

Relational and item-specific influences on generate–recognize processes in recall

Melissa J. Guynn · Mark A. McDaniel · Garrett L. Strosser ·
Juan M. Ramirez · Erica H. Castleberry · Kristen H. Arnett

Published online: 15 January 2014
© Psychonomic Society, Inc. 2014

Abstract The generate–recognize model and the relational–item-specific distinction are two approaches to explaining recall. In this study, we consider the two approaches in concert. Following Jacoby and Hollingshead (*Journal of Memory and Language* 29:433–454, 1990), we implemented a production task and a recognition task following production (1) to evaluate whether generation and recognition components were evident in cued recall and (2) to gauge the effects of relational and item-specific processing on these components. An encoding task designed to augment item-specific processing (anagram-transposition) produced a benefit on the recognition component (Experiments 1–3) but no significant benefit on the generation component (Experiments 1–3), in the context of a significant benefit to cued recall. By contrast, an encoding task designed to augment relational processing (category-sorting) did produce a benefit on the generation component (Experiment 3). These results converge on the idea that in recall, item-specific processing impacts a recognition component, whereas relational processing impacts a generation component.

Keywords Memory · Recall · Recognition · Memory models

Portions of this research were presented at the 79th Annual Meeting of the Rocky Mountain Psychological Association, Albuquerque, NM (2009) and the 50th Annual Meeting of the Psychonomic Society, Boston, MA (2009).

M. J. Guynn (✉) · J. M. Ramirez · E. H. Castleberry · K. H. Arnett
Department of Psychology, New Mexico State University, MSC
3452, P.O. Box 30001, Las Cruces, NM 88003-8001, USA
e-mail: mguynn@nmsu.edu

M. A. McDaniel
Department of Psychology, Washington University in St. Louis,
St. Louis, MO, USA

G. L. Strosser
Department of Psychology, Southern Utah University,
Cedar City, UT, USA

A central challenge for memory researchers is to characterize the processes that support recall (e.g., Hogan & Kintsch, 1971; Kintsch, 1968; Raaijmakers & Shiffrin, 1981, 1992; Tulving & Pearlstone, 1966; Tulving & Thomson, 1973). One long-standing framework is the generate–recognize model (e.g., Anderson & Bower, 1972; Bahrick, 1970; Bellezza, 2003; Jacoby, 1998; Jones, 1978; Kintsch, 1970, 1978; Watkins & Gardiner, 1979), according to which there are two phases or components of recall. Candidates for recall first are generated and then are submitted to a recognition check that excludes some of the candidates; that is, only the generated candidates that are recognized as having been encountered are recalled. Although the classic version of the generate–recognize model was discredited by results such as the recognition failure of recallable words (Tulving & Thomson, 1973; Watkins & Gardiner, 1979), Jacoby and Hollingshead (1990) introduced a revised model that accounts for the problematic results. A critical assumption of the revised model is that individuals do not inevitably attempt to recognize the generated candidates. If an item is generated relatively fluently, an individual may recall it without subjecting it to a recognition check. Recall in this revised model is supported by “generate” and “sometimes recognize” processes. For convenience, we will refer to this revised model using the simpler label *generate–recognize*. Since its inception, further evidence for this model has been reported by Guynn and McDaniel (1999; and later in this article).

Another approach to recall has developed somewhat independently of the generate–recognize model. In this framework, recall is characterized as being dependent upon two different types of information: (1) relational information, or information about similarities among stimuli, and (2) item-specific information, or information about unique or nonoverlapping features of stimuli (Einstein & Hunt, 1980; Hunt & Einstein, 1981). This relational–item-specific distinction basically integrates two venerable ideas about the types of encoding processes that benefit recall—organization of items (e.g., Mandler,

Pearlstone & Koopmans, 1969; Tulving, 1962) and elaboration of each item qua item (Craik & Lockhart, 1972; Craik & Tulving, 1975)—and has been applied to explain a number of recall phenomena (see Hunt & McDaniel, 1993, for a review).

The objective of the present study was to consider these two approaches to recall in concert. One possibility is that relational and item-specific information influence recall through their selective effects on generation and recognition, respectively. That is, relational information is used to generate candidates for recall, and item-specific information is used during the recognition phase to discriminate the studied items among the generated candidates (Einstein & Hunt, 1980; Hunt & Einstein, 1981). Thus, on this idea, a generation phase would be most sensitive to relational information, whereas a recognition phase that follows a generation phase would be most sensitive to item-specific information (Einstein & Hunt, 1980; Hunt & Einstein, 1981; McDaniel, Einstein & Lollis, 1988). Using somewhat different terminology, item-specific information could be conceptualized as impacting a “late correction” process that edits the products of an earlier generation process (Jacoby, Bishara, Hessels & Toth, 2005; Jacoby, Kelley & McElree, 1999).

