Published in final edited form as:
Psychol Aging. 2012 June ; 27(2): 418–428. doi:10.1037/a0024786.

Tests of the DRYAD theory of the age-related deficit in memory for context: Not about context, and not about aging

Aaron S. Benjamin,
University of Illinois at Urbana-Champaign

Michael Diaz,
Washington University in St. Louis

Laura E. Matzen, and
Sandia National Laboratories

Benjamin Johnson
University of Illinois at Urbana-Champaign

Abstract

Older adults exhibit a disproportionate deficit in their ability to recover contextual elements or source information about prior encounters with stimuli. A recent theoretical account, DRYAD (Benjamin, 2010), attributes this selective deficit to a global decrease in memory fidelity with age, moderated by weak representation of contextual information. The predictions of DRYAD are tested here in three experiments. We show that an age-related deficit obtains for whichever aspect of the stimulus subjects' attention is directed away from during encoding (Experiment 1), suggesting a central role for attention in producing the age-related deficit in context. We also show that an analogous deficit can be elicited within young subjects with a manipulation of study time (Experiment 2), suggesting that any means of reducing memory fidelity yields an interaction of the same form as the age-related effect. Experiment 3 evaluates the critical prediction of DRYAD that endorsement probability in an exclusion task should vary nonmonotonically with memory strength. This prediction was confirmed by assessing the shape of the forgetting function in a continuous exclusion task. The results are consistent with the DRYAD account of aging and memory judgments and do not support the widely held view that aging entails the selective disruption of processes involved in encoding, storing, or retrieving contextual information.

As the world's population ages (U. S. Bureau of the Census, World Population Profile, 1998), it is increasingly important to understand the nature and extent of memory deficits in older adults, and to provide appropriate theoretical frameworks for predicting, measuring, and characterizing problems with memory. Because these frameworks help suggest approaches to developing rehabilitation regimens and designing appropriate environmental and technological support, there is much at stake in evaluating competing theoretical perspectives. Benjamin (2010) recently proposed a theoretical model, DRYAD, of age-related deficits in tasks that involve memory judgments, including *recognition* (judging whether a stimulus was previously encountered), *source memory* (deciding where it was previously encountered), and *recognition exclusion* (deciding whether it was encountered in

Correspondence: Aaron Benjamin, ASBENJAM@ILLINOIS.EDU.

a particular context). DRYAD, which stands for *Density of Representations Yields Age-related Deficits*, postulates a global deficit in memory fidelity in older adults and, as such, contrasts with most other extant accounts of recognition, source memory, and exclusion. Those theories suggest that a specific deficit of one sort or another in the elderly is the cause of age-related impairments in those tasks.

There are two main classes of experiments that are relevant, and they differ only in how memory is tested. In both classes, subjects study a list of stimuli (words or pictures, for example) that vary in some contextual or source attribute. For example, words may be spoken by different speakers, be presented in different lists, or be presented in different colors. Usually, there are only two such contexts, but not always.

The first type of experiment assesses memory for the stimuli by requiring subjects to make a yes/no recognition decision about stimuli that were or were not actually in the list, and follows that task with an additional source memory decision, usually of the forced-choice variety, in which they choose which of the possible contextual attributes accompanied that stimulus. In these studies, it is common to find that the measure of source memory differs more between younger and older subjects than does the measure of recognition (Spencer & Raz, 1995).

The second type of study uses the exclusion task (Jacoby, 1991) to combine the sources of information that underlie recognition and source memory. In that task, subjects are asked to endorse items if and only if they had been previously studied and possessed a particular contextual attribute (e.g., they were spoken by a female, appeared in the first list, or were printed in red). Two data are relevant from this task. First, older adults appear more prone than younger adults to mistakenly endorsing previously seen items from the contraindicated context (e.g., words spoken by a male, words from the second list, or words printed in blue). Second, manipulations that enhance learning of the list sometimes increase the frequency of this type of error in the elderly but decrease the frequency in the young (Benjamin, 2001; Benjamin & Craik, 2001; Jacoby, 1999; Light, Chung, Pendergrass,& Van Ocker, 2006). Such a disordinal interaction is particularly compelling, and has been taken to imply that the strength manipulation increases the familiarity of contraindicated items but is not offset by a greater likelihood of recollecting the context of the prior encounter with those items in the elderly.

These two patterns of results converge to tell a consistent story about older adults' inability to recover the context of prior experiences. They can neither produce that context when asked directly nor use that context to guide their responding in the exclusion task. The contrast to be drawn here is between theories that explain this pattern by postulating a selective impairment in the processes or stores that support memory for context and theories that postulate a general memory deficit that, in its particulars, plays out in impaired memory for context.

Theories that postulate selective impairment

One prominent suggestion is that older adults lack the ability to engage in effective source monitoring (Hashtroudi, Johnson, & Chrosniak, 1989). The compatible aspect of this claim with the current theory is the idea that the task of resolving which prior source or context accompanied a stimulus involves accessing the same representations and using the same processes as any other memory decision. There is no need to postulate processes or memory systems that are specialized for the storage or retrieval of contextual information specifically. Where this theory diverges from DRYAD is in its statement about the relationship between the theoretical construct of source monitoring and in the empirical measurement of source attributions. The authors plainly state, "[d]ifferences in age deficits…would indicate a more selective disruption of source monitoring processes with age." (p. 107) In DRYAD, interactions between age and other variables are not taken a priori to indicate selective deficits. This is an important point that will be clarified in greater detail when we return to discuss DRYAD.

Another important theory in the literature holds that older adults exhibit impaired recollection, which is the process used to retrieve prior experiences (Jacoby, 1998, 1999). When this retrieval process is impaired, familiarity plays a larger role in memory judgments. Among its many successes, this theory provides the most straightforward explanation of why strength manipulations decrease exclusion errors in the young but increase them in older adults (Benjamin, 2001; Benjamin & Craik, 2001; Jacoby, 1999; Light et al., 2006). DRYAD challenges this theoretical suggestion by showing that this result can arise in a simpler system that presupposes no selective impairment. That is, the empirical *interaction* need not imply a selective process *dissociation*.

A third significant theory holds that failure to remember context reflects a more general deficit in "binding" between any two aspects of a complex stimulus (Naveh-Benjamin, 2000; see also Chalfonte & Johnson, 1996)—in our case, in making the appropriate connection between the "item" and the "context" of an experience. Older adults are hypothesized to suffer from an *associative deficit* that impairs their ability to bind together aspects of an experience, but this suggestion poses certain problems that are solved by DRYAD. To understand what is (or is not) successfully bound, one must have a theory about what the elements to be bound are. This requires a strong definition of what an "item" is and what "context" is. In DRYAD, there are no items and contexts, so no such theory is required. And, indeed, the experiments here demonstrate that, rather than thinking about the world in terms of "items" and "contexts," we gain better leverage on these problems simply by thinking about the distribution of attention. The net result is that DRYAD can explain the memory judgment data that the associative deficit is intended to explain, but it does so without postulating a selective impairment in older adults.

This is not to say that DRYAD stands to explain all age-related deficits in associative memory (e.g., Old & Naveh-Benjamin, 2008). Some cases are likely incompatible with the assumptions of DRYAD. However, there are also data in the context-memory literature that are difficult to reconcile with a postulated associative memory deficit. These two theoretical approaches cover somewhat different empirical ground, a point we will return to later.

The theoretical suggestions reviewed above all postulate a particular process that is impaired in older adults. These contrast with DRYAD, which assumes only that memory is less veridical in older adults. We turn now to the way in which that general memory deficit produces the relevant statistical interactions.

