한국심리학회지 : 인지 및 생물
The Korean Journal of Cognitive and Biological Psychology
2009, Vol. 21, No. 3, 167-189

Exploring episodic decision variable using adaptive biased feedback procedure: Characteristic of decision axis during recollective judgment

Sanghoon Han

Center for Cognitive Neuroscience Duke University

Recognition memory theories have focused on the distinction between single- and dual-process models that characterize episodic retrieval decisions. Dual-process models suggest that episodic decision is contingent on two independent processes: familiarity and recollection. One outstanding question is whether the recollection reflects a distinct categorical mnemonic decision variable or a stricter criterion operation on the graded strength-of-evidence decision axis. The present study examined whether the decision variable of recollective judgment (i.e., "Remember" response) is qualitatively different from familiarity-based response by employing an adaptive biased feedback procedure that was previously used to investigate the lability of decision criterion on the continuous memory trace. The data demonstrated prominent and gradual criterion shifts for Remember response suggesting a continuous characteristics of recollective decision variable but the pattern shows a qualitative difference compared to familiarity-based recognition decisions. A potential model is also discussed.

Key words : Episodic memory decision variable, Recollection, Decision criterion, Biased feedback

† Corresponding Author : Sanghoon Han, Ph.D. Research Associate (Postdoctoral Researcher), Center for Cognitive Neuroscience Duke University, Durham, NC 27708 US
Phone : 919-684-3435, Fax : 919-681-0815, E-mail : sanghoon.han@duke.edu

Episodic recognition judgment refers to the decision process that supports retrieval of personally experienced item or event. Although it is often required to retrieve the context (or source) embedded with memory probes during acquisition and this source memory depends on distinct neural mechanisms (Dobbins, Foley, Schacter, & Wagner, 2002; Donaldson & Rugg, 1998), item-based episodic recognition typically requires subjects to parse continuous memory strength signals into discrete response categories (e.g., “Old” (studied) or “New” (unstudied)). One of the more successful and straightforward models to explain item memory is Signal Detection Theory (SDT). In its simplest form, SDT assumes that recognition performance is governed by a scalar indication of the amount of global mnemonic evidence (i.e., familiarity, x-axis in Fig. 1A), which is a continuous scale of memory strength value (Macmillan & Creelman, 1991).

In the basic detection model depicted in Figure 1, the samples of target and lure items yield overlapping Gaussian distributions on a strength-of-evidence continuum, which necessitates the use of a particular evidence value or values

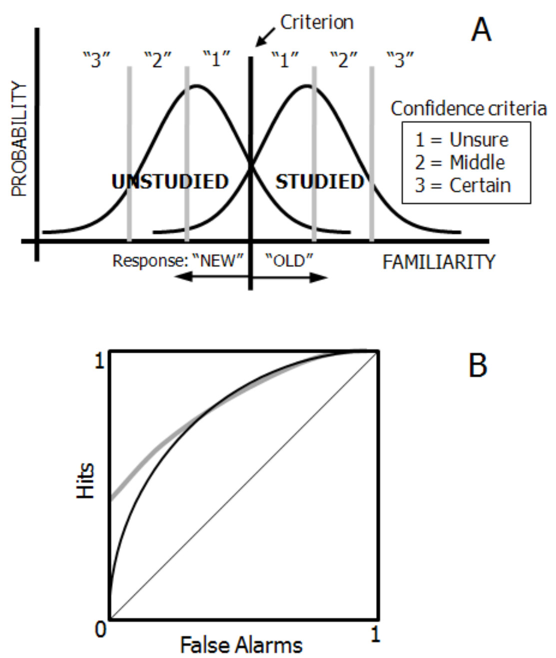


Figure 1. (A). Example of the density distributions of continuous amount of memory evidence values (i.e., familiarity) for new and old items in recognition memory. (B). Receiver Operating Characteristic (ROC) curves representing familiarity-based decision variable(black curve-curvilinear) or recollection -based component(gray curve-more linear).

as decision criteria during classification attempts. For simple “Old” or “New” classification only a single criterion is used. In contrast, when asked to make confidence ratings, the number of required criterion values is one less than the number of confidence categories available to the observer. When the correct Old response proportions (Hits) are plotted against incorrect endorsement of new items (False alarms) demarcated by the criteria and cumulated beginning with the most confidence endorsements, a curve known as the receiver operating characteristic (ROC) is traced out (Figure 1B).

The investigation of the shape (i.e., asymmetry) of ROC curves may provide a critical evidence that represents qualitatively distinct recognition decision variables during retrieval (i.e., symmetrical and curvilinear ROC curve vs. linear ROC), particularly during the high confidence responses (Fortin, Wright, & Eichenbaum, 2004; Yonelinas, 1994, 1997; Yonelinas, 2002). However the characteristics of the mnemonic decision variables accessed by retrieval operations in a standard recognition paradigm are still being debated (i.e., Greve, Donaldson, & van Rossum, in press; Mickes, Wais, & Wixted, 2009). More specifically, some item memory information can be recollected with fairly vivid contextual details whereas others leave behind only a feeling of familiarity. For example, it has been widely accepted that a symmetrical and curvilinear component ROC curve represents familiarity-based decision variable (black curve in Fig. 1B). On the other hand, a more linear component reflects recollection processes (light gray curve in Fig. 1B). The goal of the present study is to characterize the decision variable that governs more recollective item recognition judgments.

A method directly measuring different awareness states about the retrieval process during item recognition was proposed by Tulving (1985), by requiring participants to distinguish “Remember” or “Know” responses for their recognition of probe. This recognition procedure requires subjects to categorize endorsed test items as either evoking contextual recollections of prior experiences (viz., Remember) versus evoking a sense of recency, of item exposure, or item familiarity, without the recovery of specific contextual recollections (viz., Know) (Tulving, 1985). Thus, it is typically proposed that the Remember and Know responses engage two distinct retrieval states: Recollection and Familiarity, respectively, based on the distinct mnemonic evidence. However, there is currently considerable debate as to the utility of the procedure in fully isolating different underlying memory systems or processes (Dual- vs. Single-process model) (see below for more details). Therefore, despite the high degree of interest in the relation between Remember and overall Old/New recognition, it is still not clear whether the decision variables for these decisions reflect two qualitatively distinct memory traces.