An alternate possibility is that relational and item-specific information jointly enable retrieval processes to specify more precisely the items to be retrieved, thereby fostering more successful reproduction of those items (Hunt & McDaniel, 1993; Hunt & Smith, 1996; Smith & Hunt, 2000). This idea is more in line with the possibility that recall involves direct access to memory traces, with rich episodic features (including both relational and item-specific information) being used to guide this direct access (e.g., Guynn & McDaniel, 1999; Tulving & Thomson, 1973; Watkins & Gardiner, 1979; Weldon & Colston, 1995). A key implication of this idea is that a generate–recognize model would not capture the processes that support recall. Item-specific information, rather than influencing a recognition phase that follows a generation phase, would provide distinctive features that could be exploited to produce the items to be retrieved (Knoedler, Hellwig & Neath, 1999; McDaniel, DeLosh & Merritt, 2000) or, using somewhat different terminology, could impact an “early selection” process that specifies the information to be retrieved (Jacoby et al., 2005; Jacoby et al., 1999).

The experimental paradigm

To gain leverage on whether separate processes of generation and recognition are evident in recall and, if so, whether they are influenced by relational and item-specific processing, respectively, we adopted the method and assumptions described by Guynn and McDaniel (1999; based on seminal work by Jacoby & Hollingshead, 1990). Across our experiments, participants took one of three memory tests: (1) an

implicit test of category production, to reflect a process of generation; (2) an explicit test of category-cued recall, which was expected to entail generation plus a recognition check for some but not all generated candidates (i.e., the candidates that were not generated relatively fluently); or (3) an implicit test of category production plus an explicit recognition instruction, to reflect processes of generation plus recognition for all generated candidates. Thus, the three memory tests were assumed to involve the same generation process but to differ in the extent to which the recognition process played a role (see also Guynn & McDaniel, 1999; Jacoby & Hollingshead, 1990).

A priori, a pattern can be specified that would imply that participants had adopted a generate–recognize strategy for recall. Recognition should not be attempted for any generated candidate in category production, so no candidate should be excluded, and thus retrieval of encoded exemplars should be the highest when category production alone is examined. Recognition should be attempted for every generated candidate in category production+recognition, so all candidates should be subject to being excluded, and thus retrieval of encoded exemplars should be the lowest when category production followed by recognition is examined. If recognition is attempted for some but not all generated candidates in category-cued recall, retrieval should range between that for category production alone and that for category production followed by recognition (i.e., retrieval in category-cued recall should depend on how many generated candidates are submitted to a recognition check). In any case, the primary expectation is that retrieval in category-cued recall should not significantly exceed that in category production and should not be significantly exceeded by that in category production+recognition. Alternatively, recall might not be characterized by a generate–recognize strategy, which would be evidenced by significantly better category-cued recall than category production, significantly better category production+recognition than category-cued recall, or both (Guynn & McDaniel, 1999).

Experiment 1 focused on identifying the locus of an item-specific processing effect (if found) in recall, in conjunction with gathering evidence for the generate–recognize model. Experiment 2 focused on replicating and exploring the limits of the Experiment 1 results across a broader range of stimulus attributes (category exemplar frequencies). Experiment 3 extended the Experiment 2 results by including an encoding task designed to augment relational processing, to identify the locus of a relational processing effect (if found) in recall.

Experiment 1

In this experiment, we focused on illuminating the benefit to recall that arises from a particular type of item-specific processing. We built on a finding that Mulligan (2002; see also Kinoshita, 1989) reported using category production and

category-cued recall tests like those used herein (and in Guynn & McDaniel, 1999). After participants read words and generated words from letter-transposed anagrams (words with their first two letters reversed and underlined), Mulligan (and Kinoshita) found that more transposed anagrams than read words (an item-specific processing effect) were retrieved in category-cued recall, but not in category production. To account for this dissociation, Mulligan suggested that the tests were differentially sensitive to two types of conceptual information; category production was sensitive to relational information, whereas category cued recall was sensitive to both relational and item-specific information (Hamilton & Rajaram, 2001; Mulligan, 1996; Mulligan, Guyer & Beland, 1999; Smith & Hunt, 2000). Note that this interpretation of the joint influence of relational and item-specific information in recall is consistent with both the generate–recognize model and the view that recall involves direct access to memory traces, with rich episodic features promoting this access (Guynn & McDaniel, 1999; Weldon & Colston, 1995).

In this experiment, we attempted to characterize the recall process and use this as a foundation to understand the role of item-specific processing more fully. To accomplish this, we implemented the anagram-transposition generation task (Kinoshita, 1989; Mulligan, 2002) within the generate–recognize paradigm (Guynn & McDaniel, 1999; Jacoby & Hollingshead, 1990). We reasoned that if the pattern of performance implicated a generate–recognize strategy, the locus of the anagram-transposition generation effect (hereafter referred to as the *transposition effect*, to avoid confusing it with the first phase in the generate–recognize model) would isolate the item-specific influence. Specifically, if the transposition effect was not found in category production but was found on the recognition responses when separated from the production responses, the effect could be interpreted as an item-specific influence on the recognition component of recall. On the other hand, if the transposition effect was found in category production, a different interpretation would be implicated. The item-specific processing stimulated by the transposition task could be considered to act as episodic information additional to the category cue to constrain recall (or direct access) of the studied information (Jacoby et al., 2005; Jacoby et al., 1999). Furthermore, this pattern, in conjunction with significantly better category-cued recall than category production, or significantly better category production+recognition than category-cued recall, would imply that recall is not achieved via a generate–recognize strategy (Weldon & Colston, 1995).

