Published in final edited form as:

J Exp Psychol Learn Mem Cogn. 2011 July ; 37(4): 861–873. doi:10.1037/a0022798.

Variations in recollection: The effects of complexity on source recognition

Colleen M. Parks¹, Linda J. Murray², Kane Elfman², and Andrew P. Yonelinas²

¹ University of Nevada, Las Vegas

² University of California, Davis

Abstract

Whether recollection is a threshold or signal detection process is highly controversial. U-shaped zROCs observed in tests thought to rely heavily on recollection, such as source memory tests, have provided evidence in favor of the threshold assumption, but zROCs are not always as U-shaped as threshold theory predicts. Source zROCs have been shown to become more linear when the contribution of familiarity to source discriminations is increased, and this may account for the existing results. However, another way in which source zROCs may become more linear is if recollection can become more graded under certain conditions. We tested the ‘graded recollection’ account in the current study. We found that increasing stimulus complexity (i.e., changing from single words to sentences), or increasing source complexity (i.e., changing the sources from audio to videos of speakers), resulted in flatter source zROCs. In addition, conditions expected to reduce recollection (i.e., divided attention and amnesia) had comparable effects on source memory in simple and complex conditions, suggesting that differences between simple and complex conditions were due to differences in the nature of recollection, rather than differences in the utility of familiarity. The results suggest that under conditions of high complexity recollection can appear more graded and it can produce curved ROCs. The results have implications for measurement models and for current theories of recognition memory.

Recognition memory is largely agreed to rely on two processes, recollection and familiarity (e.g., Yonelinas, 2002; Diana, Reder, Arndt, & Park, 2006; Wixted, 2007). Dual process theory defines recollection as conscious memory for contextual or episodic details of prior events, whereas familiarity is defined as recognizing something as old without consciously recalling episodic details (Mandler, 1981; Jacoby, 1991; Tulving, 1985; Yonelinas, 1994). Despite the emerging consensus on the component processes of recognition memory, there is still widespread disagreement about the nature of recollection and whether it is best described as a threshold process, whereby recollective strength can fall below a memory threshold, or as a signal detection process whereby memory strength varies in a completely continuous manner. Answering this question is important for testing competing theories of recognition memory and because it affects how one measures recollection. For instance, if recollection is a threshold process, then explaining recollection deficits in aging or amnesia is a matter of explaining why recollection is less likely to occur successfully. On the other hand, if it is a continuous process as described by signal detection theory, then it might be

Address Correspondence to: Colleen Parks, Psychology Department, University of Nevada, 4505 S. Maryland Parkway, Box 455030, Las Vegas, NV 89154-5030, colleen.parks@unlv.edu.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/xlm

incorrect to conceptualize deficits as less frequent recollection; rather, it may be that the frequency is stable, but that the strength of recollection has decreased or the variability of its strength greatly increased. As it turns out, deciding between these different ways of describing recollection has not been easy because there is evidence supporting both threshold and signal detection based accounts of recollection.

The most common way of assessing the assumptions of threshold and signal detection theories of recollection has been to measure source memory and examine Receiver Operating Characteristics (ROCs). Source memory refers to the ability to discern the original source of an item, and it is typically expected to rely heavily (though not exclusively) on recollection (Johnson, Hashtroudi, & Lindsay, 1993; Parks & Yonelinas, 2007; Yonelinas 2002). In the lab, subjects will often hear two voices reading words, see the words on different sides of the computer screen, or read words in different colors. Later they are presented with the studied items and asked to remember who said it, where it was, or what color it had been and to rate their confidence in their source response. ROCs are then examined by plotting the hit rate against the false alarm rate across different levels of confidence (see MacMillan & Creelman, 2005 for details). The z-scores of each hit and false alarm rate are also typically plotted to produce zROCs. Plotting the ROCs in z-space is useful because signal detection theory predicts linear zROCs, whereas threshold theory predicts U-shaped zROC (e.g., T. Parks, 1966; see Macmillan & Creelman, 2005)¹. The zROCs observed in source memory tests are almost always U-shaped, in contrast to those seen in item recognition tests which are often very close to linear (see Yonelinas & Parks, 2007 for a review).

Although signal detection theories accounted for the linear zROCs typically found in item recognition, they could not account for the U-shaped zROCs found in source memory tests. Although pure threshold models could account for U-shaped zROCs, not all source zROCs are as U-shaped as threshold theory predicts. One interpretation of this pattern of findings is that although source memory often relies heavily on recollection, it reflects the contribution of both recollection and familiarity. In line with this interpretation, the dual process signal detection (DPSD) model predicts different degrees of zROC linearity or curvilinearity depending on the relative contribution of recollection and familiarity (e.g., Yonelinas, 1999). According to this approach, recollection is a threshold process and familiarity is a signal detection process. This combination of processes allows the model to predict zROC shapes that are different from those predicted either by pure threshold or pure signal detection models alone. Specifically, tasks that require greater proportions of recollection for successful performance are expected to result in more U-shaped zROCs. Because many source memory tests force subjects to rely primarily on recollection to make accurate source decisions, the model predicts U-shaped zROCs. However, when source tests allow familiarity to serve as a basis for accurate source judgments, the DPSD model predicts zROCs that are closer in shape to those found in item recognition (i.e., more linear).

In line with these predictions, the degree to which source memory zROCs are U-shaped can be systematically varied. For instance, if the items from two sources differ in overall strength then familiarity can support source memory discriminations. For example, if source-A items are studied five days before source-B items then successful source decisions can be supported by familiarity and the resulting zROCs become flatter (e.g., Yonelinas, 1999). Similarly, if items are encoded such that they are integrated or unitized with their sources familiarity can also support source discriminations. For example, if subjects encode

¹Importantly, threshold theories can also predict curvilinear ROCs if one adds additional free parameters for each confidence point, but the standard threshold models of recollection assume that when recollection occurs, it supports high confidence responses and therefore results in a linear ROC (for a discussion of such models see Malmberg 2002; Yonelinas & Parks, 2007)

the item ‘elephant’ as coming from the ‘red’ source by imagining a sunburned elephant, then a red elephant should be more familiar than a green elephant at test. Encoding conditions like this that promote unitization have been found to promote familiarity-based source discriminations and they lead zROCs to become flatter (e.g., Diana, Ranganath, & Yonelinas, 2008; Diana, Yonelinas, & Ranganath, 2009; for related results see Staresina & Davachi, 2006).