DRYAD and the representation of information

In a typical experiment investigating memory for items and contexts, words are presented in, for example, one of two colors, or in one of two locations on the screen. From the experimenter's perspective, the items are the words, and the contexts are the accompanying color or location information. DRYAD represents information without respect to the artificial boundaries dividing "items" and "contexts." Information in the environment that is well attended elicits dense encoding (i.e., encoding of many pieces of information). Information that is unattended or poorly attended elicits sparse encoding. What results is an attention-weighted record of experience. In DRYAD, that record is a vector of values, with greater dimensionality accorded to aspects of the stimulus that elicited greater attention. In most experiments, the contextual or source attributes of to-be-learned stimuli are less interesting and less variable, and sometimes even de-emphasized, so context is typically represented more sparsely than focal item information. This process is not hypothesized to differ between younger and older adults. Although there are certainly cohort differences in what things are interesting to younger and older people, those differences are probably not very important for such artificial experimental tasks and are not necessary to account for in understanding performance on those tasks.

What does differ between younger and older adults is the overall probability that a given piece of information, once attended, is encoded and maintained accurately in memory. Older adults are presumed to have lower memory fidelity. This is the global deficit that is at the heart of DRYAD. The relative resistance of aspects of a stimulus to this global deficit is a positive function of its encoding density; thus, sparsely encoded aspects of a stimulus are most dramatically affected on behavioral testing by this global deficit.

The contrast between the selective-deficit view and DRYAD is illustrated in Figure 1. There, the to-be-remembered stimulus is a car on a road. According to selective-deficit views, the resulting representation of the stimulus differs between younger and older adults, such that aspects that are related to the context (the road) are more poorly maintained (or accessed) in the elderly. Little difference in the representation of the item (the car) is expected. This combination of claims is depicted in the top part of the figure by the poorer contrast in the image of the background for the older adults. According to DRYAD, memory for the entire stimulus is poorer for older adults—thus, in the bottom part of the figure, contrast is poorer across the entire stimulus. However, the greater attention paid to the car during encoding (indicated by its greater size and brightness) offsets this deficit. In both cases, what results is a disproportionate deficit in memory for the background in older adults.

DRYAD attributes the apparently selective impairment in context or source memory not to the failure of a selective process or mechanism, but rather to a global deficit in memory fidelity in the elderly. In this paper we evaluate the validity of two major components, and one central prediction, of DYRAD. DRYAD asserts no representational distinction between items and contexts; rather, aspects of stimuli are encoded to the degree that they are attended to. Under most conditions, the nominal context of a stimulus elicits less attention, but the nature of what is experimentally defined as ”context” plays a major role. Evidence for this claim comes from experiments that use particular stimulus or instructional characteristics to eliminate this confound. In such tasks, age-related deficits in context memory are reduced or disappear (May, Rahhal, Berry, & Leighton, 2005; Rahhal, May, & Hasher, 2002).

Our first experiment assesses whether an instructional manipulation—to attend to one or another part of a complex stimulus—leads to a selective age-related deficit in memory for the less attended aspects of the stimulus. If an age-related deficit can be shown as a function of attention, then that finding would undermine the widely held claim that older adults are selectively deficient in memory for context. That is, such a finding redirects our theoretical focus from aspects of the stimulus—item or context—to aspects of the learner’s interaction with that stimulus—greater or lesser attention.

The second experiment reported here addresses the question of whether a selective deficit in memory for context can arise from a global reduction in memory fidelity. In that experiment, we demonstrate that, in younger adults, a reduction of study time leads to a disproportionate reduction in memory for contextual aspects of a stimulus. That finding demonstrates that a selective deficit in context memory can arise as the consequence of a global reduction in memory fidelity, rather than from a failure of specific contextual encoding or recovery mechanisms.

The final experiment in this paper tests a novel prediction of DRYAD. DRYAD accounts for the disordinal interaction between manipulations of learning and age group with a nonmonotonic relationship between memory strength and endorsement probability. This nonmonotonicity is a natural consequence of the differential representational density of item and contextual information. We test for the presence of such a nonmonotonicity by manipulating the study-test interval in a continuous exclusion task in Experiment 3.

The origin of the global deficit

DRYAD provides an existence proof of the claim that a global deficit can yield empirical interactions of the form typically taken to indicate a selective deficit. However, it does not provide strong constraints on the origin of that deficit, and consequently has little to say about the validity of a number of competing theories. The global deficit could be a result of limited attentional resources (Craik & Byrd, 1982) leading to less accurate sampling of the stimulus per unit time. An appealing aspect of that suggestion is that it provides a natural means of understanding the parallels between aging and divided attention (Castel & Craik, 2003; Naveh-Benjamin, Craik, Guez, & Kreuger, 2005), response speeding (Benjamin, 2001; Benjamin & Craik, 2001; Jacoby, 1999; Light et al., 2006), and other performance-limiting conditions.

A related suggestion is that older adults may exhibit lower processing speed (e.g., Salthouse, 1996). Similar to the description provide above, if memory fidelity reflects something about the quantity of sampling, then limiting the rate of that sampling process could lead to lower accuracy in memory. Even a variant of associative-binding deficit hypothesis (Naveh-Benjamin, 2000) is a potential candidate for understanding a global deficit, if one thinks of encoding as the act of binding all of the sampled elements from a stimulus to one another. If that process is slow or inaccurate, the resulting composite memory might suffer. Note, however, that the associative deficit hypothesis is usually characterized as binding items with contexts (or with each other), a construal that is at odds with DRYAD's lack of a distinction between "items" and "contexts."

Measurement of memory for "items" and "contexts"

One theme revisited throughout this paper is the difficulty involved in comparing memory for items and memory for context in traditional tasks. This difficulty owes to the fact that these two tests usually entail quite different formats: whereas memory for items is often assessed in a yes/no recognition task, memory for context is more typically assessed using *n*-alternative forced-choice recognition. More often than not, no care is taken to ensure compatibility between these measures. In the first experiment presented here, we will use a traditional design and compare measures using traditional techniques as well as more theoretically motivated approaches. However, even in Experiment 2, in which the full force of unequal-variance Signal Detection Theory (Green & Swets, 1966; Macmillan & Creelman, 2005) is brought to bear in deriving compatible measures, an age-related deficit in memory for context will be apparent when study time is sufficient. What is in question here is the nature of this deficit. DRYAD attributes this behavioral deficit to a global failure of memory in the elderly, not to a failure of a subset of processes or mechanisms that specifically underlie context memory.

In the first experiment, we assess whether age-related deficits can be made to appear for different aspects of stimuli based on instructions to attend to one aspect of the stimulus or another. These experiments bear on the assumption of the model that "items" and "contexts" are constructs devoid of psychological reality. Instead, reliable relationships between those experimental constructs and behavior appear only because of the differential attention that they typically induce. Notice that this claim does not suggest that context has no role to play in memory. It is clear that it plays an exceptional role from encoding to retrieval. Rather, what is contested here is whether specialized processes or systems are devoted to the maintenance or retrieval of contextual information.

In Experiment 1, we use verbal stimuli and direct subjects to attend to one of the nouns in the sentence. The critical test in both experiments is whether the less-attended aspect of the stimulus yields an age-related deficit and the more-attended does not, regardless of what those aspects actually are.

Experiment 1: Orientation of attention and memory for sentences

The stimuli in this experiment are short sentences of a common form: "NAME is a JOB." Subjects were asked to read the entire sentence, but to pay special attention to either the subject (the name) or the predicate nominative (the job). The critical contrast is whether older adults exhibit a selective deficit for the aspect of the sentence that attention was drawn away from during encoding. Pilot research using similar materials had indicated that performance on the "item memory" portion of this task was likely to be quite similar between age groups. This should help avoid complications that can arise due to differences in baseline performance.