Method

Design and participants

The design was a 2×2 mixed factorial, varying encoding task (read, transpose) within subjects and retrieval task (category-

cued recall, category production + recognition) between subjects.¹ Participants were 24 undergraduates enrolled in psychology courses at New Mexico State University, who participated either in partial fulfillment of a course requirement or for extra credit. Twelve participants performed each retrieval task.

Materials

The 96 medium frequency words arranged into the four randomized lists from Mulligan (2002) were used in this study.² One list consisted of six exemplars from each of the categories *fish*, *musical instruments*, *parts of a building*, *fruits*, *articles of furniture*, *four-footed animals*, *parts of the human body*, and *substances to flavor food*. Another list consisted of six exemplars from each of the categories *birds*, *trees*, *vegetables*, *articles of clothing*, *sports*, *colors*, *types of dances*, and *insects*. Half of the exemplars from each category were presented as intact words, and half were presented as letter-transposed anagrams (e.g., *urtile*)—words with their first two letters reversed and underlined. In an alternate form of each list, exemplars presented as intact words in the original version were presented as anagrams, and vice versa.

A sheet of two columns of 48 numbered lines was used for encoding. For category-cued recall, two sheets were used, with eight sets of six numbered lines per sheet, with the sets labeled *Category 1–Category 16*. For category production+recognition, the sheets were identical, except that next to each longer line was a shorter line where a check mark could be made. A sheet of 40 math problems was used for a distractor task.

Procedure

The experiment was advertised as a study on word perception and problem solving and required about 30 min to complete. For encoding, intact words and letter-transposed anagrams were presented one at a time for 8 s each in the center of the computer screen. Participants read each word or transposed each anagram by reversing its first two letters and then wrote the word on the encoding sheet. This encoding task for the 48 stimuli from one of the four lists was followed by a 3-min math problem distractor task.

For retrieval, the 16 category labels (8 each for studied and nonstudied categories) were presented one at a time in the center of the computer screen. For category-cued recall, participants were asked to write down as many exemplars as

¹ We used the category production results from the category production+recognition test for comparison with category-cued recall because our prior work (Guynn & McDaniel, 1999) indicated that a category production+recognition test and a separate category production test yielded equivalent category production results.

² Thanks to Neil Mulligan for providing these stimuli.

possible that were presented earlier, as either words or anagrams. If there was a category for which no exemplars were presented, participants were asked not to write anything down. For category production+recognition, participants were asked to write down the first six exemplars that came to mind. If an exemplar that was presented earlier came to mind, they should write it down, but they should not *try* (or *try not*) to think of exemplars that were presented earlier (but they should put a check mark next to each such exemplar written down).

Each category label remained on the computer screen until participants clicked on a *Ready* button to move to the next category. Participants could work on each category for as long as they liked, but they could not return to a category. One pseudorandom order of the category labels was used, with no more than two studied or two nonstudied categories tested consecutively. After the retrieval phase, participants in the category production+recognition condition were asked whether they had *tried* (or *tried not*) to think of exemplars that were presented earlier. The data from the one participant who answered *yes* were discarded, and a replacement was tested.

Results

The dependent measures for each treatment condition appear in Table 1. Each proportion reflects the mean number out of 24 encoded (read or transposed) exemplars that were retrieved in that particular condition. Thus, the category-cued recall means reflect the proportion of exemplars recalled on that test, whereas the remaining means are from the category production+recognition test: (1) The category production means reflect the proportion of exemplars generated before being submitted to a recognition check, (2) the category production+recognition means reflect the proportion of exemplars recognized after having been generated, and (3) the recognition means reflect recognition as a proportion of production. All reported effects were significant at a 0.05 level unless otherwise indicated.

To bolster our conclusions in cases of null effects, we also performed Bayesian analyses to compute the likelihood of failing to find an effect given the data. We used the method developed by Wagenmakers (2007; see also Masson, 2011), setting up the null hypothesis (effect absent) and alternative

hypothesis (effect present) as competing models, and using Bayes information criterion values to estimate a Bayes factor and compute the posterior probability for each hypothesis.

Item-specific effect in recall

To confirm that we obtained the transposition effect (an item-specific processing effect) in recall, we first conducted a one-way within-subjects analysis of variance (ANOVA) comparing retrieval of read and transposed exemplars on the category-cued recall test. There was a significant effect of encoding task, $F(1, 11) = 5.03$, $MSE = 0.005$, indicating that transposing produced significantly better recall than did reading.

Evidence for a generate–recognize strategy

The main objective of our study was to identify the locus of this item-specific processing effect in recall. To this end, we conducted several analyses to evaluate whether a generate–recognize strategy was evident. The generate–recognize model predicts that retrieval of encoded exemplars should decline as recognition plays an increasing role. Thus, category-cued recall should not significantly exceed category production, and category production+recognition should not significantly exceed category-cued recall. To evaluate this prediction, two ANOVAs were conducted, one comparing category production and category-cued recall, and one comparing category-cued recall and category production+recognition. The ANOVAs were 2×2 mixed factorials with encoding task as the within-subjects factor and retrieval task as the between-subjects factor.