Dual-process models, for example, explicitly assume participants can distinguish two distinct independent memory states (Recollection vs. Familiarity). Recollective judgment reflects a recall-like process of the memory episode consciously triggered by specific associative and contextual contents that accompanied the original memory item (e.g., Gardiner, 1988). Given that this decision recovers discrete information, most recollection research incorporated a high-threshold theory that assumes an all-or-none process (Yonelinas, 2002). The model holds that recollection either occurs or does not occur if the mnemonic evidence of a test item falls above a threshold or below the threshold. There is no intermediate event occurring between these two ends. In contrast, familiarity -based recognition involves more quantitative signal or strength-based information devoid of retrieval of details. As depicted in the SDT figure above, this decision requires a single criterion process on a continuous memory decision axis. Critically, the placement of decision criterion on the continuous decision axis determines the response tendency to endorse items as studied, and also the amount of recognized items. Despite a lack of consensus, a predominant view of distinct contributions of anatomical brain structures also support the dual-process model suggesting that activity in hippocampus and posterior parahippocampal region is linked to recollection while perirhinal cortex activation is linked to familiarity-based recognition (e.g., Davachi, Mitchell, & Wagner, 2003; Dobbins, Rice, Wagner, & Schacter, 2003; Kensinger, Clarke, & Corkin, 2003; Ranganath, et al., 2004).

In contrast to dual-process model, it has been argued that experimental dissociations of the response types can be adequately modeled by assuming they arise from the use of two different report criteria on a single decision axis (Donaldson, 1996; Dunn, 2004, 2008; Wixted & Stretch, 2004). More specifically, proponents of the one-dimensional signal detection model of Remember/Know data assume that episodic memories are retrieved by a single process operating on a single decision variable and that "Remember" reports simply reflect the use of a secondary and stricter criterion on the evidence dimension (Fig. 2). This two-criteria account has been increasingly focused on due to its clear advantage providing a relatively simple explanation for multiple mnemonic states or processes using a single decision axis. However, it is also widely accepted that the model is not suitable to explain the recent neuroanatomical evidence mentioned above that delineated the distinct neural mechanisms for either episodic retrieval states.

Adaptive biased feedback procedure and decision variables One known method to explore the decision axis is to examine the flexibility of decision criterion placed along the axis (e.g., Han & Dobbins, 2008; Morrell, Gaitan, & Wixted, 2002; Stretch & Wixted, 1998). If the decision variable reflects a threshold-like scale, then the theoretical signal detection estimates of criterion placement would not be influenced by strength manipulations of evidence since the typical SDT only assumes a continuous strength-of-evidence based decision axis. Conversely, if the decision variable for retrieval judgments reflects a single memory trace strength continuum, which is the scalar indication of quantitative mnemonic evidence level, then the placement of decision criterion on the axis should be adaptively flexible. For example, to explore observers' ability to adaptively adjust criterion during a typical standard Old/New recognition, Han and Dobbins (2008) employed a novel biased feedback methods, in which participants were systematically misinformed about incorrect responses of a certain type. More specifically, the design tacitly encourage certain errors by indicating they are the correct answers. For example, half the subjects were informed that incorrect rejection of Old items (Misses) were in fact correct whereas the other half were informed that incorrect endorsement of New items (False alarms) were indeed correct in the context of feedback that was otherwise accurate. The actual structure of the test list remained equivalent across the groups and subjects appeared to be unaware of the skewed nature of the feedback manipulation. The studies demonstrated that the feedback manipulation led to prominent and durable shifts in the relative decision criteria of the groups. That is, through the false positive feedback contingency manipulation, participants learned to adjust a criterion that would lead to a positive outcome more often (i.e., "Old" or "New") and altered the proportion of time participants report "Old" response to items.

A follow-up study by the same researcher group using the feedback manipulations further supports the idea that the effect may rely upon incremental reinforcement learning mechanisms (Han & Dobbins, 2009), which have been previously implicated in the learning of novel categories and response changes observed during probabilistic classification learning (e.g., Knowlton, Squire, & Gluck, 1994). The researchers found incrementally continued changes of criterion placements as the total amount of biased feedback accumulated within the test. Together, these findings of gradual flexibility of recognition decision criterion placement critically supports the adaptive change of response tendency on the continuous decision axis.

The goal of present study is to employ the biased feedback procedure and to determine whether Remember and overall Old/New recognition responses behave in the same manner as the criterion behavior documented in previous studies (Han & Dobbins, 2008; Han & Dobbins, 2009). If Remember and overall Old/New recognition responses represent two qualitatively discrete mechanisms (Dual-process model), then an experimental manipulation that shifts decision criterion on the strength-based decision axis should yield differential effects between Remember and Know responses. The manipulation might, for example, influence only the Know response rates based on the familiarity level of recognized items. In contrast, if Remember and Know responses work on the same strength-of-evidence based decision axis, any variables that affect recognition performance would alter the response patterns for both Old/New recognition and Remember responses in the same direction, depending on the type of feedback. This would show that they are not distinctively different response mechanisms but rather dependent on a strength-based decision variable on a single axis.

EXPERIMENT 1
Probabilistic Biased Feedback Procedure (Old/New Recognition)

The goal of Experiment 1 was to verify the biased feedback procedure to induce decision criterion shift in a direction manipulated within a test run. A serious potential drawback of previous feedback manipulation by Han and Dobbins (2008) was that the relative shift of criterion across test runs could have reflected rule-based controls between test runs rather than reflecting an adaptive adjustment in a trial-by-trial manner within the runs (for rigid and between-test criterion shifts, see Stretch & Wixted, 1998; Morrell et al., 2002). The current design employed the subtle probabilistic feedback manipulation (Han & Dobbins, 2009) but sought to detect the sustained learning effect of criterion shift within the same test list by interspersing recognition trials without feedback.