Method

Participants—Twenty-nine undergraduate students from the University of Illinois at Urbana-Champaign comprised the younger adult group, and participated in partial fulfillment of a course requirement. The older group included 27 adults from the Urbana-Champaign community who were paid for their participation. In this and in Experiments 2 and 3, the elderly subjects were drawn from a large pool of volunteers over the age of 60 (range = 60 – 83, mean age = 69) with normal scores on the MMSE (25 or above, mean score = 29). Performance on the Shipley vocabulary test ranged from 28 to 40 (mean score = 36) for the older adults and from 25 to 39 (mean = 32) for the younger adults. The older adult sample is comprised of 30% males and 70% females; the younger adult sample is 38% male and 62% female. Fourteen younger and fourteen older subjects were randomly assigned to the attend-name condition, and the rest to the attend-job condition.

Materials—Sixty common last names from a range of ethnicities were drawn from varying websites listing last names. Sixty jobs were drawn from published lists of common occupations. A study list was created individually for each subject by pairing a random subset of 30 names with a random subset of 30 jobs in sentences of the form *NAME is a JOB*. A test list was constructed individually for each subject by randomly ordering all 60 names or jobs (depending on which aspect was emphasized during study, as described below).

Procedure—Participants were told that they would be studying a set of sentences that they should read and try to remember. They were given an example sentence and told that either the name or the occupation was the most important aspect to remember. The study phase presented the sentences for 3s each with 1s between sentences. After all 30 study sentences, participants were given a distractor task in which they performed simple 1-to 2-digit addition and subtraction problems. This phase terminated after 60s.

For the test phase, subjects made yes/no recognition decisions for either all 60 names or all 60 jobs. If they had been told to pay special attention to the jobs, then the 60-item recognition list consisted of jobs, half of which had been studied and half of which had not. Immediately following each recognition decision, a 30-item forced-choice list of either names or jobs (whichever was deemphasized during study) appeared and participants were asked to pick the 'context' that had been previously paired with that 'item' (following the example described above, to pick the name associated with the queried job). The items on the forced-choice test were exclusively the ones that were viewed during study, and the entire test was self-paced. These were presented on the computer screen in the same format as the study stimuli, in two columns of 15.

Results

The results from Experiment 1 are presented in Figure 2. All results are significant at the $\alpha < .05$ level unless otherwise noted. The critical prediction is that the age-related impairment should be greater for the aspect of the stimulus to which attention is not directed. We evaluate this claim using two measures, one traditional and one less so, which we describe in the next two subsections.

Traditional analysis—As we noted before, a direct comparison of item memory and context memory is difficult in an experiment like this one, because item memory is measured using yes/no recognition and context memory is measured using forced-choice recognition. One way of placing these measures on the same scale is to convert item recognition performance to a measure of percentage correct (PC), which occupies the same range of values as context memory (namely, 0 to 1). In the next section we present an in-depth criticism of such an approach, but use it here to more closely adhere to standards in this literature. The data are shown in the top row of Figure 2.

When looking at percent correct, the main effects of age ($F[1,54] = 8.8$, $MSE = .02$) and test type ($F[1,54] = 6.1$, $MSE = .01$) were reliable. In addition, the interaction between memory test and age group was significant for both conditions put together ($F[1,54] = 11.0$, $MSE = .01$), marginally for the attend-names condition ($F[1,26] = 3.4$, $p = .08$), and for the attend-jobs condition ($F[1,26] = 7.9$, $MSE = .01$). For the latter two cases, the interaction was in the appropriate direction: memory was selectively worse in the elderly for the less well attended aspect of the sentence. These results indicate that deprioritized information during encoding is not only learned less well, it exhibits a greater age-related impairment.

Detection-theoretic analysis—This second analysis takes more seriously the need to provide a common metric on which item and context memory can be compared. It does so by relating performance on both measures to parameters in signal-detection theory (Green & Swets, 1966) which, among other things, provides a natural means of comparing yes/no and forced-choice performance. In that theory, to-be-endorsed items are discriminated from to-be-rejected ones by virtue of a deterministic criterion placed to effectively divide two overlapping distributions of evidence (cf. Benjamin, Diaz, & Wee, 2009). The actual locations and shapes of the distributions determine a decision-free measure of discriminability, d. When those distributions are of equal variance, the measure is called d' and simply measures the distance between the means in units of their common variance.

In forced-choice recognition, the task is to choose, out of n alternatives, the stimulus that yields the highest evidence value. Because the distractor stimuli are drawn from a common pool, the distribution of their evidence values can be characterized with a single distribution and the expected maximum of n events can be described by numerical means (e.g., Hacker & Ratcliff, 1979). This provides a means of estimating d' from n-alternative forced-choice recognition, and this value has the same interpretation and measurement qualities as d' as estimated from yes/no recognition.

As we shall discuss in the introduction to Experiment 2, this approach has its limitations as well. Although it is properly theoretically motivated, there are aspects of equal-variance signal-detection theory that are inadequate for recognition, and somewhat more complex procedures and analyses are required. Nonetheless, *d´* provides a more reasonable means of comparing scores directly than traditional analysis allows (Swets, 1986).

The results from this analysis are shown in the bottom row of Figure 2. The analysis using *d´* confirmed the results seen in the analysis using percent correct. The main effect of age (F[1,54] = 3.5, MSE= .54) and test type (F[1,54] = 3.9, MSE = .24) were again reliable. More importantly, the interaction between test type and age group was significant for both conditions (F[1,54] = 11.1, *MSE* =.24), for the attend-names condition (F[1,26] = 5.8, *MSE* =.27), and for the attend-jobs condition (F[1,26] = 5.0, *MSE* =.21).

Discussion: Experiment 1

The phenomenon of interest in this paper is the greater impairment evident in tests of context memory relative to tests of item memory for older adults. The first experiment evaluated whether a manipulation of attention could yield an analogous effect. This claim confirmed the results of a pilot experiment, not reported here, in which “items” and “contexts” were implemented more literally as objects (chairs) in environments (rooms), and attention was directed either towards the object or the environment. Regardless of the nature of the stimuli, in DRYAD, attention determines the relative number of features about context that are encoded, and memory fidelity—assumed to be poorer in the elderly—determines the quality of that encoding. DRYAD shows how more sparsely represented aspects of a stimulus yield a disproportionate deficit when a global decrease in memory fidelity is imposed, and explains older adults’ deficit in memory for context as an example of this more general phenomenon.

The goal of the first experiment has been to demonstrate this generality in a a paradigm that is roughly similar to the ones used in experiments investigating context memory. The hypothesis was confirmed: Older adults showed a memory deficit for the aspect of the stimulus that the instructions deemphasized.

One way of characterizing this result is that the well-known age-related deficit in memory context is not about context at all. Rather, it is about representational density. When things are densely encoded, the memory deficit that older adults exhibit affects eventual recovery of that information less dramatically than if they had been only sparsely encoded. This interpretation provides a key role for redundant coding: Having rich or redundant information can allay the problems that arise with low-quality encoding.

According to DRYAD, the age-related deficit in context memory is not only not about context, it is not a selective *age*-related deficit at all! Rather, the decreased memory fidelity associated with aging is at fault. This view differs from most theories of age-related problems with context memory, which attribute deficits to specific systems or processes. DRYAD claims that the deficit is an inherent and unavoidable consequence of low representational fidelity acting on sparsely encoded information. Thus, DRYAD predicts that any means of reducing memory fidelity—forgetting, decreased study time, distraction—should yield the interaction that is the hallmark of the age-related effect. This prediction is the focus of Experiment 2.