However, there are also cases in which no particular manipulation was used to foster the use of familiarity as a basis for source memory decisions and the resulting zROCs exhibited little if any evidence of a U-shape (e.g., Glanzer, Hilford, & Kim, 2004; Qin et al., 2001; Slotnick et al., 2000). The most dramatic example of this is a study reported by Qin et al. (2001), who found almost perfectly linear zROCs in a test of source memory for spoken statements. Findings of flat zROCs in source memory, like those found in Qin et al. (2001) have suggested to some that the DPSD model and others like it are inconsistent with the data, and as indicating that recollection is a continuous process that is more in keeping with a pure signal detection model (e.g., Wixted, 2007; Qin et al., 2001; Slotnick & Dodson, 2005). Qin et al. argued, for example, that recollection was the driving force behind the source judgments, and thus also was the process that produced the linear zROC. If true, such evidence would provide a challenge to the assumption that recollection is a threshold process.

Qin et al (2001) argued that the reason recollection acted in a more continuous manner in their study than in previous studies was because their materials were much more complex than those used previously. In that study, subjects watched videotapes of speakers making statements (schema-based statements in one study, facts and feelings in another). Both the studied items (i.e., sentences) and the sources (i.e., auditory and visual information about the speaker) in these experiments contained more information than in typical source memory experiments. Thus, even if the recollection of each aspect of the study item or source behaves in line with threshold theory, taken together they may begin to act in a more signal-detection like manner. However, those experiments differed from previous studies in various other ways as well (e.g., the analysis was based on aggregate ROCs rather than on individual subject ROCs because there were too few trials collected from each subject). Thus, it is not possible to determine with certainty whether stimulus complexity per se was responsible for the flattened zROCs. In addition, even if complexity were shown to lead to a flattening of the source zROCs, it would not be clear whether complexity led recollection to become more continuous in nature, or if complexity led familiarity to be more useful in making the source discriminations. That is, highly complex stimuli may facilitate the unitization of the item and source information, and thus the flattening of the zROC could be due to familiarity, rather than recollection.

We examined these issues in a series of experiments geared toward determining whether manipulations of complexity would produce differences like those found between Qin et al. (2001) and Yonelinas (1999). In experiments 1a and 1b we manipulated complexity of the stimulus (words versus sentences) and the complexity of the source (audio versus video). The simplest condition, in which subjects listened to male and female speakers reading single words, mimics many source experiments that have produced fairly linear source ROCs (U-shaped zROCs). The most complex condition, in which subjects watched videos of a man and a woman making statements, was designed to match the stimulus properties of the Qin et al. study. If the complexity of the item or the source is a critical factor, then we should find increasing ROC curvilinearity with increasing complexity. Experiment 1 demonstrated that both stimulus complexity and source complexity can lead to a flattening of the source zROCs. Experiment 2 was designed to determine if the flattening of the

zROCs produced by complexity arose because complexity led recollection to become more continuous, or because it increased the reliance on familiarity-based source memory.

Experiment 1

Experiments 1a and 1b were conducted to test the effects of stimulus and source complexity on the shape of source memory ROCs. Stimulus complexity (words vs. sentences) and source complexity (audio vs. video) were crossed, resulting in four experimental conditions in Experiment 1a (audio word, audio sentence, video word, and video sentence). In Experiment 1b only the two extreme conditions (audio-word [simplest] and video-sentence [most complex]) were included, but ROCs were examined under conditions in which overall performance was matched to determine whether differences in ROC shape could be attributed to differences in performance.

Method

Experiment 1a

Participants—Ninety-six students (average age 19 years, 72 women, 24 men) from the University of California, Davis participated in Experiment 1a for course credit. Twenty-four participants were randomly assigned to each of the four conditions: audio-word, audio-sentence, video-word and video-sentence.

Materials and Procedure—The stimulus pool consisted of 230 words and 230 sentences. The words were all nouns, 4 to 8 letters in length, with average Kucera-Francis frequency of 51.3. The sentences varied in length and consisted of general trivia such as *There are more Barbie dolls in Italy than there are Canadians in Canada.* and *The hippopotamus has a stomach 10 feet long.* Trivia were collected from Symons (2006), Harvey (1987), and Bacon (1979). Both sets of stimuli were divided into lists of 115 items each to create two counterbalance conditions. Six additional items (of both types) were used as buffers, 3 for recency and three for primacy.

A man and a woman were filmed reading the sentences and words. The digital video was then edited to create both the audio and the video study lists for each of the four experimental conditions (audio-word, audio-sentence, video-word, and video-sentence). Presentation of stimuli by the two speakers was counterbalanced across subjects. One order of stimuli was presented to all subjects at study; stimuli were pseudo-randomly mixed such that the man or the woman read no more than three items in a row. Words presented for test were newly randomized for each subject.

Procedure—In the study phase, participants were instructed to attend to the speaker and what he or she said in preparation for a later memory test on which speaker said which word or sentence. The audio and video word lists lasted approximately 10 minutes, whereas the sentence lists (audio-sentence and video-sentence) lasted approximately 25 minutes. Duration of the stimuli varied due to different speaking rates and stimulus lengths but there was a standard 1-s ISI between all stimuli. All study items were tested and were presented as text on a computer screen such that there was no perceptual match between study and test for any of the conditions. Items (words or sentences) were presented one at a time on a computer screen and subjects indicated whether they remembered the man or the woman presenting the item using a six-point confidence scale; 1 represented high confidence that the item had been presented by the man and 6 represented high confidence that the item had been presented by the woman. Subjects were instructed to use the entire range of response confidence and completed the test at their own pace.

Experiment 1b

Participants—Thirty-two students (average age 19.5, 23 women, 9 men) from the University of California, Davis participated for course credit.

Procedure—The materials and procedures were the same as those of Experiment 1 with the exception that only the extreme conditions from Experiment 1a were tested (audio-word and video-sentence), the manipulation was within-subjects, and in an attempt to equate performance across the conditions, audio-word stimuli were each presented twice (video sentence stimuli were presented only once as in Experiment 1a). Repetitions of audio-word stimuli were pseudo-randomly ordered such that no word was repeated immediately. Participants completed the audio-word condition first followed by the video-sentence condition; this order ensured that subjects could not visualize the man and woman in the videos while encoding in the audio-word condition.

