Research Report

An Asymmetry Between Memory Encoding and Retrieval

Revelation, Generation, and Transfer-Appropriate Processing

Neil W. Mulligan and Jeffrey P. Lozito

University of North Carolina

ABSTRACT—*Transfer-appropriate-processing accounts of memory emphasize the similarity of encoding and retrieval processes, and imply that experimental manipulations should have similar effects on encoding and retrieval. Exceptions to this expectation are thus of great interest, but extant exceptions (produced by studies using divided attention, alcohol, and benzodiazepines) are debatable, single dissociations between encoding and retrieval. The present experiments demonstrate a reversed dissociation, in which the same variable produced opposite effects when implemented at encoding and retrieval. At encoding, participants either solved anagrams of study words or read intact study words. At retrieval, participants likewise solved anagrams or read intact words prior to making recognition memory judgments. Compared with reading intact words, solving anagrams at encoding enhanced later recognition accuracy, whereas solving anagrams at test impaired accuracy. These results were obtained with old/new decisions (Experiment 1) and with confidence ratings (Experiment 2).*

One of the fundamental ideas in modern memory research is that the match between encoding and retrieval dictates memory accuracy (Nairne, 2002; Roediger, Gallo, & Geraci, 2002), an idea presupposing substantial similarity between encoding and retrieval processes (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). This idea is embodied in the notion of transfer-appropriate processing (TAP; Morris, Bransford, & Franks, 1977; Roediger, 1990; Roediger et al., 2002), in encoding specificity (Tulving & Thompson, 1973), and in the procedural view of memory and skill acquisition (Kolers & Roediger, 1984). Similarly, Craik (1983) proposed that memory retrieval represents an attempt to recapitulate encoding processes. Consistent with this view are much behavioral data and recent neuroimaging research showing that successful retrieval is associated with heightened activity in the same cortical areas active during encoding (Roediger et al., 2002).

These lines of research emphasize the similarity of encoding and retrieval processes (Craik et al., 1996; Roediger et al., 2002). Reasoning from the TAP framework, Craik et al. (1996) argued that "if encoding and retrieval processes are indeed similar, experimental manipulations that affect one set of processes should have a similar effect on the other set" (p. 159). Despite the general success of the TAP framework, Craik et al. argued that there are exceptions to this expectation, and that these exceptions are critical for constraining TAP-based theories and for delineating the relation between encoding and retrieval.

Craik et al. (1996; Naveh-Benjamin, Craik, Gavrilescu, & Anderson, 2000) focused on the role of attention, finding that dividing attention during encoding produced dramatic reductions in later memory accuracy, whereas dividing attention during retrieval produced little or no effect. If retrieval recapitulates encoding, Craik et al. asked, why does this manipulation not exert similar effects at encoding and retrieval? Craik et al. (1996; Naveh-Benjamin et al., 2000) argued that alcohol and benzodiazepines produced similar dissociations. When administered prior to encoding, these substances reduced later memory accuracy, but when administered prior to retrieval, little effect on memory accuracy resulted (Birnbaum, Parker, Hartley, & Noble, 1978; Curran, 1991).

However, these dissociations are limited in their ability to delineate qualitative differences between encoding and retrieval

Address correspondence to Neil W. Mulligan, Department of Psychology, University of North Carolina, Chapel Hill, NC 27599-3270, e-mail: nmulligan@unc.edu.

for two related reasons. First, they are single dissociations, in which the manipulation affected encoding but had no measurable effect on retrieval. Single dissociations are ambiguous because they may reflect quantitative differences in sensitivity rather than qualitative differences in processing. Second, the clarity of the basic results is open to question. For example, researchers sometimes find that dividing attention during retrieval significantly reduces memory accuracy (e.g., Hicks & Marsh, 2000; Rohrer & Pashler, 2003). Even when these variables (divided attention, alcohol, benzodiazepines) fail to significantly reduce retrieval accuracy, there is virtually always a trend in the direction of the significant encoding effect. These points indicate that single dissociations between encoding and retrieval may be uncovering quantitative differences in sensitivity rather than qualitative differences between encoding and retrieval. From the perspective of the TAP approach, it is important to note that these variables do not produce wholly dissimilar effects on encoding and retrieval.