Experiment 2: Effect of study time on memory for words and locations

In Experiment 2, we evaluated the effect of reducing study time on memory for items and memory for contexts. There are two methodological concerns that were less important in the first experiment, but are of paramount importance here. Since the goal of that experiment was to demonstrate that attention shifts the locus of the age-related deficit, attention was directly manipulated. The goal of Experiment 2 is to assess the effects of study time on memory for items and memory for contexts in as similar a fashion as possible to extant work on context memory, and consequently we chose study materials and encoding instructions that follow more directly from the modal study in the field.

Second, we deal more substantively with the measurement issues raised in the analyses of Experiment 1. Specifically, we designed our testing procedure such that equivalent valid measures of item memory and context memory could be assessed. We did so by making both tests yes/no decision tasks, and by collecting confidence measures that allow derivation of a superior detection-theoretic measure, d_a. (e.g., Diaz & Benjamin, 2008; Matzen & Benjamin, 2009) In addition, though the performance of older adults is not directly relevant to the central hypothesis of this experiment, we collected data from older adults to evaluate whether the age-related effect seen in Experiment 1 and throughout the literature generalized to a superior and more directly comparable measure of performance.

Method

Participants—Sixty-six younger and 35 older adults participated in the experiment. The younger subjects were undergraduates who participated for partial course credit. The older adults came from the pool described in Experiment 1. One younger subject was dropped from all analyses because of a failure to use sufficient range on the confidence scale and a consequent inability to fit the detection model.

Materials and Procedure—Materials consisted of 1400 medium to low frequency words. For each participant, a random subset of 240 words was selected. One-hundred and twenty of these words were used as study words, and one-hundred twenty were used as study foils. Study words were presented one at a time on either the left or the right side of a fixation point in the center of a computer monitor. Half of the study words were presented on the left and half were presented on the right. In each location, 40 words were seen for 500 msec (weak encoding condition), 40 for 2000 msec (medium encoding condition), and 40 for 5000 msec (strong encoding condition). All words had an ISI of 1000 msec.

Immediately following the study phase, participants were tested on all 240 words. For each word, participants first made a yes/no recognition judgment followed by a three-point confidence judgment for that response. After each recognition and confidence judgment, participants were given a source-exclusion judgment for that word. Half of the words for each strength x location condition were given left-exclusion judgments and the other half were given right-exclusion. For example, in left-exclusion, participants were asked if the word had appeared on the left, and were told to answer “no” if it had either appeared on the right or if it was not studied. After each exclusion judgment, participants were asked to rate their confidence in that judgment on a three point scale.

Results

Age-related deficit in memory for context—The joint effects of age and test type can be seen in Figure 3 as a function of study time. For recognition, d_a represents the discriminability between old items (which varied with study time) and new items (which did not). For exclusion, d_a represents the discriminability between to-be-endorsed (TBE) items and to-be-excluded (TBX) items (each of which varied with study time). The unequal-variance signal-detection model was fit to individual subjects using maximum-likelihood estimation.

When collapsing across study time, neither the main effect of age ($F[1,99] = 2.8$, $MSE = .35$) nor the interaction between age and test type was significant ($F[1,99] = 2.5$, $MSE = .09$, $p = .12$). However, when comparing the magnitude of the interaction at each study time, as shown in Figure 3, it can be seen that the magnitude increases with study time. In fact, although the interaction was not even numerically apparent at the shortest study time ($F[1,99] < 1$, $MSE = .13$, $p = .94$), it was in the appropriate direction but only marginally significant at the middle study time ($F[1,99] = 3.1$, $MSE = .13$, $p = .08$) and was significant and of quite impressive magnitude at the longest study time ($F[1,99] = 7.4$, $MSE = .13$). This result does not bear on the questions raised in this paper but does confirm the claim in the literature that the magnitude of the age-related deficit increases with memory strength (e.g., Spencer & Raz, 1995).

Effect of study time—The effects of study time on memory for item and contextual information in younger subjects can be seen in Figure 4. Study time positively affected memory ($F[2,130] = 14.4$, $MSE = .14$), and performance was superior on recognition than on exclusion ($F[1,65] = 106.7$, $MSE = .13$), both as expected. More importantly, as predicted by DRYAD, the effect of study time was greater on the exclusion task than on the recognition task ($F[2,130] = 4.3$, $MSE = .09$).

Discussion

The critical prediction in this experiment concerned the effect of study time on item and context memory. If the assertion of DRYAD is correct that any reduction of memory fidelity will elicit a greater effect on sparsely encoded information, like context, then the effect of study time should be greater on the exclusion task than on the recognition task. This prediction was confirmed.

An interesting and unexpected finding in this experiment was that the magnitude of the age-related deficit increased with increasing study time. This result suggests that the division of attention across the word and its location varied systematically with study time. Though DRYAD contains no mechanism by which to account for such a finding, one extant model has considered how attention to context might vary with study time (Malmberg & Shiffrin, 2005) and could be integrated into DRYAD for future exploration.

Experiment 3: Nonmonotonic false-alarm rate functions in exclusion

The experiments reported thus far in this article are relevant to DRYAD's ability to account for age-related interactions in experiments in which item and context memory are assessed separately. A more important result, however, concerns DRYAD's account of disordinal interactions in recognition exclusion. In the exclusion paradigm, subjects are asked to endorse only a subset of previously studied items (for example, previously red but not previously blue items), thus placing a burden on subjects to recover the context of that previous presentation.

The critical finding with respect to aging is that some manipulations that increase hits have opposite effects on false alarms to to-be-excluded items in younger and older subjects: they decrease with additional memory in younger subjects but increase in older subjects. This result obtains with study repetitions (Jacoby, 1999) and increased spacing of repetitions (Benjamin & Craik, 2001), and analogous results are apparent in plurality-reversal recognition (Light et al., 2006), associative recognition (Jones & Jacoby, 2001), and semantic false memory (Benjamin, 2001; Watson, McDermott, & Balota, 2004).

The traditional account of such findings is that the ability to recollect the study context is normally enhanced by additional learning, and that this recollective ability offsets the greater familiarity engendered by the additional learning. That is, subjects are able to reject high-strength contraindicated items despite their high familiarity, because they are more likely to recollect their prior study context. Subjects with weakened recollective ability, like older adults, thus reveal the unmasked effects of underlying familiarity.

DRYAD accounts for this disordinal interaction quite differently. Memory strength affects the fidelity of the memory trace for the event. Because item information is more densely represented than context information, the relationship between memory strength and the probability of recovering information differs for item information and context information. At very low levels of learning, increasing memory strength brings about the ability to recover item information but is not sufficient to support context recovery. At higher levels of learning, more context information becomes recoverable. The differential course of recovery for the two types of information leads to a nonmonotonic relationship between memory strength and exclusion performance, as shown schematically in Figure 5 (see also Matzen, Taylor, & Benjamin, in press).

DRYAD attributes the disordinal interaction between age and memory strength to the fact that older and younger adults fall at different points on the functions shown in Figure 6. Older adults start at lower levels of memory fidelity and are in a range of the function in which additional learning increases false alarms to to-be-excluded (TBX) items. Younger subjects have higher levels of memory fidelity and are in a range of the function in which additional learning decreases false alarms to these items. Benjamin (2010) simulated this result and showed that nonmonotonic false-alarm rate functions in exclusion will result when context is represented more sparsely than item information.

The purpose of Experiment 3 is to evaluate this prediction experimentally. Memory strength was manipulated by varying the retention interval, and the critical prediction is that false alarms to to-be-excluded items should rise, and then fall, with increasing retention interval.