Results

Experiment 1a

An analysis of variance of overall recognition performance (indexed by d_a , a measure of discrimination analogous to d' but appropriate for both symmetrical and asymmetrical ROCs see Table 1) showed both a main effect of stimulus complexity ($F_{(1,95)} = 6.3$, $p < 0.02$), and a main effect of source complexity ($F_{(1,95)} = 12.16$, $p < 0.002$), but no significant interaction ($F_{(1,95)} = 1.5$, $p = .229$), indicating that increases in stimulus complexity and source complexity led to increases in overall source memory performance.

The average ROCs and zROCs are presented in Figure 1. The figure shows that increasing stimulus complexity (moving from the circles to the triangles) led the ROCs to become more curved (and conversely led the zROCs to become flatter). In addition, increasing source complexity (moving from the two lower to the two upper curves) led the ROCs to become more curved (and led the zROCs to become flatter). To quantify these effects we examined the shape of the ROCs (see Figure 1) in three complementary ways. First, the DPSD model was fit to each subjects' ROCs using maximum likelihood estimation to obtain estimates of recollection and familiarity. Because recollection is modeled as a threshold process and familiarity is modeled as a signal detection process these parameters can also be interpreted as measurements of ROC shape, with familiarity providing an estimate of the curvilinearity and recollection providing an estimate of the linearity of the ROC. The parameter estimates for each condition are shown in Figure 2. The recollection estimates did not differ across conditions ($F_{(3,95)} = 1.77$, $p = 0.16$). The familiarity estimates, on the other hand, showed both a main effect of stimulus complexity ($F_{(1,95)} = 6.62$, $p < 0.02$) and a main effect of source complexity ($F_{(1,95)} = 6.95$, $p < 0.02$), such that the familiarity (curvilinearity) of the ROCs was greater when complexity increased. The interaction was not significant ($F < 1$). Thus, the analysis indicated that both stimulus and source complexity led the source ROCs to become significantly more curvilinear.

In our second analysis of ROC shape, we compared the fit of the traditional unequal variance signal detection (UVSD) model to the fit of the same model with an added threshold parameter ("UVSD+R") for the two extreme conditions (audio-word and video-sentence). The models were fit using maximum likelihood estimation and goodness of fit was measured using the log-likelihood statistic G^2 . It was expected that the addition of the extra parameter in the UVSD + R model would result in a better fit (and thus a lower G^2) for both conditions. If, however, the video-sentence ROC is more curvilinear than the audio-word ROC, then the audio-word condition should obtain a greater benefit from the addition of the threshold component than the video-sentence condition. This was the pattern that was

found: the decrease in G^2 values with the addition of the threshold parameter to the model was greater in the audio word condition than in the video sentence condition (audio word: 9.33 and 3.97, video sentence: 5.35 and 3.30, for UVSD and UVSD+R respectively, Mann-Whitney U test, $z = -2.18$, $p < .03$)).

In the third analysis of shape, zROCs (see Figure 1) for each subject were submitted to quadratic regression analyses to obtain an estimate of the quadratic coefficient, an estimate of the (U-shaped) curvature present in the zROC. Because most source zROCs are U-shaped, we expected the quadratics to be greater than zero, and in addition, that the complexity manipulation would affect the size of the coefficient, with simpler conditions having larger quadratics than complex conditions. The quadratic coefficients were greater than zero in all conditions (audio word $t(21) = 6.08$; video word $t(21) = 4.28$; audio sentence $t(21) = 2.14$; video sentence $t(22) = 6.21$; all $ps < .05$)². An analysis of variance indicated that there was an interaction between stimulus complexity and source complexity ($F(1, 85) = 13.23$, $MSe = .05$, $p < .001$). Consistent with the results of the prior sets of analysis, the quadratic in the audio-sentence condition was smaller than in the audio-word condition ($t(42) = 3.08$, $p < .005$), indicating that stimulus complexity led to flatter zROCs. Moreover, the quadratic in the video-word condition was smaller than the audio-word condition, although this difference was not significant ($t(42) = .86$, $p = .39$). However, the video sentence condition produced the largest quadratic component of any condition, suggesting that it produced the most U-shaped zROCs. This latter finding is inconsistent with the two previous sets of analyses, which indicated that the video-sentence condition produced the most curved ROCs (and thus the flattest zROCs). One potential explanation of this anomalous finding is that differences in overall performance may have influenced the quadratic analysis. That is, a quadratic value of .2 when performance is very low is not necessarily equivalent to a quadratic of .2 when performance is much higher. This possibility will be directly addressed in the next experiment.

Overall, the dual process and signal detection analyses converged in showing that increasing either stimulus complexity or source complexity led the ROCs to become more curved (and thus led to a flattening of the zROCs). The linearity analysis of the zROCs supported the same conclusions, with the exception of the video sentence condition, which was expected to have the smallest quadratic coefficients but in fact produced the highest.

Experiment 1b

The purpose of Experiment 1b was to examine the effects of complexity on ROC shape when overall performance was matched across conditions. In order to increase the performance in the audio word condition items were presented two times rather than just once. The average ROCs and zROCs overlap for the two conditions overlap (see Figure 1), and overall memory performance, measured by d_a , did not differ between the two conditions (see Table 1; $t_{(31)} = 1.04$, $p = 0.31$); thus performance was well matched between the conditions in this experiment.

The DPSD model was fit to the ROCs in order to obtain familiarity estimates as measures of curvilinearity in probability space. The parameter estimates for each condition are shown in Figure 2. Recall estimates did not differ between conditions ($t(31) = 1.68$, $p = .10$) but familiarity estimates were higher in the video-sentence than in the audio-word condition ($t(31) = 2.69$, $p < .02$), indicating that the ROCs were more curved in the more complex condition. Fits of the traditional UVSD and the UVSD+R models to the data confirmed this

²Quadratic coefficients were first submitted to outlier analyses; coefficients that were outside the bounds defined by the mean plus or minus two standard deviations were not included in further inferential analyses.

pattern. Adding a threshold parameter to the unequal variance model led to a decrease in G^2 in the audio-word condition (from 9.0 to 3.8) that was greater than the decrease in G^2 in the video-sentence condition (from 6.4 to 3.9) (Mann-Whitney U test, $z = -2.27$, $p < .03$).