A more compelling result would be a reversed dissociation in which the same experimental manipulation produces opposite effects when implemented at encoding and retrieval. In this article, we report such a dissociation, presenting what is, to our knowledge, the first such demonstration.

The present study involved two memory phenomena. The first is a retrieval phenomenon, the revelation effect (Watkins & Peynircioglu, 1990), which is found by contrasting two conditions in a recognition memory test. In the revelation condition, the test item is slowly revealed immediately before the recognition memory decision. For example, the participant may solve an anagram that corresponds to the test word just prior to making an old/new recognition judgment. This condition is contrasted with a typical test condition, in which the test item is simply presented in its normal, intact form. The revelation condition produces more "old" responses (irrespective of study status) than the intact condition. Preceding the recognition decision with an unrelated task (e.g., solving an anagram that does not correspond to the subsequent test item) produces a similar, though not identical, effect (Verde & Rotello, 2004). Interestingly, the revelation effect is typically larger for new than old items, increasing false alarm rates more than hit rates (especially when retrieval conditions are manipulated between subjects; Hicks & Marsh, 1998). Thus, a second effect of the revelation manipulation is a net decrease in recognition accuracy.

The second phenomenon of interest is an encoding phenomenon, the generation effect, in which material that is self-generated during encoding produces better memory than material that is merely perceived (Mulligan & Lozito, 2004). Typical generation manipulations include generating words to semantic cues (e.g., antonyms) or generating words by transposing reversed letters (e.g., generating "horse" given "ohrse"). It is interesting to note that the typical revelation manipulation (solving an anagram prior to the recognition decision) could be fashioned into a generation manipulation if implemented at encoding. This raises the possibility that the same manipulation could produce opposite results when implemented at encoding versus retrieval. Specifically, compared with reading a word in its intact form, solving an anagram at encoding should enhance later memory, whereas solving an anagram prior to a recognition decision should impair memory. Critically, this empirically derived prediction is at odds with the TAP framework, which typically assumes that similar overt processing induces the transfer-appropriate advantage (e.g., Craik et al., 1996; Morris et al., 1977; Roediger et al., 2002).

EXPERIMENT 1

Method

Participants

The participants were 48 undergraduates at the University of North Carolina.

Design and Materials

Study condition (intact word vs. anagram) and test condition (intact word vs. anagram) were manipulated between subjects. The critical items were 80 familiar eight-letter words. These words were randomly divided to create two study lists, each consisting of 40 critical items plus 4 buffer items (2 primacy and 2 recency). The two study lists were presented to the same number of participants, so that the critical items were equally often old and new at test. The recognition test consisted of all 80 critical items: 40 old and 40 new.

Procedure

The experiment consisted of a study phase, a distractor task, and the recognition test. In the *study-anagram* condition, each study word was presented as an anagram, with a number below each letter indicating the position of the letter in the anagram's solution (e.g., a "1" indicated the first letter of the solution). The letters of the anagram were always presented in the same pattern to facilitate successful solution (the pattern was 54687321). The study list was preceded by two practice trials to familiarize the participant with the solution rule. Each anagram remained on the screen until the participant responded with a solution (whereupon the experimenter advanced the screen) or for 10 s (a duration long enough for participants to respond to the vast majority of the anagrams). Next, the study word (i.e., the solution for the anagram) was presented for 2 s. In the *study-intact* condition, each study word was presented for 2 s,[1] and participants were asked to read each word aloud. In both conditions,

[1]Studies of the generation effect equate either functional presentation time (i.e., the duration from the point of identification of the study word) or nominal study-trial duration. In the present experiment, the need to match the encoding variable with the revelation manipulation dictated the former strategy. For present purposes, the critical issue is that the encoding and retrieval variables were identical.

TABLE 1
Mean Hit and False Alarm (FA) Rates and d′ as a Function of Study and Test Conditions in Experiment 1

	Study condition					
	Anagram			Intact		
Test condition	Hit	FA	d'	Hit	FA	d'
Anagram	.91	.20	2.38	.77	.36	1.12
Intact	.89	.11	2.68	.74	.24	1.58

participants were asked to try to remember the study words for a later test. After the study phase, participants completed a 3-min distractor task (a task in which they completed word stems with the names of U.S. cities).