Method

In the experiment reported here, we use a variant of the continuous recognition procedure (Shepard & Teghtsoonian, 1961), combined with the exclusion paradigm (Jacoby, 1991), to measure the effects of memory fidelity (operationalized as retention interval) on hit and false-alarm rates to previously studied words. Words were studied in either blue or yellow, and test words were presented, after a variable interval, in either blue or yellow. Subjects were instructed to endorse items that were presented in the same color as on previous trials and to reject both new items and items for which the color had changed.

Subjects—Forty-seven undergraduates from the University of Illinois participated in partial fulfillment of a course requirement.

Materials—Words used in the experiment were four-to eight-letter nouns selected from the MRC Psycholinguistic Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm) and constrained on Kucera-Francis written frequency (50–500), imagability (200–500), and concreteness (200–500). Words were randomly assigned to be studied or not prior to test; of those assigned to be studied, half were designated to be TBX items (different color) at test and half were designated to be TBE items (same color). Half of each of these categories were studied in blue; the other half was studied in yellow. Half of the new items were assigned to be blue and half to be yellow. Old items were randomly assigned without replacement to one of nine retention-interval values (1, 2, 5, 10, 25, 40, 50, 70, 100), and all items were randomly allotted to positions within a pre-set schedule of positions in which each quartile contained an equal number of test items of each type. Lag was equated across each test half. There were eight items of each type (TBE, TBX) at each lag and 112 new items, for a total of 400 presentations (including test trials) in the experiment.

Procedure—Subjects viewed instructions informing them that they would be studying words and making memory judgments. Each test trial was indicated by a "???" underneath the word, and subjects were told to endorse the item if and only if they remember having seen it before in the color that it was currently presented in. Each study item was presented for 1500 msec with a 1500 msec inter-stimulus interval. Test items remained on the screen until a response was made. After the experiment, subjects were debriefed and thanked for their participation.

Results

The results are presented in Figure 6, in which hit rates (HR) and false-alarm rates to TBX (FAR-TBX) and to new (FAR-new) items are shown. Here it is evident that that FAR-TBX function initially rises during early forgetting (on the right side of the graph) and then later drops, in parallel with the HR function, during later forgetting (on the left side of the graph). This result confirms the prediction of DRYAD that the representation that underlies accurate exclusion (namely, memory for color) is more dramatically affected by a global decrease in memory than is memory for the item itself.

To test for the reliability of the nonmonotonic trend visually apparent in the FAR-TBX data, the data for each subject were collapsed into retention-interval tertiles (1, 2, 5; 10, 25, 40; 50, 70, 100) and the number of subjects with the maximum value in each tertile was counted. For the FAR-TBX data, 70% of subjects exhibited a maximum in the middle tertile (21% and 28% were in the first and final tertiles, respectively; these probabilities sum to a value greater than 100% because of tied scores). This value is greater than would be expected based on a random distribution of maxima across the tertiles, as estimated by the multinomial distribution ($p < .001$). The values for the HR data were 98%, 4%, and 4%for the three tertiles, which also rejected the null hypothesis of equal distribution ($p < .001$).

Discussion

The data shown in Figure 6 clearly reveal the presence of a nonmonotonic relationship between memory strength (test interval) and endorsement probability. This confirms the central prediction of DRYAD, and thus supports the view that disordinal interactions between age and memory strength reflect a nonmonotonic relationship between learning and false-alarm rates to to-be-excluded items, rather than the selective absence of a process that counters the effect of familiarity in the elderly.

Indeed, one way of thinking about these data is to imagine the various outcomes of an experiment in which only one pair of neighboring retention intervals from our set had been chosen. Depending on which two intervals were chosen, FAR might be found to increase or to decrease with increasing memory strength—and this would have occurred in a sample of exclusively younger adults. From that perspective, this experiment has taken the same tack as the previous one: younger adults were shown to exhibit the same effects as older subjects simply by reducing memory fidelity.

This result also helps explain why younger subjects under deadlined conditions (Benjamin, 2001; Benjamin & Bjork, 2000; Benjamin & Craik, 2001; Light et al., 2004, 2006) exhibit qualitatively similar patterns to the elderly. It is not because such manipulations selectively affect recollection. It is because they reduce the effective quality of the memory trace used for the recognition or exclusion decision.

One important limiting condition is worth noting. At the present time, unpublished experiments from our laboratory and one published paper (Light, Patterson, Chung, & Healy, 2004) have failed to find nonmonotonic functions relating memory strength to false alarms to rearranged pairs of words in associative recognition. Thus, it remains to be seen whether DRYAD will provide a satisfactory account of associative recognition or whether a density explanation fails to generalize to such paradigms.

General Discussion

The experiments reported here follow directly from the theoretical perspective in DRYAD, in which the effect of aging is presumed to be global, not specific, and in which nonfocal information, typically including sources and contexts, is represented more sparsely than focal information. According to that view, the age-related impairment should follow the distribution of attention: older adults should exhibit greater deficits, compared to younger subjects, for material that is less well attended. Experiment 1 evaluated this prediction by directing attention to different aspects of stimuli and showing that older adults exhibited less deficit for the focal than the nonfocal aspects of the stimulus.

In Experiment 2, we evaluated the hypothesis that any means of reducing memory fidelity—rather than specifically age—should elicit the interaction revealing selectively poor memory for context. We took great care to ensure compatible measures of item and context memory in that experiment, and showed that even a very basic variable (study time) affected memory for context more dramatically than memory for items. This result supports the view that whatever specific aspects of aging elicit poorer memory can be profitably analogized with many other variables that affect memory, including study time.

Finally, in Experiment 3, the model's ability to account for disordinal interactions was tested. In the model, a nonmonotonic relationship between memory strength and endorsement probability for to-be-excluded items arises because of the different courses for recovery of item and context information. This claim was qualitatively confirmed by demonstrating a nonmonotonic response function in human data with a simple manipulation of retention interval.

Theoretical perspectives on the age-related deficit in context memory

In many perspectives on aging and memory, the effects of aging are characterized as a *selective* deficit. That is, aging is thought to affect one putative process or system but not another. Views that postulate deficits in source monitoring (Johnson, Hashtroudi, & Lindsay, 1993), associative binding (Naveh-Benjamin, 2000), and recollection (Jacoby, 1999) all characterize the age-related deficit in memory for context in this manner.

In contrast, DRYAD attributes the age-related deficit in context memory to a global reduction in memory fidelity with age. Aspects of stimuli are encoded to the degree that they are attended to. One major advantage of this view is that it relieves theorists of the burden of determining how to ecologically define items and contexts, and directs them instead towards the ecological, motivational, and intellectual factors that drive attention. This perspective helps explain, for example, why the use of contexts that are known to be meaningful to older adults eliminates the deficit (May, et al., 2005; Rahhal et al., 2002), as does the use of contexts that are perceptually highly discriminable (Ferguson, Hashtroudi, & Johnson, 1992).

However, it does not provide a satisfactory way of understanding the circumstances under which deficits obtain for characteristics of stimuli to which attention is specifically directed. Though we argue here that paradigms involving source memory and context memory do not fall into this category—because sources and contexts are typically homogeneous, uninteresting, and irrelevant—other results do. For example, older adults show an inability to remember word-font associations even when memory for the words and fonts *individually* is roughly equivalent between younger and older adults. Such results are at the heart of the claim that an inability to construct and retain associations is central to aging (Chalfonte & Johnson, 1996; Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). The assumptions of DRYAD appear to be inadequate to deal with such a result, and the claim of a associative deficit appears to be incompatible with effects of attention (in the current experiments) and meaningfulness (e.g., Rahhal et al, 2002) that are emphasized here. Clearly more work is needed to relate contextual memory paradigms to associative memory paradigms.