Finally, the zROCs were submitted to quadratic regression analyses to obtain estimates of quadratic curvature in z-space. Quadratic coefficients in both conditions were greater than zero ($t(28) = 8.8$; video sentence $t(28) = 3.65$; $ps < .002$), but they were greater in the audio word condition than in the video sentence condition ($t(28) = 2.85$, $p < .009$). Thus, when performance is matched in the two extreme conditions, all three methods of ROC shape analysis concur: ROCs were more curvilinear (flatter in z-space) in the complex video sentence condition than in the simple audio word condition.

The critical pattern of results from Experiment 1a was replicated: the average ROC for complex condition (video-sentence) was more curvilinear than the ROC for the simple condition (audio-word). This pattern was found (in reverse) in z-space as well: the zROC for the complex condition was more linear (less U-shaped) than in the simple condition (see Figure 1). Importantly, Experiment 1b results demonstrate that these differences are not attributable to differences in performance level.

Combined, Experiments 1a and 1b indicate that increasing the complexity of the studied item or increasing the complexity of the source information can lead to more curvilinear ROCs and flatter zROCs. Clearly both stimulus and source complexity can affect the shape of ROCs, but the most dramatically curved ROCs were found when both item and source complexity increased together. Increasing complexity generally led to an increase in overall performance in Experiment 1a, particularly when both stimulus and source complexity increased together. Experiment 1b addressed this issue by examining the effects of complexity while controlling for levels of overall performance and showed the same pattern: the simple condition produced relatively linear ROCs whereas the complex condition produced curved ROCs.

Experiment 2

Experiment 1 showed that complex study events can lead source ROCs to become more curved, but they do not indicate why differences in complexity produce differences in ROC shape. It could be that source memory judgments in complex conditions rely more on familiarity if the richer information allows the subjects to encode the item and source in an integrated manner. Previous work has indicated that familiarity can support source and associative recognition discriminations if the item and source information are unitized or integrated during study (e.g., Diana et al., 2008; Diana et al., 2009; Quamme, Yonelinas, Norman, 2007; Yonelinas et al, 1999). Thus, encoding sentences from a video of the speakers may provide a richer study event that more readily allows an integration of the item and source information, thereby increasing the utility of familiarity in source tests.

Another possibility is that the increase in the curvilinearity of the ROC might occur because increasing complexity causes recollection to behave in a more continuous or Gaussian manner (Yonelinas, 1999a; Qin et al., 2001). For example, even if the retrieval of each detail is a threshold process (it may simply fail), the culmination of many such details may give rise to graded recollection for the complex event (e.g., Malmberg, 2002). If so, then recollection may influence ROC shape differently depending on the complexity of the stimuli.

To test these alternative accounts, we examined the effects of divided attention (Experiment 2a) and amnesia (Experiment 2b) on source discriminations for the simple and complex source conditions from Experiment 1b. In Experiment 1a we examined divided attention at

test because previous work has indicated that recollection is particularly sensitive to the disruptive effects of divided attention (e.g., Gruppos, Lindsey, & Kelley, 1997; for a review see Yonelinas, 2002). Thus, if the differences seen in the ROCs in the previous experiment were due to differences in the nature of recollection, but the *contribution* of recollection was the same, then dividing attention should reduce performance equally in the simple and complex conditions. In contrast, if the exaggerated curve seen in the complex condition was due to an increased reliance on familiarity in that condition, then dividing attention should have a smaller effect in the complex condition than in the simple condition.

We applied similar logic in Experiment 2b; we compared the performance of hypoxic patients to that of healthy controls on the simple and complex source tasks. Previous work has indicated that patients suffering mild hypoxic episodes exhibit specific deficits in recollection, (e.g., Yonelinas et al., 2002; for evidence that severe hypoxic patients can exhibit deficits in both processes see Manns, Hopkins, Reek, Kitchener, & Squire, 2003; Yonelinas et al., 2004). If recollection contributes equally to the simple and complex source tasks, then the patients should show similar deficits relative to controls on the two tasks. In contrast, if performance on the complex source task is supported by familiarity, then the patients should show less of a deficit on the complex than the simple source task.

Method

Experiment 2a: Divided Attention

Participants—Sixty-four students (average age 19.3, 47 women, 17 men) from the University of California, Davis participated for course credit. Thirty-two participants were randomly assigned to the Full Attention (FA) condition and thirty-two participants were randomly assigned to the Divided Attention (DA) condition.

Materials and Procedure—Each participant completed the audio-word task and then the video-sentence task from Experiment 1b. In the DA condition a digit-monitoring task was added during the test phase of each task. The digit-monitoring task consisted of an auditory presentation of single digits ranging from 1 to 9 in a random order, with a one second interval between digits. Two different orders were created- one for each test phase. The participants' task was to monitor the series of digits and say "odd" out loud whenever two odd digits were presented consecutively. The experimenter monitored the performance of each participant and hit a desk bell whenever an error was made.

The confidence scale was eliminated from the test phase in both tasks; participants simply pressed a key labeled "man" or a key labeled "woman" for each word or sentence presented. In the DA condition, each response timed out after 5 seconds if the participant did not respond. This was done to minimize task-switching between the memory test and the digit-monitoring task. Participants were instructed to treat both tasks as important and to try to make as few mistakes as possible. Before each test phase, participants were given a short practice session on the digit-monitoring task.

Experiment 2b: Amnesia

Amnesic Patients—Two amnesic patients were recruited for participation and paid for their time. Both patients showed normal intelligence and attention scores but exhibited memory impairments, primarily with verbal materials (see Table 2). Patient 1 was a survivor of sudden cardiac arrest and so was expected to have bilateral damage relatively limited to the hippocampus as a result of cerebral hypoxia (Rempel-Clower, Zola, Squire & Amaral, 1996), however due a defibrillator the patient could not be scanned. Patient 2 suffered a traumatic head injury acquired in a motorcycle accident and was in a coma for one month. A

structural MRI revealed no apparent brain damage, suggesting that his memory impairment was hypoxia-related.

Healthy Controls—Nine healthy control participants were recruited for participation and paid for their time. The average years of education ($M = 16.6$) and the average age ($M = 38.8$) of the controls was similar to that of the amnesics ($M = 16$ and $M = 38.5$, respectively).

Materials and Procedure—The materials and procedures were identical to those used in Experiment 2a full attention condition except that participants were instructed to use a strategy during the study phase that involved thinking about the feminine or masculine aspects of a word or sentence when trying to link stimuli to the female or male speaker. The strategy instruction was added in order to reduce the variance between participants, because debriefing in previous studies suggested that some subjects tried to use encoding strategies but others did not. Finally, participants in Experiment 2b made confidence responses on a six-point scale at test.