More specifically the experiment was designed to, on average, provide the biasing manipulation (i.e., false positive feedback) on 85 percent of a certain type of incorrect responses (i.e., only false alarms or only misses). This probabilistic manipulation was designed not only to discourage detection of the biasing feedback, but it simultaneously tests the robustness or sensitivity of the subjects to the feedback contingencies. It is important to note that the biased feedback manipulation only occurred on a small minority of total trials (for example, proportion of biased feedback trials * false alarm = .85 * .15 = .13). It was predicted that it would, given that the feedback effect is related to the incremental reinforcement learning rather than to the use of strategic rule or control. A secondary consequence of this probabilistic manipulation is that it reduces the overall frequency of biased feedback trials even further, again, testing the sensitivity of the subjects to the feedback contingencies. In Experiment 1, the same biased feedback condition were given to subjects across all three blocks (i.e., Strict -Strict-Strict or Lax-Lax-Lax) to induce more consistent learning effects towards the last run if it is the gradual reinforcement learning that governs the shift of criterion.

Methods

Subjects. Thirty Duke undergraduates participated in return for partial course credit and two separate groups of 15 were included in the analyses. Informed consent was obtained as required by the human subjects review committee of Duke University. The groups were randomly assigned to different feedback conditions. Four reported on the post-experiment questionnaire that they noticed irregularities in the feedback and also exhibited poor discrimination (d' < .5). Five subjects did not report any knowledge of the feedback manipulation, however, their performance also fell below the cutoff of a d' value of .5. All of these subjects were removed and replaced.

Procedures. A total of 600 nouns were drawn randomly for each subject (average of 7.09 letters and 2.34 syllables, with a Kucera-Francis corpus frequency of 8.85). From this, three lists of 200 items (100 old, 100 new items for each cycle) were constructed for use in three sequential study/test cycles. During study, the participants were instructed to rate words on the computer screen for the number of syllables ("Counting syllables 1/2/3 more than 4"). The study was identical for all cycles and for both groups of subjects, and forewarned testing immediately followed each study. Participants were not forewarned that feedback would be given during testing. During testing, old and new items were randomly intermixed and subjects made self -paced Old/New recognition judgments ("Is this OLD of NEW? 1= OLD 2 = NEW"). Following this, the subjects reported confidence on a scale of 1-3 ("Confidence? Unsure = 1 2 3 = Certain"). Feedback, when given, was immediately presented after the confidence report. Twenty percent of items (20 old, 20 new items) for each cycle did not include confidence reporting and feedback. The biased feedback manipulation was given to 85% of false alarm or miss trials accompanied by confidence reporting and feedback (see below).

Biased Feedback Manipulation: For half the subjects, all three tests selectively encouraged lax responding by probabilistically giving false positive feedback “That is CORRECT” for, on average, 85% of false alarm responses. All other response types (i.e., hits, misses and correct rejections) and 15% of false alarm trials were correctly identified by the feedback (Lax condition - L). For the other half of subjects, 85% of miss responses received false positive feedback (Strict condition - S) across all tests. Thus, the order of manipulation was feedback encouraging lax (LLL) response or feedback encouraging strict (SSS) response. A gray box surrounding the probe changed to green when the feedback for correct response (“That is CORRECT”) was presented for 200ms. The feedback for incorrect answers (“That is INCORRECT”) was presented with a red background screen for 200 ms.

Results and Discussion

Collecting confidence rating enabled the current analyses to inspect the shape of ROCs, which indicates that the recognition data in Experiment 1 follow the unequal variance signal detection (UEVSD) model (see Ratcliff, Sheu, & Gronlund, 1992 for more details). Thus, for more accurate investigation, the analyses employed the UEVSD estimates of accuracy and decision criterion A_z and c_a (Macmillan & Creelman, 1991; Rotello, Masson, & Verde, 2008) (see Appendix 1 for more details), although the standard equal variance signal detection (EVSD) estimates such as d’ or c yielded a similar pattern of results. *Accuracy* (A_z):

Table 1. Experiment 1 accuracy and decision criterion across groups and tests. Note: Values in parentheses indicate standard deviations. caNoFB denotes decision criterion for no-feedback trials.

	STRICT-STRICT-STRICT						LAX-LAX-LAX					
	TEST1		TEST2		TEST3		TEST1		TEST2		TEST3	
Hit	.64	(.12)	.59	(.16)	.51	(.16)	.73	(.10)	.71	(.13)	.75	(.13)
False Alarm	.19	(.09)	.16	(.11)	.16	(.12)	.26	(.12)	.36	(.17)	.39	(.18)
A_z	.80	(.08)	.75	(.11)	.77	(.08)	.81	(.05)	.76	(.09)	.78	(.08)
c_a	.23	(.22)	.34	(.31)	.45	(.36)	.03	(.33)	-.09	(.40)	-.24	(.50)
c_aNoFB	.27	(.34)	.33	(.33)	.25	(.30)	.08	(.41)	-.11	(.46)	-.35	(.53)

ANOVA for A_z with factors of Group (SSS or LLL) and Test (First, Second or Third) yielded no significant effects (Table 1).

Decision Criteria (c_a): The estimates for the Old/New decision criterion are shown in Table 1. A two-way ANOVA with factors of Group (SSS or LLL) and Test (First, Second or Third) yielded a main effect of Group ($F(1,27) = 12.06$, $MSe = 4.24$, $p < .005$) with the SSS group being more conservative than the LLL group. There was an interaction between Group and Test ($F(2,54) = 17.55$, $MSe = .42$, $p < .001$) indicating that the group differences in criterion estimates differed as a function of test. Post-hoc t-tests demonstrated that the difference in the criterion estimates approached significance ($t(27) = 1.95$, $p = .062$) for the groups in the first test and that the SSS group was significantly more conservative than the LLL group during the second test ($t(27) = 3.24$, $p < .005$) and the final test ($t(27) = 4.23$, $p < .001$).

For trials without feedback estimate of c_a was also analyzed to ensure the effect carried over to these trials. A two-way ANOVA with factors of Group (SSS or LLL) and Test (First, Second or Third) yielded a main effect of Group ($F(1,27) = 10.09$, $MSe = 3.71$, $p < .005$) with the SSS group being generally more conservative (mean $c_a = .28$) than the LLL group (mean $c_a = -.13$). There was also a main effect of Test suggesting that the criterion, collapsed across the groups, became more lax across the three tests ($F(1,54) = 6.21$, $MSe = .40$, $p < .005$). Importantly, there was an interaction between Group and Test ($F(2,54) = 4.58$, $MSe = .30$, $p < .05$) indicating that the group differences in criterion estimates differed as a function of test. This finding demonstrates the same pattern of criterion shifts as in feedback trials, reflecting that the differentially learned position of decision criterion across groups was sustained within the same test.