The memory test followed. Participants were informed that they would be presented with a series of words, some of which were old words (from the study phase) and some of which were new. In the *test-anagram* condition, each test item was first presented as an anagram to be solved. The anagram presentation was the same as in the study-anagram condition. Then, after the anagram was solved (or 10 s elapsed), the test item was presented for an old/new recognition decision. In the *test-intact* condition, each test item was first presented in its intact form to be read aloud (presentation the same as in the study-intact condition). Then, the old/new recognition decision was made.

Results and Discussion

First, we verified that our test manipulation produced the expected revelation effect on the proportion of "old" responses in the recognition test (Table 1). The effect of test condition indicated that solving an anagram prior to recognition increased the proportion of "old" responses (i.e., the revelation effect was obtained), $F(1, 44) = 6.43$, $MSE = 0.0143$ ($\alpha = .05$ for this and subsequent analyses). Furthermore, test condition interacted with old/new status, $F(1, 44) = 4.67$, $MSE = 0.0088$, indicating that the revelation effect was more pronounced for false alarms than hits.

Recognition accuracy was assessed with d' and submitted to a 2 (study condition) × 2 (test condition) analysis of variance (ANOVA). The analysis revealed an effect of study condition, $F(1, 44) = 41.99$, $MSE = 0.397$, indicating that the study-anagram condition enhanced accuracy, and an effect of test condition, $F(1, 44) = 4.35$, $MSE = 0.397$, indicating that the test-anagram condition disrupted accuracy relative to the intact condition. The interaction was nonsignificant ($F < 1$).

Test condition produced the expected revelation effect on the proportion of old responses in two ways. The test-anagram condition increased the number of "old" responses, and the effect was greater on false alarms than hits. More critical are the accuracy results. The effect of study condition can be conceived of as a generation effect: Solving the anagram produced greater memory than simply reading the study item. The same manipulation at test produced the opposite result: Solving the anagram at retrieval disrupted recognition accuracy relative to simply reading the test item. Together, these results represent a reversed dissociation: The same manipulation produced opposite results when implemented at encoding and retrieval (Fig. 1).

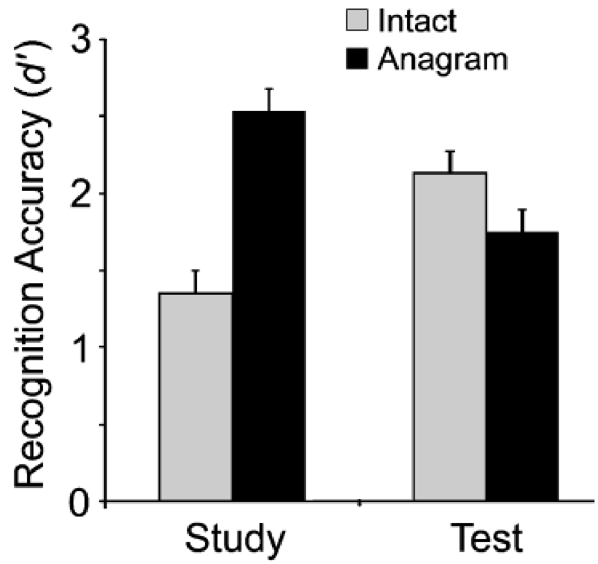

Fig. 1. Effects of the study and test manipulations on recognition accuracy in Experiment 1. The graph shows mean d' in the two study conditions averaged over test conditions and in the two test conditions averaged over study conditions.

EXPERIMENT 2

Verde and Rotello (2003) recently questioned the use of d' to assess accuracy in recognition memory when bias varies (as in the revelation effect), instead recommending the d_a measure. The computation of d_a requires multiple points on the receiver-operator-characteristic (ROC) curve and cannot be obtained from a single pair of hit and false alarm rates as produced by standard old/new recognition. In Experiment 2, we used confidence ratings during the recognition test to obtain ROC curves and compute d_a, and verified the results of Experiment 1.

Method

Experiment 2 used the same method as Experiment 1 with the following modification. In lieu of old/new recognition decisions, participants made confidence ratings as follows: 1 = very sure new; 2 = sure new; 3 = somewhat sure new; 4 = somewhat sure old; 5 = sure old; 6 = very sure old. These labels appeared on the bottom of the computer screen for each test item. Sixty-four undergraduates participated.