The ANOVA involving category production and category-cued recall did not reveal a significant main effect of encoding task, $F(1, 22) = 1.72$, $MSE = 0.005$, indicating no general benefit of transposing over reading. Although neither the main effect of retrieval task ($F < 1$) nor the interaction, $F(1, 22) = 2.68$, $MSE = 0.005$, was significant, pairwise comparisons were conducted to test the predictions of the generate–recognize model. These comparisons confirmed a significant advantage for category production relative to category-cued recall for read words, $F(1, 22) = 5.88$, $MSE = 0.005$, and no significant difference for transposed words ($F < 1$). Moreover, the probability of the null (effect absent) model for transposed words, given the data, was 0.77, which constitutes positive evidence for this null effect (Raftery, 1995). The generate–recognize model predicts that category-cued recall should not significantly exceed category production, and thus both comparisons are consistent with this model.

The ANOVA involving category-cued recall and category production + recognition did reveal a significant main effect of encoding task, $F(1, 22) = 4.34$, $MSE = 0.005$, indicating a general benefit of transposing over reading in the conditions

Table 1 Mean proportions of retrieval (and recognition as a proportion of production) for Experiment 1

	Read Words	Transposed Words
Category production	0.25	0.24
Category-cued recall	0.18	0.25
Category production+recognition	0.20	0.23
Recognition	0.78	0.94

in which recognition was expected to play a role. Although neither the main effect of retrieval task nor the interaction was significant (both F s < 1), pairwise comparisons were again conducted to test the predictions of the generate–recognize model. These comparisons indicated no significant difference between category-cued recall and category production + recognition for either read or transposed words (both F s < 1). The probability of the null (effect absent) model for both read and transposed words, given the data, was 0.75, again constituting positive evidence for these null effects (Raftery, 1995). The generate–recognize model predicts that category production + recognition should not significantly exceed category-cued recall, and thus both comparisons are consistent with this model.

Item-specific effect in category production and recognition

The prior analyses were consistent with the idea that a generate–recognize strategy guided category-cued recall. Thus, we next examined the category production and recognition performances separately, to assess the transposition effect (the item-specific processing effect) on the separate processes of generation and recognition, respectively.

For category production, there was no hint of an item-specific processing effect; indeed, there was a nominal benefit of reading over transposing, and the one-way ANOVA with encoding task as the within-subjects factor revealed no significant difference ($F < 1$). For recognition, a one-way ANOVA with encoding task as the within-subjects factor revealed a marginally significant item-specific processing effect, with a recognition benefit for transposing over reading, $F(1, 11) = 3.22$, $MSE = 0.051$, $p = 0.10$.

Discussion

We obtained a significant transposition effect in recall, thereby replicating Mulligan (2002; and Kinoshita, 1989) in finding this item-specific processing effect in category-cued recall but not category production. Importantly for present purposes, the results were also consistent with the generate–recognize model, in that category-cued recall was not significantly better than category production, and category production + recognition was not significantly better than category-cued recall. Furthermore, evaluating recall in terms of the separate generation and recognition components provided evidence that the item-specific processing evoked by the transposition task influenced the recognition phase and not the generation phase of recall. Specifically, there was no hint of a transposition effect on the production phase of category production+recognition, but there was an effect on recognition as a proportion of production, suggesting that item-specific processing impacts recall at the recognition phase in a generate–recognize model.

Experiment 2

The goals of Experiment 2 were primarily (1) to reinforce the evidence from Experiment 1 for the generate–recognize model and for the locus of the item-specific processing effect at the recognition phase of recall (including obtaining a significant transposition effect on recognition as a proportion of production) and also (2) to explore possible boundary conditions under which the model can account for recall (as suggested in Guynn & McDaniel, 1999). In partial service of the first goal, we added a test of category production that was not followed by a recognition decision, to isolate better the generation phase. In service of the second goal, we varied the category exemplar frequency of the stimuli, such that some were high-frequency and some were low-frequency exemplars of their category. If a generate–recognize strategy generally guides recall, the patterns obtained in Experiment 1 should again emerge, for both types of stimuli.

A more complex pattern is possible, however, given Guynn and McDaniel's (1999, Experiment 3) finding that although a generate–recognize strategy guided recall of high-frequency exemplars, low-frequency exemplars were recalled better than would be expected with this strategy. This finding suggests that the theoretical explanation might differ for the different stimuli. One possibility is that recall is not invariably guided by a generate–recognize strategy (e.g., for low-frequency exemplars), and for present purposes, the implication is that an anagram-transposition effect will not inevitably arise in recall (e.g., for low-frequency exemplars), if the locus of this item-specific processing effect is at the recognition phase of recall in a generate–recognize model. On this view, we would expect that a generate–recognize strategy would be used for high- but not low-frequency exemplars, and a transposition effect would occur in recall for high- but not low-frequency exemplars. This pattern might not necessarily imply that a generate strategy was not used for low-frequency exemplars, however, and we mention another possible interpretation following the next experiment.

Method

Design and participants

The design was a $2 \times 2 \times 3$ mixed factorial, varying encoding task (read, transpose) and category exemplar frequency (high, low) within subjects and retrieval task (category production, category cued recall, category production + recognition) between subjects. Participants were 48 undergraduates enrolled in psychology courses at New Mexico State University who participated either in partial fulfillment of a course requirement or for extra credit. Sixteen participants performed each retrieval task.