EXPERIMENT 2
Biased Feedback Procedure
(Remember/Know Judgment)

Given that the biased feedback manipulation successfully adjusted the decision criterion in Experiment 1, the main purpose of Experiment 2 was to directly investigate the feedback manipulation in the context of the Remember /Know procedure (see more details below). If the Old/New and Remember/Know criteria are in the same dimension and move in lockstep (Figure 2), then changes occurring in the overall recognition rates should be mirrored in the remember rates. In contrast, the dual-process theory predicts that "Remember" responses should be relatively insensitive to the feedback, in comparison to those associated with a

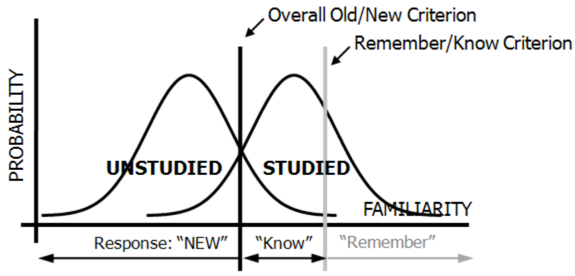


Figure 2. Two-criteria signal detection model for Remember/Know response

subjective sense of guessing (Overall Old/New bound). Given this, there should be an interaction between the overall Old/New recognition and Remember hit rates.

Methods

Subjects. Thirty undergraduate subjects were paid $10 for participation in the study. Informed consent was obtained as required by the human subjects review committee of Duke University.

Procedures. The nature of the judgment following the Old/New response was different from Experiment 1. Whereas Experiment 1 required a three-point confidence rating following Old/New discrimination, the current study required subjects to instead make a Remember /Know distinction following "Old" responses. More specifically, following an "Old" response, subjects were cued with "REMEMBER SPECIFICS? 1 = YES 2 = NO". They were instructed to respond "Remember" whenever they recollected seeing the study item. If recollection was absent, they were instructed to respond "NO" then the item was scored as "Know". Test responses were self-paced (see Appendix 2 for the experimental instruction).

Biased Feedback Manipulation: Given that Experiment 1 repeated the same type of feedback manipulation across all three test blocks and that four subjects still indicated the potential awareness of the manipulation on a post-experiment questionnaire, it was unclear whether the probabilistic nature of the biased feedback did in fact reduce participants' awareness in Experiment 1. This is because although the manipulation would be subtler than deterministically giving biased feedback to incorrect responses, repeating the same type of feedback across all three blocks may have made it easier to detect. In Experiment 2, the same feedback condition was not repeated across tests but instead employed the deterministic manipulation of feedback as in the previous study by Han and Dobbins (2008). Therefore, all errors of commission (False alarms) in the Lax condition or errors of omission (Misses) in the Strict condition received false positive feedback in Experiment 2, whereas correct responses were always indicated as such. Additionally, the unbiased correct feedback contingency run was included in the second test and the type of feedback manipulation was changed in the final test run to reduce the chance of subjects' awareness of manipulation towards the end of experiments. That is, for half the subjects, for example, feedback during the first test encouraged lax responding by giving false positive feedback, during the second test only correct feedback was given (Neutral condition, N), and the during last test the feedback encouraged strict responding (LNS). The order was reversed for the other half of the subjects (SNL).

Results and Discussion

Accuracy (Az). A two-way ANOVA for A_z with factors of Group (LNS or SNL) and Test (First, Second or Third) yielded no significant effects, indicating overall accuracy was unaffected by the feedback manipulations or test factor.

Decision Criteria (ca). In contrast to accuracy, the two-way ANOVA for the criterion estimate of Remember responses demonstrated a significant interaction between Group and Test ($F(2,52) = 5.43$, $MSe = .07$, $p < .005$) indicating that the group differences in criterion varied as function of test (Fig. 3A). Post-hoc t-tests demonstrated that the criterion difference

Table 2. Experiment 2 decision criterion across groups and tests. Note: Values in parentheses indicate standard deviations. ca Rem.: Criterion for Remember response.

	STRICT-NEUTRAL-LAX (SNL)						STRICT-NEUTRAL-LAX (SNL)					
	TEST1		TEST2		TEST3		TEST1		TEST2		TEST3	
(A). Remember												
Hit	.36	(.14)	.39	(.16)	.40	(.17)	.35	(.19)	.28	(.16)	.22	(.14)
False Alarm	.03	(.03)	.04	(.04)	.07	(.07)	.05	(.07)	.03	(.04)	.02	(.03)
c_a Rem.	1.14	(.38)	1.09	(.34)	0.98	(.57)	1.16	(.62)	1.38	(.52)	1.46	(.36)
(B). Overall "Old/New"												
Hit	.53	(.14)	.60	(.13)	.68	(.11)	.71	(.12)	.70	(.11)	.56	(.14)
False Alarm	.09	(.06)	.12	(.05)	.21	(.11)	.19	(.18)	.19	(.11)	.10	(.06)
c_a Overall	.54	(.28)	.46	(.25)	.13	(.16)	.20	(.41)	.17	(.26)	.53	(.26)

was not significant during Test 1 ($p > .48$) and although the difference was in the right direction, the criteria did not differ during Test 2 ($p > .09$). The SNL group, however, was significantly more liberal than the LNS group during Test 3 ($t(28) = 3.17$, $p < .005$), in which the feedback contingencies were changed.

For the overall Old/New response criterion, there was also a significant interaction of Group by Test ($F(2,52) = 24.56$, $MSe = .05$, $p < .001$) (Fig. 3B) indicating that the relative group differences in criterion changed across the tests. Simple post-hoc t-tests demonstrated that in Test 1, the SNL group was significantly more conservative than the LNS group ($t(28) = 2.64$, $p < .05$). This pattern remained during Test 2 in which all feedback was correct and unbiased ($t(26) = 3.05$, $p < .01$). Finally, during Test 3, in which the nature of the biased feedback was reversed for the groups, there was a significant reversal of the groups' criteria with the SNL group now demonstrating a more liberal criterion than the LNS ($t(28) = 5.14$, $p < .001$). Consistent with Experiment 1, the feedback manipulation effectively altered the relative position of the groups' overall Old/New decision criterion.

