Controlled Effects of Study/Test Compatibility. *Journal of Memory and Language*, *35*, 32–52. https://doi.org/10.1006/jmla.1996.0002

Jacoby, L. L. (1999a). Deceiving the Elderly: Effects of Accessibility Bias in Cued-Recall Performance. *Cognitive Neuropsychology*, *16*(3–5), 417–436. https://doi.org/10.1080/026432999380861

Jacoby, L. L. (1999b). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 3–22. https://doi.org/10.1037/0278-7393.25.1.3

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005a). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, *134*(2), 131–148. https://doi.org/10.1037/0096-3445.134.2.131

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005b). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review*, *12*(5), 852–857. https://doi.org/10.3758/BF03196776

Kellen, D., Klauer, K. C., & Bröder, A. (2013). Recognition memory models and binary-response ROCs: A comparison by minimum description length. *Psychonomic Bulletin and Review*, *20*(4), 693–719. https://doi.org/10.3758/s13423-013-0407-2

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-class approach. *Psychometrika*, *75*(1), 70–98.

Koen, J. D., & Yonelinas, A. P. (2014). The Effects of Healthy Aging, Amnestic Mild Cognitive Impairment, and Alzheimer's Disease on Recollection and Familiarity: A Meta-Analytic Review. *Neuropsychology Review*, 332–354. https://doi.org/10.1007/s11065-014-9266-5

Kopelman, M. D. (1999). Varieties of False Memory. *Cognitive Neuropsychology*, *16*(3–5), 197–214. https://doi.org/10.1080/026432999380762

Macmillan, N. A, & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, *98*(1), 185–199. https://doi.org/10.1037/0033-2909.98.1.185

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222–43. https://doi.org/10.1037/a0017619

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*, 939–944. https://doi.org/10.1212/WNL.34.7.939

Millar, P. R., Balota, D. A., Maddox, G. B., Duchek, J. M., Aschenbrenner, A. J., Fagan, A. M., … Morris, J. C. (2017). Process Dissociation Analyses of Memory Changes in Healthy

Aging, Preclinical, and Very Mild Alzheimer Disease: Evidence for Isolated Recollection Deficits. *Neuropsychology*, *31*(7), 708–723. https://doi.org/10.1037/neu0000352

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–4.

Morris, R. G., & Kopelman, M. D. (1986). The memory deficits in Alzheimer-type dementia: A review. *Quarterly Journal of Experimental Psychology*, *38A*(October), 575–602. https://doi.org/10.1080/14640748608401615

Moscovitch, M. (1989). Confabulation and the Frontal Systems: Strategic versus Associative Retrieval in Neuropsychological Theories of Memory. In *Varieties of memory and consciousness: essays in honour of Endel Tulving* (pp. 133–161).

Norman, D. A., & Bobrow, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, *11*(1), 107–123. https://doi.org/10.1016/0010-0285(79)90006-9

Norman, K. A., & Schacter, D. L. (1996). Implicit memory, explicit memory, and false recollection: A cognitive neuroscience perspective. In L. M. Reder (Ed.), *Implicit memory and metacognition*. Hillsdale, NJ: Erlbaum Associates.

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), 1173–203. https://doi.org/10.1037/a0033044

Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer ' s disease A critical review, 383–404.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, *95*(3), 318–339. https://doi.org/10.1037//0033-295X.95.3.318

Rouder, J. N., Lu, J., Morey, R. D., Sun, D. C., & Speckman, P. L. (2008). A hierarchical process-dissociation model. *Journal of Experimental Psychology: General, 137*(2), 370–389. https://doi.org/10.1037/0096-3445.137.2.370

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annu Rev Psychol*, *49*, 289–318. https://doi.org/10.1146/annurev.psych.49.1.289

Singmann, H., & Kellen, D. (2013). MPTinR: Analysis of multinomial processing tree models in R. *Behavior Research Methods*, *45*(2), 560–75. https://doi.org/10.3758/s13428-012-0259-0

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop Performance in Healthy Younger and Older Adults and in Individuals With Dementia of the Alzheimer's Type, *22*(2), 461–479.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

West, R. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292. https://doi.org/10.1037/0033-2909.120.2.272

Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition*, *40*(5), 663–680. https://doi.org/10.3758/s13421-012-0205-5