Materials

Ninety-six exemplars were selected from the Battig and Montague (1969) category norms (and are reproduced in Guynn & McDaniel, 1999). Four high-frequency exemplars (average category frequency of 194.48) and four low-frequency exemplars (average category frequency of 16.60) were selected from each of 12 categories, and four lists were constructed from these exemplars. One list consisted of four high- and four low-frequency exemplars from each of the categories *a four-footed animal, an article of clothing, a ship, a sport, a fruit, and an article of furniture*. Another list consisted of four high- and four low-frequency exemplars from each of the categories *a kitchen utensil, a bird, an occupation, a musical instrument, a part of the human body, and a vegetable*. Half of the high- and low-frequency exemplars from each category were presented as intact words, and half were presented as letter-transposed anagrams. In an alternate form of each list, exemplars presented as intact words in the original version were presented as anagrams, and vice versa. The exemplars for each list were presented in one pseudorandom order, with no exemplars from a given category presented consecutively and neither encoding task performed more than twice consecutively.

The encoding and retrieval sheets and math problems were the same as those in Experiment 1, except that there were six sets of eight numbered lines per retrieval sheet. The retrieval sheets for category production were the same as those for category-cued recall.

Procedure

The encoding and retrieval activities were the same as those in Experiment 1, except for the addition of the category production retrieval task. For this test, participants were asked to write down the first eight exemplars that came to mind for each category label. Following Experiment 1 and Guynn and McDaniel (1999), participants were also instructed that if exemplars presented earlier came to mind, they should write them down, but they should not *try* (or *try not*) to think of exemplars presented earlier. Two category production participants and two category production + recognition participants reported that they had either *tried* or *tried not* to think of these exemplars, and so their data were discarded and replacements were tested.

Results

The dependent measures for each treatment condition appear in Table 2. The values are similar to those reported for Experiment 1, except that the proportions reflect the mean number retrieved out of 12 encoded stimuli. The values from the category production test appear in row 1, and the

production phase values from the category production+recognition test appear in row 4. The analyses conducted were similar to those for Experiment 1, but with the additional within-subjects factor of category exemplar frequency.

Item-specific effect in recall

Considering just cued recall, there was a significant benefit of transposing over reading, $F(1, 15) = 12.00$, $MSE = 0.009$. There was also a significant benefit for high-frequency over low-frequency exemplars, $F(1, 15) = 21.74$, $MSE = 0.015$, and a trend toward an interaction, $F(1, 15) = 3.65$, $MSE = 0.017$, $p < 0.08$, reflecting a greater transposition effect for high-frequency exemplars. As was anticipated in the introduction to this experiment, pairwise comparisons confirmed a significant transposition effect for high-frequency exemplars, $F(1, 15) = 10.59$, $MSE = 0.017$, but not for low-frequency exemplars ($F < 1$; null model probability of 0.78). The following analyses help to inform this effect.

Evidence for a generate–recognize strategy

To evaluate whether a generate–recognize strategy was evident and to identify the locus of the item-specific processing effect for high-frequency exemplars, we conducted analyses similar to those for Experiment 1, but with the additional within-subjects factor of category exemplar frequency.

For the $2 \times 2 \times 2$ mixed ANOVA involving category production and category-cued recall (rows 1 and 2), there was a significant exemplar frequency \times retrieval task interaction, $F(1, 30) = 35.19$, $MSE = 0.013$. Pairwise comparisons indicated that this interaction reflected that category-cued recall ($M = 0.31$) was significantly worse than category production ($M = 0.46$) for high-frequency exemplars, $F(1, 30) = 27.69$, $MSE = 0.013$, whereas the pattern reversed for low-frequency exemplars ($M_s = 0.17$ and 0.08 for category-cued recall and category production, respectively), $F(1, 30) = 9.97$, $MSE = 0.013$. This pattern is consistent with the generate–recognize model for high-frequency but not low-frequency exemplars.

There was also a significant encoding task \times retrieval task interaction, $F(1, 30) = 7.28$, $MSE = 0.010$, reflecting a benefit for transposed words in category-cued recall, but not in category production, and a trend toward a three-way interaction, $F(1, 30) = 3.08$, $MSE = 0.014$, $p < 0.09$, such that the transposition effect in category-cued recall was limited to high-frequency exemplars. The main effects of encoding task and exemplar frequency were significant, $F(1, 30) = 4.40$, $MSE = 0.010$, and $F(1, 30) = 166.30$, $MSE = 0.013$, respectively, reflecting benefits for transposed words and high-frequency exemplars.