However, critically, visual inspection of criterion shifts in Figure 3 suggest that the shift patterns are qualitatively different. For example, whereas the Remember response decision criterion shows sluggish changes across test runs, the overall Old/New recognition criterion was fairly sensitive to the feedback manipulation showing the relative group difference starting from Test 1. To directly investigate the distinct patterns across response types, a mixed three-way ANOVA with factors of Group, Test, and Response Type (Overall Old/New response or Remember) was conducted. This analysis demonstrated a significant interaction across all three factors ($F(2, 52) = 8.62$, $MSe = .02$, p

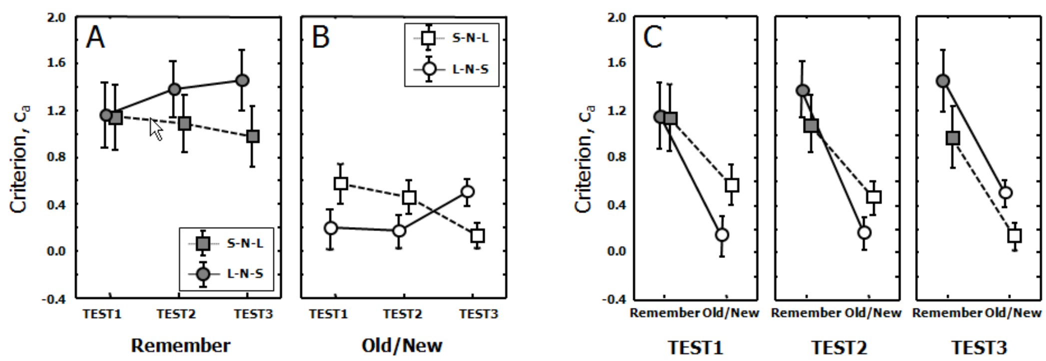


Figure 3. Experiment 2 decision criterion across groups and tests for Remember response (A) and overall Old/New response (B). (C) Group by Response Type interaction across test runs.

$< .001$). To decompose this interaction, the two-way ANOVA with factors of Group and Response Type was separately examined for each test (Fig. 3C). For Test 1, there was a significant interaction between Group and Response Type ($F(1,28) = 10.28$, $MSe = .008$, $p < .005$) which resulted because although the overall Old/New response criterion was significantly stricter for the SNL versus LNS groups ($t(28) = 2.64$, $p < .05$), the Remember response criterion did not differ between the feedback conditions ($t < 1$). Thus the feedback manipulation selectively affected the overall Old/New recognition rates whilst Remember responses were relatively less sensitive to the manipulation during Test1. For Test 2, in which all subjects received unbiased feedback, the pattern remained similar. There was a Group by Response Type interaction ($F(1,26) = 9.45$, $MSe = .13$, $p < .005$) which again resulted because whereas the overall Old/New response criterion was stricter for the SNL versus LNS group ($t(26) = 3.06$, $p < .01$) this ordinal pattern was not reflected in Remember criterion. In fact, the Remember criterion for the SNL group was only numerically higher than the LNS group and this approached significance ($t(28) = 1.75$, $p = .09$). Finally, in Test 3, there was no interaction between Group and Response Type ($F(1,28) = 0.81$, $MSe = .13$, n.s.) which as shown in Figure 3C (TEST3) resulted because both the overall Old/New criterion and Remember criterion were relatively lower for the group receiving the feedback encouraging lax responding (see Supplemental Data for the analogous patterns in individual hit and false alarm rates analyses).

Overall, the pattern of the overall Old/New criterion estimates is consistent with adaptive criterion shifts revealed in Experiment 1. Namely, the SNL group demonstrated a suppressed overall Old/New recognition criterion compared to the LNS group in Test 1, and this continued during Test 2 when the feedback was neutral. In Test 3, when the feedback manipulation was reversed across the groups, the pattern of overall Old/New criterion also reversed. More importantly however, the Remember criterion using a SDT estimate showed a qualitatively different criterion shift pattern demonstrating that the criterion determining Remember reports is not strictly linked to that governing overall Old/New recognition response rates.

Discussion

The question addressed in the present study is whether Remember responses during episodic recognition reflects decisions that are based on threshold-like recollective judgments, as is typically assumed in the standard dual-process models, or more strength-of-evidence based decisions. This issue was investigated by employing a biased feedback procedure and the findings demonstrated an adaptive criterion shift during recollective judgments. More specifically, Experiment 1 demonstrated the adaptive lability of decision criterion during the course of testing using a biased feedback technique designed to encourage certain type of response by indicating they were correct choices even when wrong. Crucially, the findings of criterion shifts even in the No-Feedback trials in Experiment 1 indicate that the learning effects induced by the feedback manipulation is not governed by rule-based controls between test runs but more related to adaptive adjustment in a trial-by-trial manner during testing (e.g., Han & Dobbins, 2008; Han & Dobbins, 2009). The critical contribution of the adaptive criterion shift is that the finding is entirely consistent with an assumption that the decision variable during an item recognition judgment, when the Remember/Know decisions were not required, reflects a continuous strength-based decision axis.

Experiment 2 demonstrated that the feedback manipulation differentially altered the location of the decision criterion of both overall "Old/New" and "Remember" responses across the two groups without producing notable differences in accuracy across the groups. The findings suggest that recollective responses need not rely on multiple independent memory traces. More specifically, the Remember judgments do not index a discrete retrieval state of memory but reflect a memory process with more stringent criterion on the continuous familiarity axis. These results are consistent with the previous findings that showed the manipulation that influenced criterion placement influenced both Remember and Know judgments (Higham & Vokey, 2004; Hirshman & Henzler, 1998). For example, Hirshman and Henzler (1998) assumed a model that has a single memory process (e.g. a single strength-based decision on familiarity decision axis) with two criteria separating the Old from New as well as Remember from Know responses on the same axis, as depicted in Figure 2. In their studies, test instructions about the proportion of study items (30% vs. 70%) were manipulated between subjects to see if this manipulation influences both Remember and Know response rates. This empirical test demonstrated that subjects who were told that 70% of the test items were study items produced more Remember and Know responses than subjects who were told that 30% of the test items were study items. The evidence of continuous decision variable during recollective judgments presented here offers potential advantages over more traditional dual-process approaches in that it does not assume that contextually specific information (viz. recollection) is retrieved in a high threshold or discrete state manner; an assumption that has been widely criticized, and that is potentially inconsistent with several findings (Dunn, 2004, Wixted & Stretch, 2004).