Results and Discussion

Experiment 2a: Divided Attention

Figure 3 shows d' averages for the two tasks in both the full and divided attention conditions. Dividing attention reduced source memory performance, and had almost identical effects on the simple and complex tasks (i.e., leading to a reduction in d' of 0.51 and 0.56 in the simple and complex conditions, respectively). The decrease in performance due to divided attention was significant ($F_{(1,62)} = 17.2$, $p < 0.001$), however d' remained above 0 in both tasks (audio-word: $t_{31} = 4.51$, $p < 0.001$; video-sentence: $t_{31} = 6.8$, $p < 0.001$). Results across the simple and complex memory tests did not differ significantly, nor was there a significant interaction between test type and attention condition (both $F_s < 1$).

The similarity in the effects of divided attention on the simple and complex conditions is most consistent with the idea that recollection was supporting source decisions to the same degree in the two conditions. Had the increase in curvature found in the complex condition ROCs been due to familiarity, performance in that condition should have been less vulnerable to the effects of divided attention. Thus, the patterns in Experiment 2 suggest that differences in the complexity or richness of the stimuli may affect the behavior of recollection such that it acts in a probabilistic manner with relatively simple stimuli but becomes more signal-detection like with more complex stimuli.

Experiment 2b: Amnesia

Recognition performance for patients and controls is shown in Figure 4. Overall performance of the two patients was significantly lower than that of the control group in each condition ($F_{(1,8)} = 16.8$, $p < 0.004$), but there were no differences in performance as a function of condition ($F_s < 1$). This pattern of results suggests that the combination of recollection and familiarity that supports source memory in the simple condition is about the same as that which supports source memory in the complex condition.

Because there were only two patients, statistical analyses were limited to the overall recognition scores (above). However, visual inspection of the ROCs and zROCs (see Figure 5) supports the conclusions based on the d_a analysis that the amnesic patients' deficits were comparable in the two conditions, thus suggesting that the tasks were equally dependent on recollection. DPSD model estimates were also consistent with this conclusion, showing no

evidence that the amnesics were able to rely more on familiarity in the video sentence condition than in the audio condition (see Figure 5).

Overall, the patients showed similar impairments in the simple and complex conditions. These data are consistent with the claim that recollection is driving performance in both conditions despite the difference in ROC shapes, thus supporting the findings of Experiment 2a.

General Discussion

The aim of the current study was twofold: to determine whether increasing the complexity of stimuli in a source memory experiment would lead the ROCs to become more curved (and the zROCs to become flatter), and if so, whether those changes in shape were due to recollection, familiarity or both. We varied complexity by mimicking the extreme conditions found in the literature that seem to have been associated with ROCs of different shapes: our simplest condition (audio word) matched the typical source conditions that are found to produce fairly linear ROCs, and our most complex condition (video sentence) matched the richer source conditions found in the Qin et al. (2001) study. The results of Experiment 1a indicated that both stimulus complexity and source complexity increased the curvilinearity of the ROCs (and decreased the U-shape of zROCs). Experiment 1b showed that the effects of complexity on ROC shape remained even when overall performance was matched, thus ruling out the possibility that the increased ROC curve was an artifact produced by high levels of performance. Experiments 2a and 2b examined whether the change in shape was due to changes in recollection or familiarity, by examining the effects of dividing attention and mild hypoxia – two factors known to undermine recollection. Recognition performance in the simple and complex conditions was similarly affected by both of these factors suggesting that complexity influenced recollection rather than familiarity. Together the results of the current experiments demonstrate that increasing complexity leads to an increased ROC curve (i.e., a flattened zROC) and suggests that this influence on shape is most likely due to changes in the nature of recollection, rather than in the extent to which familiarity contributes to performance.

The debate over the shape of source ROCs has been fueled in part by the mix of findings ranging from very linear to very curvilinear ROCs. The current study addressed this debate by examining one factor that may have led to some of the discrepancies in previous studies. The results make it clear that the shape of ROCs in source memory studies is dependent on the complexity of the stimuli. When the source and stimuli were simple we observed a fairly linear ROC that was highly U-shaped in z-space, but increasing the complexity of either the source or the stimulus led to more curvilinear ROCs and flatter zROCs. Thus, in addition to factors such as unitization (Diana et al., 2008; Diana et al. 2009) and item strength differences (Yonelinas, 1999), it is now clear that increasing complexity can lead to curved source ROCs.

Unlike unitization and strength differences though, which have been shown to affect ROC shape by increasing the usefulness of familiarity in making source decisions, complexity appears to influence the shape of source ROCs by changing the nature of recollection. If the differences in ROC curvature found in Experiments 1a and 1b had been due to greater familiarity in the complex condition, then dividing attention at test in Experiment 2a should have affected performance in the simple condition more than in the complex condition. In fact though, dividing attention had the same effect on performance in the two conditions. Thus, the results suggest that contributions of recollection and familiarity were the same regardless of complexity and that the differences in ROC curvilinearity in Experiment 1 were due to differences in the behavior of recollection. In addition, the amnesic patients in

Experiment 2b who were expected to have severe recollection impairments performed similarly in the simple and complex conditions, suggesting that contributions of recollection and familiarity were similar in the two conditions. Together, these results suggest that recollection generally leads to linear ROCs, but that when materials become highly complex, recollection can produce more curved ROCs. One might argue that the failure to find effects of divided attention and amnesia on ROC shape reflected a lack of statistical power. However, both of these manipulations had sizable effects on overall recognition performance, and there was no indication in the results that the complexity manipulation influenced familiarity. Future studies using convergent methods to verify these conclusions will be useful, but the current results certainly favor the graded recollection account of complexity over a familiarity account of the effect.

These results present a challenge for models such as the DPSD model (Yonelinas, 1994) which assume that recollection is a threshold process that predicts linear ROCs. Importantly, however, even though complexity was found to lead the ROCs to become more curved, they were inconsistent with pure signal detection models as well. That is, although the zROCs became more linear with increased complexity, they still retained a U-shape across all four conditions which is inconsistent with the linear zROCs predicted by signal detection theory.

Why does complexity lead recollection to produce curved ROCs?