The difference between these perspectives has implications beyond the theoretical ones reviewed here, as discussed at greater length by Benjamin (2010). It also poses a starting point for how we think about rehabilitation regimens or providing environmental support for the elderly. For example, Jennings and Jacoby (2003; Jennings, Webster Kleykamp, & Dagenbach, 2005) argued that training older adults to use recollection should ameliorate the deficit, and provided evidence that it did. From the perspective of DRYAD, however, those results reveal not the training of recollection but rather the training of attention: when older adults attend more to contextual aspects of a stimulus, the deficit will be less. This is, of course, the lesson of the first experiment presented here.

According to DRYAD, age-related deficits in recovering context will be reduced to the degree that contexts elicit greater attention. In addition, since the medium is typically less important than the message, and older adults have some sort of deficit in memory (in *all* theoretical perspectives), it is not clear that retraining older adults to always pay more attention to context is a good idea, since it will likely involve a trade-off with memory for focal information (cf. Light, Berger, & Bardales, 1975).

Thinking about the context memory deficit in older adults as a global deficit also has the advantage of reconciling results from the experimental literature with results from the individual-differences literature. That literature has spoken with a unanimous voice in claiming a single factor mediating age-related deficits on memory tasks. Yet the experimental literature typically characterizes empirical interactions between younger and older adults as revealing of selective problems—failures of one system or process offset by intact functioning of another. As discussed by Benjamin (2010), global deficits are not inconsistent with empirical interactions. Rather, those interactions provide clues as to the ways in which the global deficit acts on information of varying density. The findings in this paper—all of which suggest the viability of the claim of a global memory deficit in older adults—can help us better understand the ways in which global changes in memory fidelity can yield empirical interactions.

Acknowledgments

The authors thank Fergus Craik, Jason Finley, Larry Jacoby, Leah Light, and the members of the Human Memory and Cognition Laboratory at the University of Illinois for their useful comments and feedback on this work. This work was funded by grant NIA 026263 from the National Institutes of Health.

References

Benjamin AS. On the dual effects of repetition on false recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2001; 27:941–947.

Benjamin AS. Representational explanations of “process” dissociations in recognition: The DRYAD theory of aging and memory judgments. Psychological Review. 2010; 177:1055–1079. [PubMed: 20822289]

Benjamin AS, Craik FIM. Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. Memory & Cognition. 2001; 29:691–697.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Benjamin AS, Diaz ML, Wee S. Signal detection with criterion noise: Applications to recognition memory. Psychological Review. 2009; 116:84–114. [PubMed: 19159149]

Castel AD, Craik FIM. The effects of aging and divided attention on memory for item and associative information. Psychology & Aging. 2003; 18:873–885. [PubMed: 14692872]

Chalfonte BL, Johnson MK. Feature memory and binding in young and older adults. Memory & Cognition. 1996; 24:403–416.

Craik, FIM.; Byrd, M. Aging and cognitive deficits: Therole of attentional resources. In: Craik, FIM.; Trehub, SE., editors. Aging and cognitive processes. New York: Plenum; 1982. p. 191-211.

Denney NW, Miller BV, Dew JR, Levav AL. An adult developmental study of contextual memory. Journal of Gerontology: Psychological Sciences. 1991; 46:44–50.

Ferguson SA, Hashtroudi S, Johnson MK. Age differences in using source-relevant cues. Psychology & Aging. 1992; 7:443–452. [PubMed: 1388866]

Green, DM.; Swets, JA. Signal Detection Theory and Psychophysics. New York: Wiley; 1966.

Hacker MJ, Ratcliff R. A revised table of d′ for M-alternative forced choice. Perception and Psychophysics. 1979; 26:168–170.

Hashtroudi S, Johnson MK, Chrosniak LD. Aging and source monitoring. Psychology and Aging. 1989; 4:106–112. [PubMed: 2803603]

Jacoby LL. A process dissociation framework: Separating automatic from intentional uses of memory. Journal of Memory and Language. 1991; 30:513–541.

Jacoby LL. Ironic effects of repetition: Measuring age-related differences in memory. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1999; 25:3–22.

Jennings JM, Jacoby LL. Improving memory in older adults: Training recollection. Neuropsychological Rehabilitation. 2003; 13:417–440.

Jennings JM, Webster LM, Kleykamp BA, Dagenbach D. Recollection training and transfer effects in older adults: successful use of a repetition-lag procedure. Aging, Neuropsychology, and Cognition. 2005; 12:278–298.

Johnson MK, Hashtroudi S, Lindsay D. Source monitoring. Psychological Bulletin. 1993; 114:3–28. [PubMed: 8346328]

Jones TC, Jacoby LL. Feature and conjunction errors in recognition memory: Evidence for dual-process theory. Journal of Memory & Language. 2001; 45:82–102.

Light LL, Berger DE, Bardales M. Trade-off between memory for verbal items and their visual attributes. Journal of Experimental Psychology: Human Learning and Memory. 1975; 1:188–193. [PubMed: 1141830]

Light LL, Chung C, Pendergrass R, Van Ocker JC. Effects of repetition and response deadline on item recognition in young and old adults. Memory & Cognition. 2006; 34:335–343.

Light LL, Patterson MM, Chung CC, Healy MR. Effects of repetition and response deadline on associative recognition in young and older adults. Memory & Cognition. 2004; 32:1182–1193.

Macmillan, NA.; Creelman, CD. Detection Theory: A User's Guide. 2. Mahwah, N.J.: Lawrence Erlbaum Associates; 2005.

Malmberg KJ, Shiffrin RM. The "one-shot" hypothesis for context storage. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2005; 31:322–336.

Matzen LE, Taylor EG, Benjamin AS. Contributions of familiarity and recollection rejection to recognition: Evidence from the timecourse of false recognition for semantic and conjunction lures. Memory. in press.

May CP, Rahhal T, Berry EM, Leighton EA. Aging, source memory, and emotion. Psychology and Aging. 2005; 20:571–578. [PubMed: 16420132]

Naveh-Benjamin M. Adult age differences in memory performance: Tests of an associative deficit hypothesis. Journal of Experimental Psychology: Learning, Memory, & Cognition. 2000; 26:1170–1187.

Naveh-Benjamin M, Craik FIM, Guez J, Kreuger S. Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. Journal of Experimental Psychology: Learning, Memory and Cognition. 2005; 32:520–537.

Salthouse TA. The processing-speed theory of adult age differences in cognition. Psychological Review. 1996; 103:403–428. [PubMed: 8759042]

Shepard RN, Teghtsoonian M. Retention of information under conditions approaching a steady state. Journal of Experimental Psychology. 1962; 62:302–309. [PubMed: 13911664]

Spencer WD, Raz N. Differential effects of aging on memory for content and context. Psychology and Aging. 1995; 10:527–539. [PubMed: 8749580]

Swets JA. Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. Psychological Bulletin. 1986; 99:181–198. [PubMed: 3515382]

Watson JM, McDermott KB, Balota DA. Attempting to avoid false memories in the Deese/ Roediger-McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. Memory & Cognition. 2004; 32:135–141.