Although the previous findings has demonstrated the continuous criterial process during Remember judgments; nonetheless an outstanding question was whether the changes would be equivalent between Remember and Know responses. The present study addressed the questions using the manipulation of response -feedback contingency across three test runs. Our novel finding is that Remember criterion sluggishly and gradually changed relative to the overall "Old/New" criterion across tests (Figure 3). This finding implies that even though the Remember response rate was determined by the criterion manipulation through the biased feedback procedure Remember and Know criteria were influenced in different ways (i.e., sluggishly vs. immediately) and that the two responses may be represented on the different characteristics of either decision variables or dimensions. Importantly, the ANOVA for Remember false alarm rates also revealed significant Group by Test interaction in Experiment 2 indicating that the pattern that group differences in criterion estimates differed as a function of test may not be accommodated to the single high-threshold model (Yonelinas, 1994; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998) since the model does not provide the account for evoking a false recollection (i.e., Remember false alarm). This finding is consistent with the idea that although subjects appear to use a criterion for Remember report similar to the way they use one for providing Old/New judgments, Remember and Old/New endorsements are not necessarily based on the same type of content – merely the both types of content would have to be continuous or graded.

One potential approach to improve the generality of this continuous recollective judgment account might be to adopt two -dimensional (2D) signal detection models. For example, Rotello and colleagues proposed a 2D-SDT model to accommodate the performance of subjects using the Remember/Know procedure. The model assumes that items vary along both global and specific memory characteristics that are potentially reflective of the distinction between contextual recollection and item familiarity made in the dual-process models. However, unlike the traditional dual-process models both sources of memory information are assumed continuous (Rotello, Macmillan, & Reeder, 2004). Under the model, termed STREAK, when observers are simply asked to judge items as Old or New, evidence from both global and specific dimensions is summed in a weighted manner prior to criterial evaluation.

In contrast, when asked if they remember specifics about the prior encounter, the model assumes a differencing decision rule whereby endorsements require that the difference between specific and global information exceed a criterion value.

The current design was not intended to extend any models to more explicitly target hypothetical recollection and familiarity processes. Nonetheless, the data suggest a flexible basis for responding that is not consistent with either a one-dimensional approach or a high-threshold approach to modeling recollective recognition judgments (e.g., Donaldson et al., 1996; Dunn, 2004; Wixted & Sretch, 2004; Hirshman & Henzler, 1998). Furthermore, two-dimensional signal detection approaches do not require threshold assumptions and are also potentially compatible with functional imaging studies demonstrating large differences in response of the prefrontal cortex (PFC) across memory judgments that differ in specificity of the to-be-recovered content, even when materials are held constant (Dobbins et al., 2002; Nolde, Johnson, & D'Esposito, 1998; Rugg, Fletcher, Chua, & Dolan, 1999). To the extent various PFC regions are critical for the decision operations carried out on recovered memory content steering subjects towards or away from considering multiple types of memory evidence (potentially on the continuous 2D decision axes) would be expected to yield differential PFC responses. The modeling approach with the support of functional neuroimaging evidence will be an important area for the future investigation.

References

Davachi, L., Mitchell, J., & Wagner, A. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceeding of National Academic Sciences: USA,* 100, 2157-2162.

Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: multiple prefrontal processes subserve source memory. *Neuron,* 35(5), 989-996.

Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia,* 41(3), 318-333.

Donaldson, D. I., & Rugg, M. D. (1998). Recognition memory for new associations: Electrophysiological evidence for the role of recollection. *Neuropsychologia,* 36(5), 377-395.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, 24(4), 523-533.

Dunn, J. C. (2004). Remember-know: a matter of confidence. *Psychological Review*, 111(2), 524-542.

Dunn, J. C. (2008). The dimensionality of the Remember-Know task: A state-trace analysis. *Psychological Review*, 115(2), 426-446.

Fortin, N., Wright, S., & Eichenbaum, H. (2004). Recollection-like memory retrieval in rats is dependent on the hippocampus. *Nature,* 431, 188-191.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, 16(4), 309-313.

Greve, A., Donaldson, D., & van Rossum, M. (in press). A single-trace dual-process models of episodic memory: A novel compulational account of familiarity and recollection. *Hippocampus.*

Han, S., & Dobbins, I. (2009). Regulating recognition decisions through incremental reinforcement learning. *Psychonomic Bulletin & Review*. 16(3), 469-474.

Han, S., & Dobbins, I. G. (2008). Examining recognition criterion rigidity during testing using a biased feedback technique: Evidence for a adaptive criterion learning. *Memory & Cognition*, 36(4), 703-715.

Higham, P. A., & Vokey, J. R. (2004). Illusory recollection and dual-process models of recognition memory. *Quarterly Journal of Experimental Psychology: A*, 57(4), 714-744.

Hirshman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science*, 9(1), 61-65.

Kensinger, E. A., Clarke, R. J., & Corkin, S. (2003). What Neural Correlates Underlie Successful Encoding and Retrieval? A Functional Magnetic Resonance Imaging Study Using a Divided Attention Paradigm. *Journal of Neuroscience,* 23(6), 2407-2415.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory*, 1(2), 106-120.

Macmillan, N. A., & Creelman, C. (1991). Detection theory: A user's guide. New York, NY: Cambridge University Press.

Mickes, L., Wais, P., & Wixted, J. (2009). Recollection is a continuous process: Implications for dual-process theories of recognition memory. *Psychological Science,* 20(4), 509-515.

Morrell, H. E., Gaitan, S., & Wixted, J. T. (2002). On the nature of the decision axis in signal-detection-based models of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(6), 1095-1110.

Nolde, S. F., Johnson, M. K., & D'Esposito, M. (1998). Left prefrontal activation during episodic remembering: An event-related fMRI study. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, 9(15), 3509-3514.

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia,* 42(1), 2-13.

Ratcliff, R., Sheu, C.-f., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review,* 99(3), 518-535.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-Difference Theory of Remembering and Knowing: A Two-Dimensional Signal-Detection Model. *Psychological Review,* 111(3), 588-616.