To answer this question it is useful to examine the response distributions for the simple and complex conditions. Figure 6 shows the distribution of confidence responses for simple and complex conditions when overall performance is matched (i.e., Experiment 1b). Interestingly, the proportion of correct and incorrect high confidence source responses (i.e., 1s and 6s) remained constant across simple and complex conditions, suggesting that complexity did not alter the number of high confidence source recollections. Rather, complexity increased the number of source responses that were associated with the second-to-highest level of confidence (the 5s and 2s, demarcated by arrows in Figure 6) and more importantly, the proportion of correct-to-incorrect responses at this level shifted to favor the correct source. However, because performance was still better for the highest confidence responses than for the 5 and 2 responses, performance was graded across the scale, which was necessary to produce the curved ROCs. Thus, complexity appears to have led to an increase in the recollection of lower confidence and less diagnostic source information.

It is not clear why complexity led to an increase in less diagnostic source recollection, but we suspect that it might be due to an increase in the numbers of features and associations available in the stimuli. That is, in the complex condition there are many different aspects of the item and the source that the subjects might attend to, and some of these aspects might not be entirely diagnostic of source (e.g., “the sentence is about animals and I remember that the male talked a lot about animals”). Information that is only weakly diagnostic of source may lead to an increase in the proportion of items attributed to a source with lesser confidence. A related possibility is that in the complex condition subjects may be more likely to remember different numbers of associations that can be useful in making the source discriminations, such that they remember two details on one trial (voice and visual information) but only one on another trial (voice). Confidence levels that differ as a function of number of details or associations recalled could lead to the observed pattern of results, if it is the case that details are falsely recollected in a similar way. Future studies will be necessary in order to further test these accounts.

How can current models of recollection account for these effects?

As discussed earlier, the DPSD model does not account for the effect of complexity because it assumes that recollection produces linear ROCs (U-shaped zROCs). Moreover,

recollection cannot be a pure signal detection process, because as with many previous studies, the source zROCs in the current experiments were consistently U-shaped. One potential solution is to use models that integrate the characteristics of these two classes of models. For example, the Variable Recollection Dual Process (VRDP) model (Sherman et al., 2003) is based on the DPSD model, but it assumes that the above-threshold items (i.e., recollected items) form a Gaussian recollection strength distribution (for similar models see DeCarlo, 2003; Kelley & Wixted, 2001; Onyper, Zhang, & Howard, 2010). The DPSD model simply assumes that above-threshold recollection supports high confidence responses, and it does not specify the shape of the recollection strength distribution. Threshold strength distributions are sometimes illustrated as being square, but they can take on any number of different shapes (e.g., see MacMillan & Creelman, 2005; Yonelinas, Aly, Wei-Chun & Koen, 2010). We currently don't know what the shape of the recollection distribution is, but the Gaussian assumption is a reasonable working hypothesis. The VRDP model includes the standard dual process model which has two memory parameters (i.e., recollection is estimated as a probability (R) and familiarity strength is measured as d' value (Fd')). But the model also includes another d' parameter to measure the strength of the above-threshold recollection distribution (Rd') and a parameter to measure the variance of the recollection distribution (V_R). When Rd' is large or when V_R is small, the model collapses into the standard two-parameter model. Because the model $[Hit = R * (\Phi(c_k, Rd', V_R) + (1-R) * \Phi(c_k, Fd', 1)) \text{ and } FA = (1-R) * \Phi(c_k, 0, 1)]^3$ has four free memory parameters it tends to be under-constrained by the current type of ROC data sets. For example, when the model is fit to ROC data there are often multiple good solutions, so it is not possible to provide a statistically meaningful assessment of the model. Because of this, our approach was to make some simplifying assumptions to fit the average ROCs in Experiment 2 in order to illustrate how such an approach could *in principle* account for the effects of complexity. We will leave a full assessment of the model for future studies.

Figure 7 illustrates how the VRDP model could account for the results of Experiment 1b. In the simple condition, familiarity is not very useful for source discriminations and so the two familiarity strength distributions (i.e., the “male” items and “female” items) are largely overlapping and centered quite close to the 3 and 4 guessing responses. Recollection is expected to be more useful in supporting source memory discriminations so the recollection distributions are spaced further to the left and right. The recollection distributions are illustrated as having less variance than the familiarity distributions, and as falling slightly above the 6 response criterion, but we could have chosen to represent them as having much greater variance, accompanied by much higher memory strength. In either case, if one assumes that subjects make their responses on the basis of recollection, and respond on the basis of familiarity when recollection fails, then these strength distributions would produce the observed response distributions seen in simple condition in Figure 6.

To account for data in the complex condition, the VRPD model either has to increase the variance of the recollection distribution or decrease the average strength of the recollection distribution (or both). The functional effect, regardless of the way this is achieved, is that more of the recollection distribution falls below the highest confidence level than in the simple condition. We illustrate a combination of both ways to achieve this effect in the complex condition panel in Figure 7 (i.e., a slight decrease in the strength of the recollection distribution and a slight increase in its variability). In this way, most experiences of recollection would still support high confidence source responses, but a proportion of recollection responses would now be associated with lower levels of confidence. Mixing the

³ R is the probability of recollection, Φ is the normal distribution function, c_k is the criterion for confidence level k , Rd' is the strength of the recollection distribution, V_R is the variance of the recollection distribution, Fd' is the strength of the familiarity distribution, and FA is false alarm.

recollection and familiarity distributions would produce the empirical response distributions in the complex condition in Figure 6. Note that the decision rule governing responses is the same as that underlying most dual process models whereby recollection is used as a basis for responding and familiarity is used when recollection fails (e.g., Mandler, 1980; Jacoby, 1991; Yonelinas, 1994). However, the model differs from the DPSD model in the sense that it allows for the possibility that there may be items for which familiarity is more diagnostic of source than is recollection. That is, recollection might support only a 5 response whereas familiarity could support a 6 response. Whether this occurs in practice is not known, but if the model is correct it should be possible. Overall, the VRDP model can account for complexity by assuming that the recollection distribution falls largely within the highest confidence bin in the simple condition, but that increasing complexity produces a weaker and more variable recollection distribution. This change in the recollection distribution results in more recollected items receiving lower confidence ratings than in the simple condition, which in turn produces more curved ROCs (or flatter zROCs).