For the $2 \times 2 \times 2$ mixed ANOVA involving category-cued recall and category production+recognition (rows 2 and 3), there were significant main effects of encoding task, exemplar

Table 2 Mean proportions of retrieval (and recognition as a proportion of production) for Experiment 2

	Read Words		Transposed Words	
	High Frequency	Low Frequency	High Frequency	Low Frequency
Category production	0.47	0.08	0.45	0.08
Category-cued recall	0.24	0.16	0.39	0.18
Category production+recognition	0.17	0.06	0.33	0.05
Category production	0.43	0.07	0.46	0.05
Recognition	0.51	0.80	0.78	1.00

frequency, and retrieval task, $F(1, 30) = 26.44$, $MSE = 0.007$, $F(1, 30) = 52.14$, $MSE = 0.017$, and $F(1, 30) = 4.33$, $MSE = 0.060$, respectively, reflecting benefits for transposed words, high-frequency exemplars, and category-cued recall. Consistent with the generate–recognize model, the latter two factors did not interact, $F(1, 30) = 1.38$, $MSE = 0.017$. Pairwise comparisons confirmed that category-cued recall was better than category production+recognition for both high-frequency exemplars ($M = 0.31$ and 0.25 , respectively; $p < 0.08$), $F(1, 30) = 3.39$, $MSE = 0.017$, and low-frequency exemplars ($M = 0.17$ and 0.05 , respectively), $F(1, 30) = 13.55$, $MSE = 0.017$, which is consistent with the generate–recognize model for both types of stimuli. Finally, there was a significant encoding task \times exemplar frequency interaction, $F(1, 30) = 12.24$, $MSE = 0.014$, reflecting a greater transposition effect for high-frequency exemplars.

The prior analyses suggest that at least for high-frequency exemplars, category-cued recall involves a recognition process that category production does not and, moreover, that the item-specific processing effect in category-cued recall arises from this recognition process. To gather converging evidence that this item-specific processing effect arises at the recognition phase in a generate–recognize model, as in Experiment 1, we analyzed this effect separately for the category production and recognition components of the category production + recognition task.

Item-specific effect in category production and recognition

For the category production analysis, we included the values from both the category production and category production+recognition tasks (rows 1 and 4), to gain maximal information on category production. A $2 \times 2 \times 2$ mixed ANOVA with encoding task and category exemplar frequency as the within-subjects factors and retrieval task as the between-subjects factor revealed no significant main effects or interactions involving retrieval task (largest $F = 1.41$), indicating that category production was similar for the two retrieval conditions. Importantly, there was no significant transposition effect (item-specific processing effect) in category production ($F < 1$).

For the recognition analysis, although there were just five observations,³ a 2×2 ANOVA with encoding task and category exemplar frequency as the within-subjects factors revealed a significant benefit of transposing over reading, $F(1, 4) = 37.02$, $MSE = 0.008$, and a marginally significant benefit for low-frequency over high-frequency exemplars, $F(1, 4) = 5.09$, $MSE = 0.065$, $p < 0.09$. Thus, there was a significant transposition effect (item-specific processing effect) on the recognition responses when separated from the production responses.

Discussion

Using different materials from Experiment 1, we replicated the support for the generate–recognize model. Specifically, for high-frequency exemplars, category-cued recall was not significantly better than category production, and category production + recognition was not significantly better than category-cued recall. We also replicated the significant transposition effect in category-cued recall for high-frequency exemplars. Most important, following from a generate–recognize analysis, we again were able to localize this item-specific processing effect to the recognition phase of recall. Specifically, there was no significant transposition effect on either the category production task or the category production phase of the category production+recognition task, but there was a significant effect on the recognition phase of this task.

In contrast, the results regarding a generate–recognize strategy for low-frequency exemplars were less clear, as category-cued recall was significantly better than category production for these stimuli. Also, there was not a significant transposition effect in category-cued recall for low-frequency exemplars, in contrast to previous demonstrations of anagram-transposition generation effects (Flory & Pring, 1995; Kinoshita, 1989; Mulligan, 2002; Roediger & McDermott, 1993; but see Nairne, Pusey & Widner, 1985, Experiment 3).

³ There were 16 participants in the category production+recognition condition, but just 5 participants generated all four types of word at test.

One possibility is that a generate–recognize strategy was not used for low-frequency exemplars, because a strategy involving generation followed by recognition would not be useful for their recall (i.e., a more direct retrieval strategy would be more useful; see Guynn & McDaniel, 1999; Weldon & Colston, 1995), because they would be at a disadvantage in generation to category label cues. And without extensive involvement of a recognition process, the opportunity for the transposition task to influence performance would be attenuated. By this interpretation, were a recognition process forced for low-frequency exemplars, as it was in the category production + recognition task, a transposition effect should arise for both high- and low-frequency exemplars, which is the pattern that obtained when recognition was isolated from production.

Another possibility is that a generate–recognize strategy was used for low-frequency exemplars but the functional generation cues differed between category production and category-cued recall, which conferred an advantage to low-frequency exemplars in cued recall.⁴ This is inconsistent with an assumption of the Jacoby and Hollingshead (1990) revised generate–recognize model. But they tested their model with single-completion cues, whereas the present study featured multiple-completion cues, where the temptation to generate exemplars until enough are recognized might be greater. On this idea, in category production, the category label cue and instructions constrained generation to exemplars from the category, but in category-cued recall, the cue and instructions could have constrained generation to exemplars from the category that were seen at study. Such a difference would have little consequence for high-frequency exemplars, which are retrieved at high levels with a category label cue. But such a difference could have consequences for low-frequency exemplars, which are not retrieved at particularly high levels with a category label cue. For these exemplars, the joint cue for category-cued recall could serve as a better retrieval cue and, thereby, support higher levels of retrieval. Our study does not permit us to distinguish between these possibilities for low-frequency exemplars, but finding different patterns of performance across retrieval tasks for the high- and low-frequency exemplars leaves open the possibility of different retrieval strategies being effective for the different stimuli.