Rotello, C. M., Masson, M. E. J., & Verde, M. (2008). Type I error rates and power analyses for single-point sensitivity measures. *Perception & Psychophysics,* 70(2), 389-401.

Rugg, M. D., Fletcher, P. C., Chua, P. M., & Dolan, R. J. (1999). The role of the prefrontal cortex in recognition memory and memory for source: an fMRI study. *Neuroimage,* 10(5), 520-529.

Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 24(6), 1397-1410.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology,* 26(1), 1-12.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, 11(4), 616-641.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20(6), 1341-1354.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition,* 25(6), 747-763.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 20 years of research. *Journal of Memory and Language,* 46(3), 441-517.

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness & Cognition: An International Journal,* 5(4), 418-441.

Yonelinas, A. P., Kroll, N. E., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology,* 12(3), 323-339.

1 차원고접수 : 2009. 6. 6
최종게재결정 : 2009. 9. 15

한국심리학회지 : 인지 및 생물
The Korean Journal of Cognitive and Biological Psychology
2009, Vol. 21, No. 3, 167-189

회상재인기억에서 의사결정변인의 특성에 대한 연구: 편향된 피드백 절차를 이용한 순응적 변화에 근거해

한 상 훈

듀크대학교 인지신경과학센터

재인기억이론들은 그동안 일화기억 인출과정을 설명하기 위해 단일작용 모형 혹은 이중작용 모형을 제안해왔다. 특히 이중작용 모형에서는 일화기억과 관련된 의사결정이 두 가지 독립적인 과정(친숙함 혹은 회상)에 기인한다고 설명하였다. 하지만, 회상재인기억이 실제 범주적인 특성을 갖는 개별적이고 독립적인 기억 의사결정 변인을 사용하는지, 아니면 연속적인 정도차를 보이면서 기억준거의 강도를 나타내는 의사결정 축을 사용하는지에 대해서는 아직 명확히 밝혀진 바가 없다. 본 연구에서는 회상재인기억 판단(예, “Remember” 반응)의 의사결정변인이 친숙도에 근거한 재인기억반응과 질적으로 다름을 보이는지를 연구하였는데, 이를 위해 앞선 연구들에서 연속적 기억변인을 바탕으로 재인의사결정준거의 순응적 이동성향을 보이는데 사용되었던 편향된 피드백 조건이 회상재인과제에서도 이용되었다. 본 실험의 결과에서는 회상재인기억반응에서도 점진적이지만 분명한 의사결정준거의 이동이 나타남을 보였는데, 이는 회상기억판단의 의사결정변인도 연속적인 정도차를 보임을 의미하는 것이라 볼 수 있다. 하지만 친숙함에 근거한 재인반응과 회상기억의 직접적인 비교에서는 두 기억반응의 피드백 조작에 따른 변화가 질적으로 서로 다른 양상이 나타났다. 이를 설명하기 위한 가능한 모형 또한 본문에서 논의되었다.

주요어 : 일화기억 의사결정 변인, 회상, 의사결정 준거, 편향된 피드백 절차

Appendix

1. To calculate the relative difference of variance for the old and new item distributions, the UEVSD model was individually fitted for each participant's recognition data using Excel's Solver routine. The logic under the fitting routine was to minimize the sum of squared errors of prediction for response proportions at each confidence criterion level. Each subject's new item distribution standard deviation was fixed to 1 given that the new item distribution is typically not affected by testing manipulations. And the Solver algorithm adjusted the distance between the distributions, the five decision confidence criteria, and the standard deviation of the old item distribution minimizing the sum of squared errors of prediction. The obtained variance estimates from the fit were entered the formulas below to calculate A_z and c_a (Rotello, et al., 2008; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996).

$$A_z = \Phi\left(\frac{\mu_{old}}{\sqrt{1+\sigma_{old}^2}}\right)$$

$$c_a = \frac{-\sqrt{2}s}{(1+s^2)^{1/2}(1+s)}[z(H)+z(F)]$$

where Φ indicates the transformation of the value by the unit normal distribution function, and μ_{old} and σ_{old} denote distance between the distributions and estimated standard deviations of old item distribution, respectively. s: new-to-old variance ratio. H: hits, F: false alarms, z: inverse of the unit normal distribution function (for details see Macmillan and Creelman, 1991).

2. Experimental instruction for Remember judgments in Experiment 2

"You will now be asked about your memory for the words that you have just seen. During this task, half of the words will be from the earlier syllable counting task and the other half will be new words that we did not show you before. For each word you will be asked two subsequent questions. The first will simply ask you to decide if you think the word was drawn from the syllable counting list (Is this OLD? 1=YES 2=NO). Press the '1' key for a "YES" response and the '2' key for a "NO" response. For those items that you judge as old, the computer will then ask a second question which requires you to rate the quality of your memory (Remember Specifics? Yes/No) This question is asking you if you can REMEMBER specific information regarding your prior viewing of the word. For example, you might remember accidentally pressing the wrong key for this word, or that the experimenter sneezed when it was on the screen, of that you thought that it was a strange word to use in an experiment, or that you found the syllable judgment difficult for this item. Any of these would count as a case of REMEMBERING and you should respond yes. In general, if you can remember specific information regarding your thoughts or actions during the prior encounter with that word, respond "YES". If instead you have a sense that the word is