The account that we have provided for the effects of complexity on ROC shape is similar to an account provided for the effects of another variable that changes ROC shape, feature overlap. Elfman et al. (2008) reported similar findings in testing the Complementary Learning Systems (CLS) model's prediction that increases in feature overlap would lead the threshold nature of recollection to break down (Norman & O'Reilly, 2003). When feature overlap was very high amongst stimuli (stimuli were all pictures of suburban houses, for instance), recollection's typical threshold nature broke down into a more graded process. The complexity manipulation in the current experiments would seem to have a similar effect on the nature of recollection—the more complex the stimuli, the more graded the recollection process appears to be.

To determine whether the CLS model could also account for the current results, we simulated the simple and complex conditions investigated in the current study with the model to test whether it would accurately predict that increasing complexity would result in more curved source ROCs. The CLS model uses biologically constrained parameters to simulate medial temporal lobe functioning. It consists of a recollection component which is based on the hippocampus, and a familiarity component which is based on the surrounding cortex. The hippocampal component of the model is specialized to rapidly encode associations between various features of a stimulus (e.g. item-source associations). Thus, when a test item is presented to the network the hippocampal component can retrieve the associated source information. Under standard conditions, the hippocampal component acts like a threshold process in the sense that it accurately retrieves the studied source, or it fails to provide any source information (i.e., recollection fails). We conducted the simulations with the standard hippocampal component of the model, using parameter values based on previous publications (e.g., Elfman et al., 2008). In the simple condition, an item ('A') was represented as a string of features (e.g. color, shape, size etc) which was paired with a second string of features that represented a source ('X'). In the complex condition each item was made up of multiple components (e.g., AB), and each source was made up of multiple components (e.g., XY). The network then had to learn two pairings (e.g., A–X and B–Y) in order to encode all the features. This piecewise approach was meant to capture the fact that as complexity increases, the number of specific associations that the subject can form for a given item-source pairing will increase).

To match overall performance across conditions, the item-source pairings in the simple condition were encoded twice each. The source decisions were modeled by presenting an item and allowing the hippocampal network to pattern complete the missing source information. The strength distributions and the derived ROCs for the simple and the

complex conditions are presented in Figure 8. For specific modeling details see Elfman et al., 2008 (also see Elfman & Yonelinas in preparation).

The strength distributions produced in the simple condition are consistent with results of previous simulations; namely, the model produces high confidence correct source responses (akin to 1s and 6s), but also shows a large proportion of retrieval failures (i.e. the middle peak). The resulting ROCs exhibit a large linear component, consistent with the empirical results. Most importantly however, is that the threshold shaped strength distributions took on a more Gaussian shape in the complex condition, and the resulting ROCs were more curvilinear, as was seen in the empirical study. Notably, the model also correctly predicted that increasing complexity results in increased source performance, despite the breakdown in the threshold nature of recollection (evidenced by the fact that it was necessary to present “simple” items to the network twice in order to achieve matched performance). It is important to emphasize that the CLS model was not designed to produce the specific types of ROCs observed here, but was designed to capture the neuroanatomical properties of the hippocampus. Thus, the fact that it was able to produce the observed ROC effects attests to the power of this neuro-computational modeling approach.

What are the implications of the current results for current measurement models of recollection and familiarity?

Our results suggest that when estimating recollection and familiarity in source memory tests, one needs to be cautious when interpreting parameter estimates whenever the materials become highly complex. In the current study, when we used simple stimulus materials, the VRDP model produce parameters that were effectively the same as those produced by the standard DPSD model. That is, because recollection led to high confidence responses, the full VRDP model effectively collapsed into the standard DPSD model. Given that most laboratory studies use simple materials like these it would suggest that the parameter estimates derived using the simple model would not be altered if the more complex model were adopted. A previous study comparing these models led to similar conclusions (Sherman et al., 2003; see also Onyper et al., 2010). Thus, we recommend using relatively simple materials when using the DPSD model in studies of source memory. When complex stimuli are of interest, it would be best to keep the source as simple as possible given the finding that complexity effects were much larger when both item and source were complex. Alternatively one could adopt a more complex model like the VRDP model, and use that to estimate recollection and familiarity. Although there are important issues with identifiability that will need to be addressed with these models, we feel that this approach is very promising.

The effect of complexity on recollection and familiarity in item recognition tests has not been formally assessed, and will need to be examined in future studies. However, to date, complexity does not appear to affect item recognition in the same way as it was found to affect source recognition. That is, the ROC method and the DPSD model have produced estimates that converge with those from other measurement procedures for complex stimuli like faces, sentences, and scenes (see Yonelinas, 2002 for a review), which one would not expect if the ROC estimates were biased by complexity. Moreover, in preliminary studies we have not found any noticeable effects of stimulus complexity on standard item recognition ROCs.

Overall, the data demonstrate that recollection can act simultaneously as a threshold and a graded process and that at least one critical factor determining this behavior is the complexity of the stimuli. Clearly, understanding memory for a range of stimuli from the simplest to the most complex is important, but if considered separately they may well lead to different conclusions about the nature of the processes supporting recognition. The effects of

complexity shown here along with those of Elfman et al. (2008) indicate that the behavior of recollection at the level of events is a variable feature that is dependent at least in part on the nature of the stimuli being remembered.

Acknowledgments

This work was funded by National Institute of Mental Health Grant MH59352. We thank Julie Jorgensen for help collecting data.