Experiment 3

Despite the uncertain theoretical interpretation for low-frequency exemplars, the results for high-frequency exemplars were clear. Accordingly, the primary goal of Experiment 3 was to expand our focus for high-frequency exemplars to include localizing the effect of relational processing in the generate–recognize model (we also included low-frequency exemplars

to try to get some clarity on that theoretical situation). To this end, we added a category-sorting task to augment relational processing at encoding. In this condition, paralleling the anagram-transposition task, half of the exemplars were read, and half were read and sorted into taxonomic categories. We assumed that category-sorting would stimulate processing of relational information (among category exemplars and/or between the category label and the exemplars), on the basis of prior work showing that category-sorting increases both free recall and clustering in free recall (e.g., Einstein & Hunt, 1980; McDaniel et al., 1988). On the view that relational processing enhances a generation process, we expected that category-sorting, unlike anagram-transposition, would benefit category production. Furthermore, assuming that the generation process in category production is the same as that in category-cued recall, category-sorting should provide a concomitant benefit to category-cued recall. Such a result would inform and extend Rappold and Hashtroudi's (1991) finding that relational processing enhanced performance on both explicit and implicit conceptual tests.

With regard to recognition, the most straightforward expectation is that category-sorting will not have an effect (Einstein & Hunt, 1980). However, because recognition is assessed after generation and because a recognition decision might be based in part on the fluency (Jacoby & Dallas, 1981; Jacoby, Kelley & Dywan, 1989) of generation, it is possible that recognition could show some effect of category-sorting in the present paradigm. We also dropped the category production test and used the results from the category production phase of the category production+recognition test (as in Experiment 1), because Experiment 2 and our prior work (Guynn & McDaniel, 1999) showed that these conditions yielded equivalent category production results.

Method

Design and participants

The design was a $2 \times 2 \times 2 \times 2$ mixed factorial, varying encoding task (read, elaborate) and category exemplar frequency (high, low) within subjects and elaboration type (transpose, sort) and retrieval task (category-cued recall, category production+recognition) between subjects. Participants were 48 undergraduates enrolled in psychology courses at New Mexico State University, who participated for partial fulfillment of a course requirement, extra credit, or a \$5 monetary payment. Twelve participants performed each combination of elaboration type and retrieval task.

Materials

In the anagram-transposition condition, the high- and low-frequency exemplars and randomized lists were the same as

⁴ Thanks to Reed Hunt for suggesting this possible interpretation.

those in Experiment 2. In the category-sorting condition, the same high-frequency and low-frequency exemplars were used, but all were presented as intact words and blocked by category. Half of the high-frequency and low-frequency exemplars were presented with a question (“Category ?”) indicating that the word was to be written under its category label. The other exemplars were to be written in a list. Thus, the relational processing manipulation involved participants’ attention being drawn to the category label and each exemplar as a member of that category. Because of the blocked presentation, we expected some relational processing for the category exemplars to be read, but we chose not to present them in a random order, because we were concerned that this would produce differential item-specific processing for the items to be read and sorted, and we wanted a manipulation of only relational processing.

In the anagram-transposition condition, the encoding sheet was the same as in Experiments 1 and 2. In the category-sorting condition, two sheets were used for encoding: a sheet with a column of 24 numbered lines for the words to be read and a sheet with the six category labels and four numbered lines under each category label for the words to be sorted. The retrieval sheets and math problems were the same as those in Experiment 2.

Procedure

For anagram-transposition, the encoding procedure was the same as that in Experiments 1 and 2. For category-sorting, all of the words were presented intact and were to be written down in a list if *Category ?* did not appear on the screen and under their appropriate category label if *Category ?* did appear on the screen. The data for three anagram-transposition participants and one category-sorting participant were discarded for not following the category production+recognition instructions, and replacements were tested.

Results

The dependent measures for each treatment condition appear in Table 3. The values are similar to those reported for Experiment 2, in that the proportions reflect the mean number retrieved out of 12 encoded stimuli.

Item-specific and relational effects in recall

We conducted a 2 (encoding task) \times 2 (exemplar frequency) \times 2 (elaboration type) mixed ANOVA to evaluate the anagram-transposition and category-sorting effects in cued recall. All main effects were significant. Elaborating (transposing or sorting) produced better recall than did reading, $F(1, 22) = 36.44$, $MSE = 0.017$, high-frequency exemplars were recalled better than were low-frequency exemplars, $F(1, 22) = 28.23$,

$MSE = 0.019$, and category-sorting produced better recall than did anagram-transposition, $F(1, 22) = 5.70$, $MSE = 0.075$. The critical comparisons evaluated the benefit of transposition ($M = 0.33$) over reading ($M = 0.19$) and the benefit of sorting ($M = 0.48$) over reading ($M = 0.30$). Both comparisons were significant, $F(1, 22) = 13.84$, $MSE = 0.017$, and $F(1, 22) = 22.87$, $MSE = 0.017$, respectively, indicating that both encoding tasks produced a benefit to recall. And similar to Experiment 2, the anagram-transposition effect was significant for high-frequency exemplars, $F(1, 22) = 10.84$, $MSE = 0.016$, and less robust for low-frequency exemplars, $F(1, 22) = 3.75$, $MSE = 0.016$, $p < 0.07$.