Results and Discussion

Hit and false alarm rates were computed by coding confidence ratings 1 through 3 as "new" and 4 through 6 as "old" (Table 2). Relative to the intact test condition, the anagram test condition

TABLE 2
Mean Hit and False Alarm (FA) Rates and d_a as a Function of Study and Test Conditions in Experiment 2

	Study condition					
	Anagram			Intact		
Test condition	Hit	FA	d_a	Hit	FA	d_a
Anagram	.86	.26	2.12	.73	.37	1.08
Intact	.86	.16	2.50	.70	.20	1.60

significantly increased "old" judgments, $F(1, 60) = 11.36$, $MSE = 0.015$, and did so more for false alarms than hits (i.e., test condition interacted with old/new status), $F(1, 60) = 9.92$, $MSE = 0.018$.

The confidence ratings were submitted to maximum-likelihood estimation of the ROC curve for each participant, allowing the computation of d_a (Macmillan & Creelman, 1991). The average ROC curves are presented in Figure 2. Sensitivity (d_a) was submitted to a 2 (study condition) × 2 (test condition) ANOVA. The analysis revealed main effects of study condition, $F(1, 60) = 32.46$, $MSE = 0.464$, and test condition, $F(1, 60) = 6.81$, $MSE = 0.464$, with no interaction ($F < 1$). These results indicate that solving the anagram at study increased recognition accuracy, whereas solving the anagram at test disrupted accuracy.

Experiment 2 replicated Experiment 1 in all particulars. A revelation effect was found, primarily on false alarms. The identical anagram manipulation produced opposite effects on encoding and retrieval. Obtaining confidence ratings allowed us to verify that the reversed dissociation was not an artifact of the particular accuracy measure used in Experiment 1 (cf. Verde & Rotello, 2004).

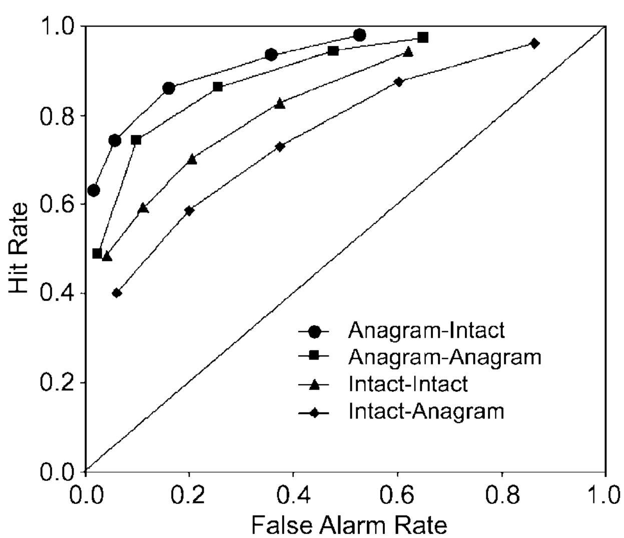

Fig. 2. Mean receiver-operator-characteristic (ROC) curves obtained from the confidence ratings in Experiment 2, as a function of study condition (anagram vs. intact) and test condition (anagram vs. intact). In the legend, the first word in each label refers to the study condition, and the second refers to the test condition.

GENERAL DISCUSSION

TAP and related theoretical frameworks emphasize encoding-retrieval match, arguing that successful memory retrieval recapitulates encoding processes, which in turn implies that experimental manipulations ought to have similar effects on encoding and retrieval. Consequently, it is critical to document variables having different effects on encoding and retrieval because such dissociations are important limiting conditions on the general TAP view and because they delineate differences between encoding and retrieval processes (Craik et al., 1996; Naveh-Benjamin et al., 2000). The single dissociations identified thus far are debatable because they may demonstrate only quantitative differences in sensitivity rather than qualitative differences.

The present experiments demonstrate a more compelling reversed dissociation, in which the same manipulation produced opposite effects on memory accuracy when implemented at encoding and retrieval. Solving anagrams at study produced an advantage in later recognition accuracy (akin to the generation effect), whereas solving anagrams at test disrupted accuracy (in addition to producing the typical revelation effect).[2] To our knowledge, this is the first behavioral evidence of a reversed dissociation between encoding and retrieval (see research on the hemispheric encoding/retrieval asymmetry, or HERA, model for neuroimaging evidence of dissociation; e.g., Habib, Nyberg, & Tulving, 2003). This result constitutes an exception to the typical expectations of the TAP account.