References

- DeCarlo LT. Source monitoring and multivariate signal detection theory, with a model for selection. *Journal of Mathematical Psychology*. 2003; 47(3):292–303.
- Diana RA, Reder LM, Arndt J, Park H. Models of recognition: a review of arguments in favor of a dual-process account. *Psychon Bull Rev*. 2006; 13(1):1–21. [PubMed: 16724763]
- Diana RA, Yonelinas AP, Ranganath C. The effects of unitization on familiarity-based source memory: testing a behavioral prediction derived from neuroimaging data. *J Exp Psychol Learn Mem Cogn*. 2008; 34(4):730–740. [PubMed: 18605864]
- Diana RA, Yonelinas AP, Ranganath C. Medial Temporal Lobe Activity during Source Retrieval Reflects Information Type, not Memory Strength. *J Cogn Neurosci*. 2009
- Dodson CS, Bawa S, Slotnick SD. Aging, source memory, and misrecollections. *J Exp Psychol Learn Mem Cogn*. 2007; 33(1):169–181. [PubMed: 17201560]
- Elfman KW, Parks CM, Yonelinas AP. Testing a neurocomputational model of recollection, familiarity, and source recognition. *J Exp Psychol Learn Mem Cogn*. 2008; 34(4):752–768. [PubMed: 18605866]
- Elfman, K.; Yonelinas, AP. Testing recognition memory predictions of the complementary learning system model. (in preparation)
- Glanzer M, Hilford A, Kim K. Six regularities of source recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 30:1176–1195.
- Glanzer M, Kim K, Hilford A, Adams JK. Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1999; 25(2):500–513.
- Hautus MJ, Macmillan NA, Rotello CM. Toward a complete decision model of item and source recognition. *Psychon Bull Rev*. 2008; 15(5):889–905. [PubMed: 18926981]
- Hilford A, Glanzer M, Kim K, DeCarlo LT. Regularities of source recognition: ROC analysis. *Journal of Experimental Psychology: General*. 2002; 131(4):494–510. [PubMed: 12500860]
- Jacoby LL. A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*. 1991; 30(5):513–541.
- Johnson MK, Hashtroudi S, Lindsay DS. Source monitoring. *Psychological Bulletin*. 1993; 114(1):3–28. [PubMed: 8346328]
- Kelley R, Wixted JT. On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27(3):701–722.
- Macmillan, NA.; Creelman, CD. *Detection theory: A user's guide*. 2. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2005.
- Mandler G. Recognizing: The judgment of previous occurrence. *Psychological Review*. 1980; 87(3):252–271.
- Manns JR, Hopkins RO, Reed JM, Kitchener EG, Squire LR. Recognition memory and the human hippocampus. *Neuron*. 2003; 37(1):171–180. [PubMed: 12526782]
- Norman KA, O'Reilly RC. Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological Review*. 2003; 110(4):611–646. [PubMed: 14599236]
- Onyper SV, Zhang Y, Howard MW. Some-or-none recollection: Evidence from item and source memory. *Journal of Experimental Psychology: General*. (In press).

- Qin J, Raye CL, Johnson MK, Mitchell KJ. Source ROCs are (typically) curvilinear: comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning Memory and Cognition*. 2001; 27(4):1110–1115.
- Rempel-Clower NL, Zola SM, Squire LR, Amaral DG. Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *J Neurosci*. 1996; 16(16):5233–5255. [PubMed: 8756452]
- Sherman SJ, Atri A, Hasselmo ME, Stern CE, Howard MW. Scopolamine impairs human recognition memory: data and modeling. *Behavioral Neuroscience*. 2003; 117(3):526–539. [PubMed: 12802881]
- Slotnick SD, Klein SA, Dodson CS, Shimamura AP. An analysis of signal detection and threshold models of source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2000; 26(6):1499–1517.
- Slotnick SD, Dodson CS. Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition*. 2005; 33(1):151–170.
- Shimamura AP, Wickens TD. Superadditive memory strength for item and source recognition: the role of hierarchical relational binding in the medial temporal lobe. *Psychol Rev*. 2009; 116(1):1–19. [PubMed: 19159146]
- Staresina BP, Davachi L. Differential encoding mechanisms for subsequent associative recognition and free recall. *J Neurosci*. 2006; 26(36):9162–9172. [PubMed: 16957073]
- Tulving E. Memory and consciousness. *Canadian Psychology*. 1985; 26(1):1–12.
- Wechsler. Wechsler Memory Scale-Revised Manual. The Psychological Corporation; San Antonio: 1987D. 1987
- Wixted JT. Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*. 2007; 114(1):152–176. [PubMed: 17227185]
- Yonelinas AP. Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1994; 20(6): 1341–1354.
- Yonelinas AP. The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1999; 25(6):1415–1434.
- Yonelinas AP. The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*. 2002; 46(3):441–517.
- Yonelinas AP, Aly M, Wang WC, Koen JD. Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*. 2010 Article first published online: 16 SEP 2010. 10.1002/hipo.20864
- Yonelinas AP, Kroll NE, Quamme JR, Lazzara MM, Sauve MJ, Widaman KF, et al. Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience*. 2002; 5(11):1236–1241.

A. Experiment 1a **B. Experiment 1b**

Figure 1.
 ROCs (left panels) and zROCs (right panels) for Experiments 1a and 1b.

A. Experiment 1a **B. Experiment 1b**

Figure 2. Average recollection and familiarity parameters from the DPSD model for Experiments 1a (top) and 1b (bottom). Error bars represent standard error of the mean.

Figure 3. Average d' for source binary decisions in the full and divided attention conditions as a function of encoding condition, with audio word the simple and video sentence the complex condition. Error bars represent standard error of the mean.

A. **B.** **Figure 4.**

A. Average d' for the controls and patients in each condition. B. Average recollection and familiarity parameter estimates for controls and patients in each condition. In both panels, bars represent standard error and individual scores of the patients are represented by the circles (GH) and triangles (DB).

Figure 5.
A. ROCs for controls and patients in each condition. B. zROCs for controls and patients in each condition.

Figure 6. Confidence response distributions from the simple and complex conditions of Experiment 1b.

Figure 7.
Familiarity and recollection distributions from the VRDP model for the simple and complex conditions of Experiment 1b.

Figure 8. Response distributions and associated ROCs simulated for simple and complex conditions by the hippocampal component of Complementary Learning Systems model.

Table 1Overall recognition performance (d_a)

	Audio Word	Audio Sentence	Video word	Video Sentence
Experiment 1a	.61 (.46)	.77 (.67)	.90 (.66)	1.36 (.62)
Experiment 1b	1.22 (.67)			1.34 (.51)

Note: Mean sensitivity (d_a) for encoding conditions in Experiments 1a and 1b (standard deviations in parentheses).

Table 2

Demographic Information and Standardized Test Results for Amnesic Patients

	Sex	Age	Edu	IQ	Att	WMS-R			
						Ver	Vis	Gen	Delayed
Patient 1	M	49	16	96	-0.20	-2.53	2.00	-1.40	-1.13
Patient 2	M	24	16	111	0.20	-1.33	0.33	-0.87	-2.13

Note: Age: age at testing; Edu: years of education; IQ: estimate of Wechsler Intelligence Scale; WMS-R: Wechsler memory scale- revised (Wechsler, 1987), all scores presented are z-scores; Att: Attention/Concentration; Ver: Verbal Memory; Vis: Visual Memory; Gen: General Memory; Delayed: Delayed Memory; Rep: delayed logical memory; Rep: delayed visual reproduction.