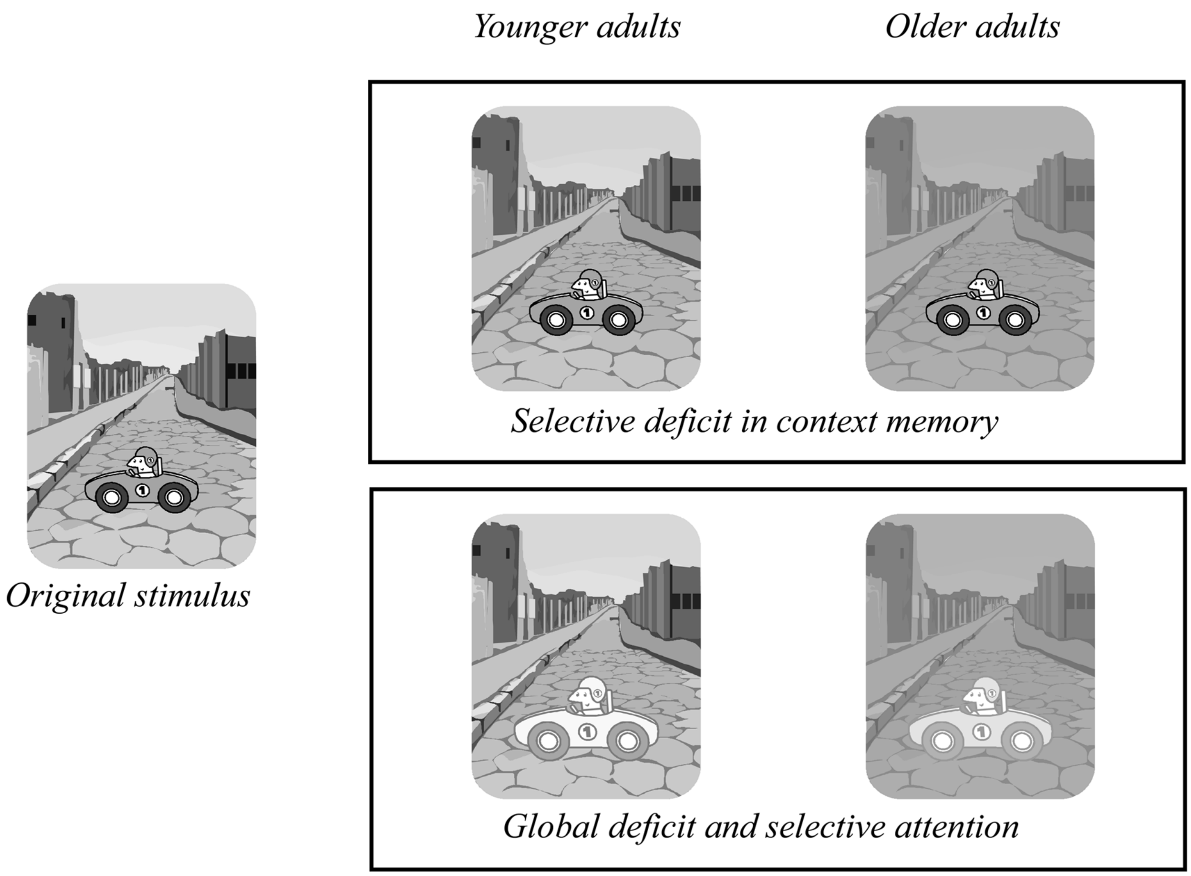

Figure 1.
Different theoretical views of the age-related deficit in context memory. Under the selective-deficit view (top portion of figure), older adults encode, store, or retrieve the contextual aspects of the stimulus less faithfully than the item (the car). Under the global-deficit view (DRYAD), the entire stimulus is maintained less faithfully in the elderly, but attention during encoding moderates the resulting deficit. Thus, greater attention to the car (indicated by its greater size and brightness) and lesser attention to the background leads to a disproportionate deficit in memory for the background.

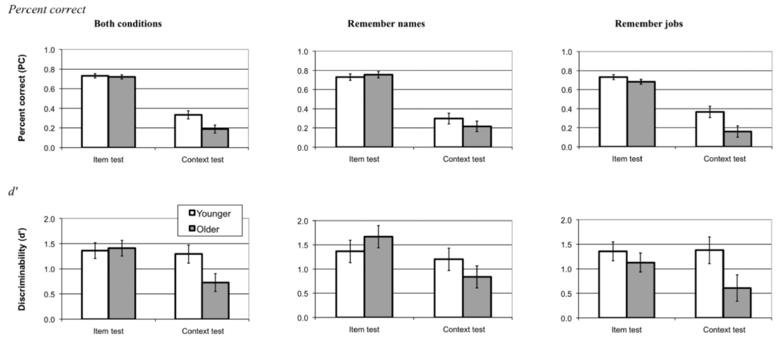

Figure 2.
Memory performance by younger and older adults on tests of item and context memory in the chair-room experiment (Experiment 1). Top graphs plot performance as percent correct and bottom graphs plot performance as d'. Left graphs shows performance from both conditions combined; middle graphs shows the condition in which memory for chairs was emphasized; right graphs shows the condition in which memory for rooms was emphasized. Error bars represent standard error pooled over age groups for each type of test.

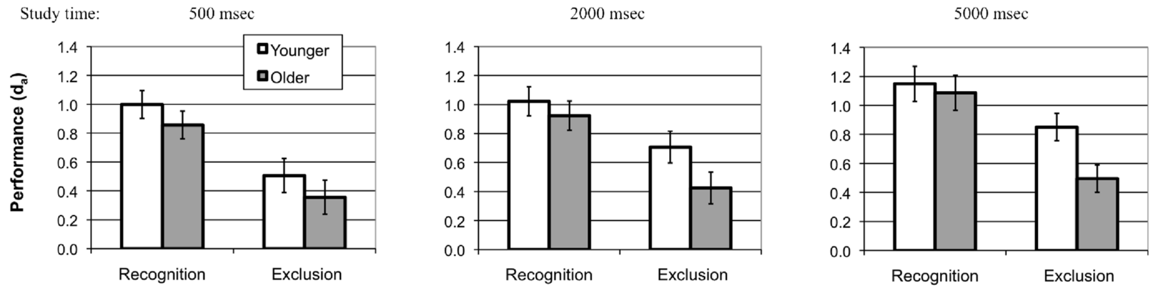

Figure 3.
The effects of age group and memory test type on discriminability (d_a) as a function of study time (Experiment 2). Error bars represent standard error pooled over age groups for each type of test.