One possible approach to reconciling these results with the broader TAP principle would be to question the extent to which the same overt processing task actually recruits the same set of cognitive operations during encoding and retrieval. However, if invoked selectively, such a tack runs the risk of rendering TAP-based accounts unfalsifiable unless embedded in a more detailed analysis of encoding and retrieval subprocesses, allowing for principled explanation of when similar overt tasks do and do not invoke similar sets of cognitive operations (see Craik et al., 1996, and Naveh-Benjamin et al., 2000, for examples of such an analysis).

In the present case, a preliminary analysis can be built on existing research. Research on the generation effect indicates it is likely that solving anagrams during encoding enhances item-

[2]An early generation study might seem relevant at first. Glisky and Rabinowitz (1985) manipulated generation at encoding and retrieval, and found, in some cases, that repetition of the generation task at retrieval enhanced accuracy (a result consistent with the TAP view). However, this study generally failed to demonstrate a revelation effect, possibly because the generation task was relatively easy (supply a missing letter to a word) and implemented within subjects, both conditions that minimize the revelation effect (e.g., Hicks & Marsh, 1998; Watkins & Peynircioglu, 1990). Our experiments demonstrate that a more challenging revelation task produces both the revelation effect and the concomitant reduction in accuracy.

specific processing, augmenting distinctiveness in memory (see Mulligan & Lozito, 2004, for review). There is less agreement about the basis of the revelation effect (Hicks & Marsh, 1998; Verde & Rotello, 2003). One possibility is that solving anagrams prior to recognition not only affects the perceived familiarity of the test item, but also increases the variability of the familiarity distributions, resulting in increased signal-to-noise ratio and decreased recognition accuracy (Verde & Rotello, 2004). This specific account remains to be evaluated.

REFERENCES

Birnbaum, I.M., Parker, E.S., Hartley, J.T., & Noble, E.P. (1978). Alcohol and memory: Retrieval processes. *Journal of Verbal Learning and Verbal Behavior*, *17*, 325–335.

Craik, F.I.M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society B*, *302*, 341–359.

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

Curran, H.V. (1991). Benzodiazepines, memory, and mood: A review. *Psychopharmacology*, *105*, 1–8.

Glisky, E.L., & Rabinowitz, J.C. (1985). Enhancing the generation effect through repetition of operations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 193–205.

Habib, R., Nyberg, L., & Tulving, E. (2003). Hemispheric asymmetries of memory: The HERA model revisited. *Trends in Cognitive Sciences*, *7*, 241–245.

Hicks, J.L., & Marsh, R.L. (1998). A decrement-to-familiarity interpretation of the revelation effect from forced-choice tests of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1105–1120.

Hicks, J.L., & Marsh, R.L. (2000). Toward specifying the attentional demands of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1483–1498.

Kolers, P.A., & Roediger, H.L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, *23*, 425–449.

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

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533.

Mulligan, N.W., & Lozito, J.P. (2004). Self-generation and memory. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 45, pp. 175–214). San Diego, CA: Elsevier Academic Press.

Nairne, J.S. (2002). The myth of the encoding-retrieval match. *Memory*, *10*, 389–395.

Naveh-Benjamin, M., Craik, F.I.M., Gavrilescu, D., & Anderson, N.D. (2000). Asymmetry between encoding and retrieval processes: Evidence from divided attention and a calibration analysis. *Memory & Cognition*, *28*, 965–976.

Roediger, H.L. (1990). Implicit memory: Retention without remembering. *American Psychologist*, *45*, 1043–1056.

Roediger, H.L., Gallo, D.A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory*, *10*, 319–332.

Rohrer, D., & Pashler, H.E. (2003). Concurrent task effects on memory retrieval. *Psychonomic Bulletin & Review*, *10*, 96–103.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.

Verde, M.F., & Rotello, C.M. (2003). Does familiarity change in the revelation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 739–746.

Verde, M.F., & Rotello, C.M. (2004). ROC curves show that the revelation effect is not a single phenomenon. *Psychonomic Bulletin & Review*, *11*, 560–566.

Watkins, M.J., & Peynircioglu, Z.F. (1990). The revelation effect: When disguising test items induces recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1012–1020.

(Received 3/22/05; Revision accepted 5/23/05; Final materials received 6/1/05)