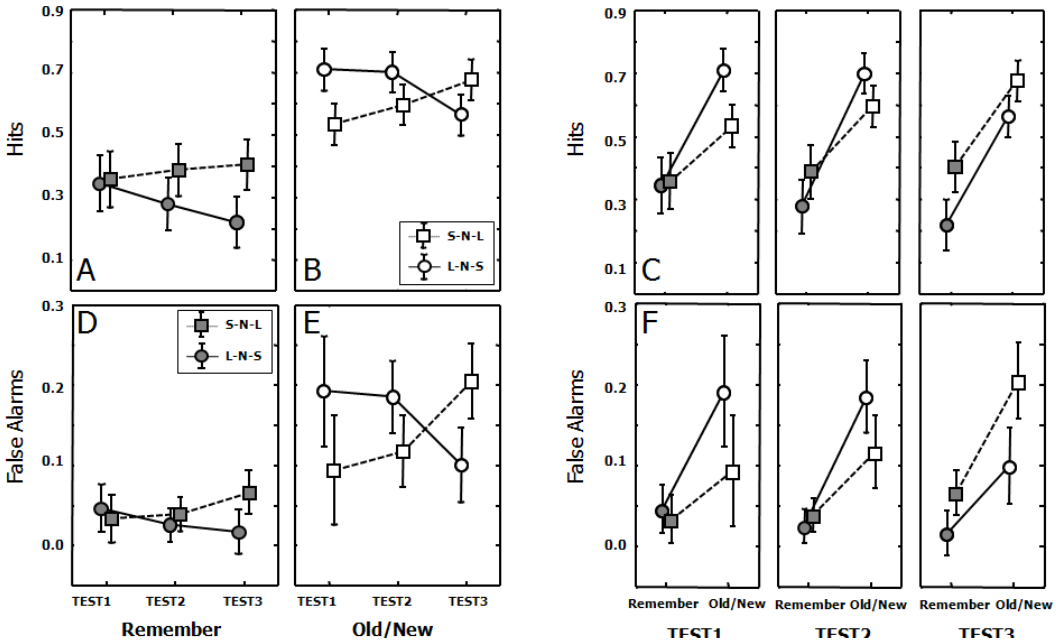
familiar and was presented, but you do not specifically remember the details of the prior judgment, you would respond "NO". You will have as long as you like to answer each question. Please be as accurate as you can. If you have any questions about this task, or are unsure which keys to press please ask the experimenter for more information."

Supplemental Data

Despite the demonstration of two qualitatively distinct criterion patterns in Experiment 2, one might object to the use of SDT theoretic criterion estimates (e.g., c_a) for Remember responses in the current analysis. For example, similar to previous studies, the false Remember response rate of the subjects was generally low in the present study, 17 of the subjects had a false Remember rate of .02 or lower during Test 1. Given this low detection theoretic statistics might be potentially unreliable at these extremes. For completeness, and to guard against analytic biases I jointly analyzed the overall recognition and remember hit rates to determine if they demonstrated the same pattern across the feedback manipulation and tests as the detection theoretic estimates did.

Hit rates: A two-way ANOVA for hit response rates with factors of Group (SNL, LNS) and Test (First, Second, Third) demonstrated a significant interaction in the Remember responses ($F(2,56) = 8.32$, $MSe = .01$, $p < .001$) (Suppl. Fig. 1A) as well as the overall Old/New recognition ($F(2,56) = 24.04$, $MSe = .01$, $p < .001$) (Suppl. Fig. 1B) indicating that the group differences in criterion varied as a function of test

A mixed three-way ANOVA with factors of Group, Test, and Response Type (Overall Old/New recognition or Remember) demonstrated a significant interaction across all three factors ($F(2, 56) = 7.54$, $MSe = .003$, $p < .01$). To decompose this interaction, I examined the two-way ANOVA with factors of Group and Response Type separately for each test (Suppl. Fig. 1C). For Test 1, there was a significant interaction between Group and Response Type ($F(1,28) = 14.44$, $MSe = .009$, $p < .001$) which resulted because although the recognition rate was significantly lower for the SNL versus LNS groups ($t(28) = 3.77$, $p < .001$), the correct Remember rate did not differ between the feedback conditions ($t < 1$). Thus the feedback manipulation only affected the overall Old/New recognition rates, the Remember responses were insensitive to the manipulation on this first test run. For Test 2, in which all subjects received unbiased feedback, the pattern remained similar. There was a Group by Response Type interaction ($F(1,28) = 12.38$, $MSe = .01$, $p < .01$) which again resulted because whereas the recognition hit rate was significantly lower for the SNL versus LNS group ($t(28) = 2.36$, $p < .05$) this ordinal pattern was not reflected in the correct Remember rates. In fact, the Remember rate for the SNL group was only



Supplemental Figure 1. Experiment 2 hits and false alarms as an index of criterion estimates across groups and tests. (A) Hits for Remember response (B) Hits for overall Old/New response (C) Group by Response Type interaction for hits across test runs. (D) False alarms for Remember response (E) False alarms for overall Old/New response (F) Group by Response Type interaction for false alarms across test runs.

numerically higher than the LNS group and this approached significance ($t(28) = 1.88$, $p = .070$). Finally, in Test 3, there was no interaction between Group and Response Type ($F(1,28) = 1.37$, $MSe = .01$, n.s.) which resulted because both the overall Old/New recognition rates and Remember rates were elevated for the group receiving the feedback encouraging lax responding.

False alarm rates: Further evidence suggesting the criteria for overall Old/New recognition and Remember rates are largely independent comes from considering incorrect recognition and remember responses (Suppl. Fig. 1D, E, F). As can be seen, the pattern for these is virtually identical to that observed for correct responding. Namely, a two-way ANOVA for hit response rates with factors of Group (SNL or LNS) and Test (First, Second, Third) demonstrated a significant interaction in the Remember responses ($F(2,56) = 5.66$, $MSe = .001$, $p < .01$) (Suppl. Fig. 1D) as well as the overall Old/New recognition ($F(2,56) = 14.58$, $MSe = .01$, $p < .001$) (Suppl. Fig. 1E) indicating that the group differences in criterion varied as function of test. There is a significant three-way interaction between Group, Test, and

Response Type ($F(2,56) = 11.94$, $MSe = .002$, $p < .001$). Again, this resulted because the pattern of two-way interactions for recognition and remember reports differed across the three tests (Suppl. Fig. 1F). Similar to the hit rate analysis above, there was a two-way interaction between Group and Response Type in Test 1 ($F(1,28) = 7.51$, $MSe = .004$, $p < .05$) which resulted because the decline in false recognition for the SNL compared to LNS groups ($t(28) = 2.08$, $< .05$) was not matched by a change in the false Remember rates ($t < 1$). Again, during Test 2 this pattern remained the same, with a Group by Response Type interaction ($F(1,28) = 7.67$, $MSe = .003$, $p < .01$) which occurred because the false recognition rate for the SNL group remained lower than the LNS group ($t(28) = 2.21$, $p < .05$) with no difference in the accompanying false Remember rates ($t < 1$). Finally, during Test 3 there was no interaction between the Group and Response Types ($F(1,28) = 3.06$, $MSe = .004$, $p = .091$) indicating that both the false recognition and false Remember rates were now similarly enhance for the SNL group in comparison to the LNS group.