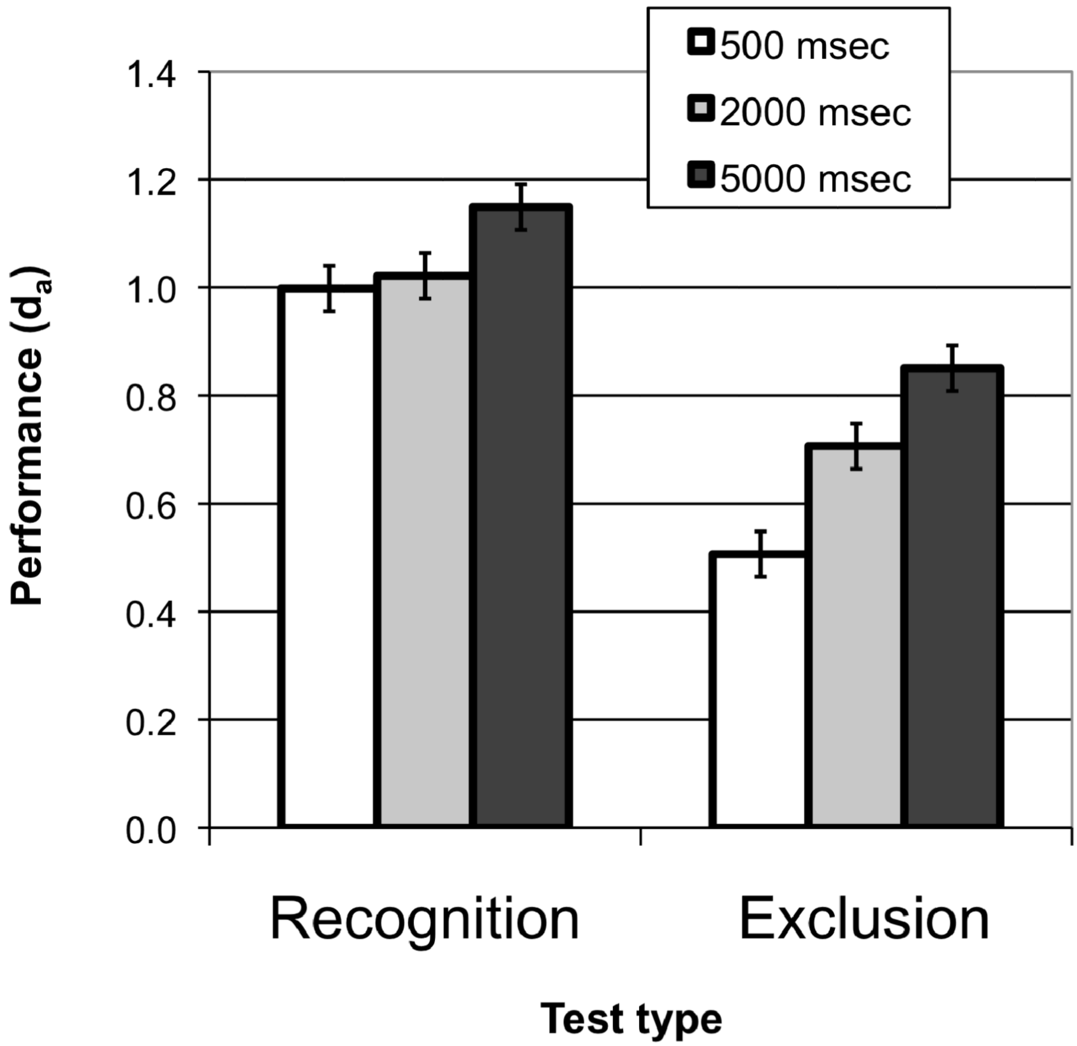

Figure 4.
Discriminability (d_a) for recognition and exclusion as a function of study time for younger subjects (Experiment 2). Error bars represent standard error pooled over all cells in the 2×3 within-subjects design.

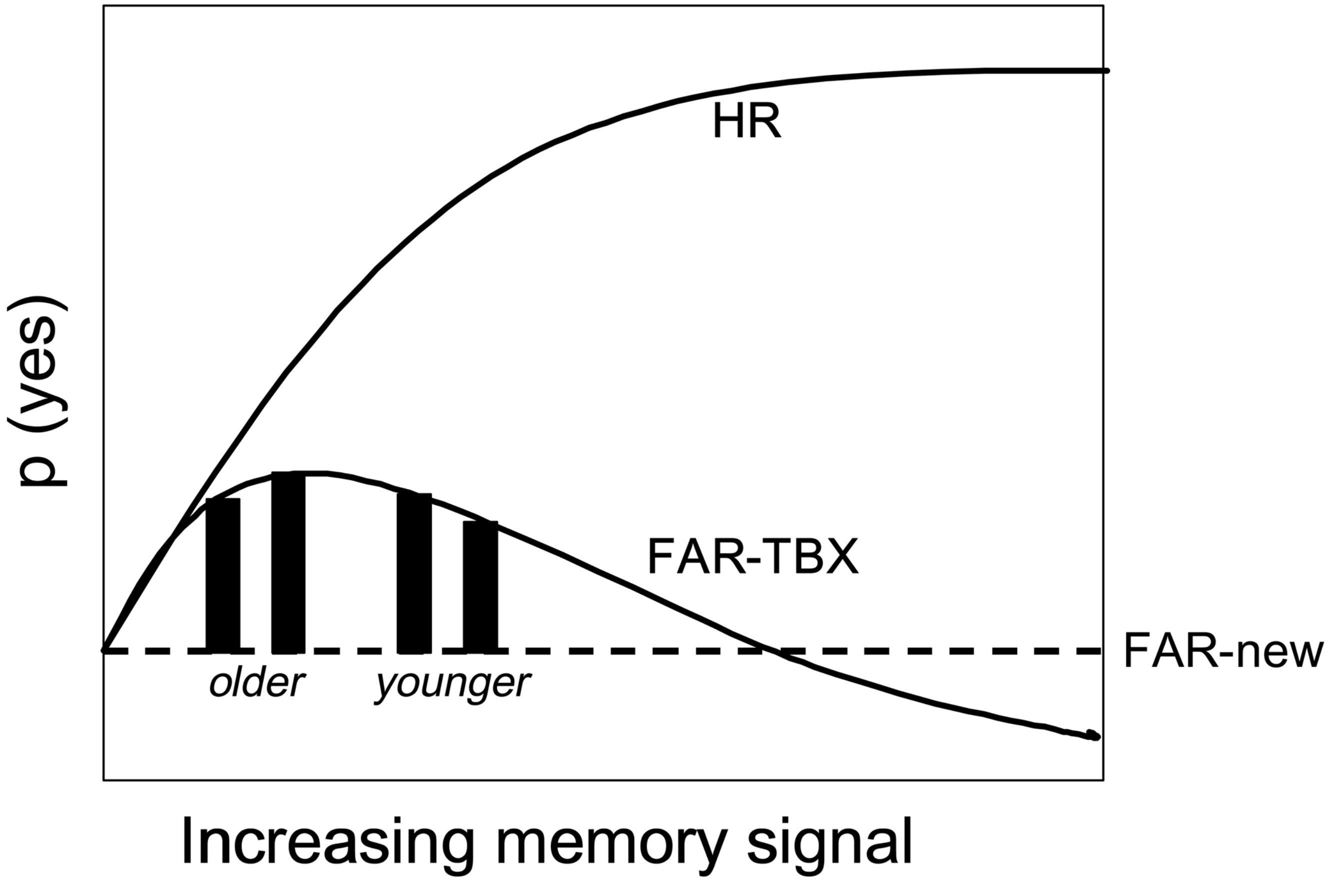

Figure 5.
A plot of the theoretical relationship between memory strength and endorsement probability for to-be-endorsed old items (HR), to-be-excluded old items (FAR-TBX), and new items (FAR-new). FAR-TBX can either rise or fall with additional memory strength depending on the position within the function.

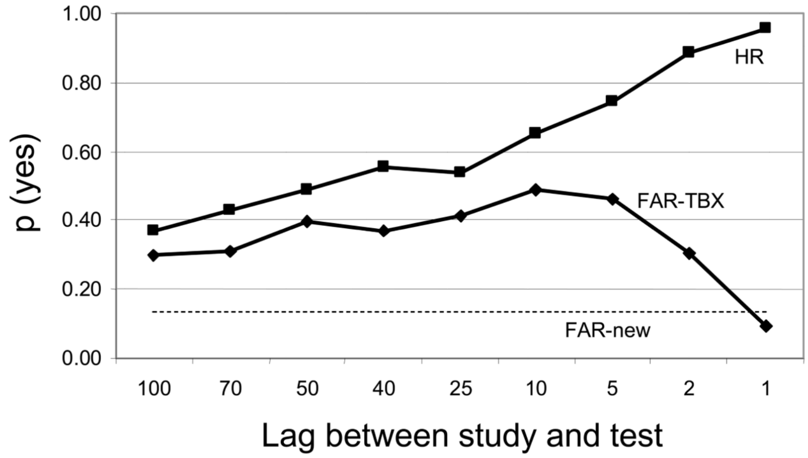

Figure 6.
Empirical confirmation of the theoretical relationship shown in Figure 5. Hit rates (HR), false-alarm rates to to-be-excluded items (FAR-TBX), and false-alarm rates to new items (FAR-new) as a function of the study-test retention interval in the continuous exclusion task (Experiment 3).