Evidence for a generate–recognize strategy

The $2 \times 2 \times 2 \times 2$ mixed ANOVA involving category production and category-cued recall revealed a significant exemplar frequency \times retrieval task interaction, $F(1, 44) = 19.12$, $MSE = 0.019$, reflecting a smaller benefit for high-frequency exemplars in category-cued recall. Critically, the interaction also indicated that fewer high-frequency exemplars were retrieved in category-cued recall than in category production. Pairwise comparisons ($MSE = 0.018$) confirmed that for high-frequency exemplars, category-cued recall was either significantly worse [the read exemplars in the transposition and sorting conditions; $F(1, 44) = 13.33$ and $F(1, 44) = 5.63$, respectively] or not significantly different [the elaborated exemplars in the transposition and sorting conditions; $F(1, 44) = 1.20$ and $F(1, 44) = 0.03$, and null model probabilities of 0.75 and 0.78, respectively] than category production, whereas for low-frequency exemplars, category-cued recall was either significantly better [the read and sorted exemplars in the sorting condition; $F(1, 44) = 5.63$ and $F(1, 44) = 8.53$, respectively] or not significantly different [the read and transposed exemplars in the transposition condition; $F(1, 44) = 0.00$ and $F(1, 44) = 0.30$, and null model probabilities of 0.78 and 0.77, respectively] than category production. As in Experiment 2, this pattern is consistent with the generate–recognize model for high-frequency but not low-frequency exemplars. There was also a significant encoding task \times retrieval task interaction, $F(1, 44) = 5.60$, $MSE = 0.017$, reflecting a greater benefit of elaborating in category-cued recall. The main effects of encoding task, elaboration type, and exemplar frequency were significant, reflecting benefits for elaborating, $F(1, 44) = 39.22$, $MSE = 0.017$, category-sorting, $F(1, 44) = 6.51$, $MSE = 0.052$, and high-frequency exemplars, $F(1, 44) = 141.15$, $MSE = 0.019$.

The comparable $2 \times 2 \times 2 \times 2$ mixed ANOVA involving category-cued recall and category production+recognition revealed significant main effects of encoding task, exemplar frequency, elaboration type, and retrieval task. Elaborating produced better retrieval than did reading, $F(1, 44) = 54.86$, $MSE = 0.018$, high-frequency exemplars were retrieved better

Table 3 Mean proportions of retrieval (and recognition as a proportion of production) for Experiment 3

	Read Words		Elaborated Words	
	High Frequency	Low Frequency	High Frequency	Low Frequency
	Anagram-Transposition			
Category production	0.46	0.12	0.49	0.19
Category-cued recall	0.26	0.12	0.43	0.22
Category production+recognition	0.23	0.10	0.39	0.16
Recognition	0.60	0.69	0.80	0.75
	Category-Sorting			
Category production	0.47	0.13	0.55	0.24
Category-cued recall	0.34	0.26	0.56	0.40
Category production+recognition	0.24	0.10	0.40	0.22
Recognition	0.50	0.61	0.75	0.81

than were low-frequency exemplars, $F(1, 44) = 87.98$, $MSE = 0.014$, sorting produced better retrieval than did transposition, $F(1, 44) = 4.44$, $MSE = 0.067$, and category cued recall was better than was category production+recognition, $F(1, 44) = 6.14$, $MSE = 0.067$. Pairwise comparisons ($MSE = 0.016$) confirmed that for both high- and low-frequency exemplars, category-cued recall was either marginally or significantly better [all four sorting conditions; high-frequency read exemplars, $F(1, 44) = 3.75$, $p < 0.06$; low-frequency read exemplars, $F(1, 44) = 9.60$; high-frequency sorted exemplars, $F(1, 44) = 9.60$; low-frequency sorted exemplars, $F(1, 44) = 12.15$] or not significantly different [all four transposition conditions; high-frequency read exemplars, $F(1, 44) = 0.34$; low-frequency read exemplars, $F(1, 44) = 0.15$; high-frequency transposed exemplars, $F(1, 44) = 0.60$; low-frequency transposed exemplars, $F(1, 44) = 1.35$; null model probabilities of 0.77, 0.77, 0.76, and 0.74, respectively] than category production+recognition. As in Experiment 2, this pattern is consistent with the generate–recognize model for both types of stimuli. Finally, there was a marginally significant ($p < 0.06$) encoding task \times exemplar frequency interaction, $F(1, 44) = 3.75$, $MSE = 0.016$, reflecting a greater benefit of elaborating for high-frequency exemplars.

Item-specific and relational effects in category production and recognition

To aid in localizing the effects of item-specific and relational processing to the phases of cued recall, we also analyzed whether there was a transposition effect or a sorting effect on category production. A $2 \times 2 \times 2$ mixed ANOVA with encoding task and exemplar frequency as the within-subjects factors and elaboration type as the between-subjects factor revealed a significant benefit for elaborated exemplars, $F(1, 22) = 7.76$, $MSE = 0.016$, and high-

frequency exemplars, $F(1, 22) = 131.88$, $MSE = 0.019$. The critical comparisons confirmed a significant benefit of sorting ($M = 0.40$ and $M = 0.30$, for the sorted and read exemplars, respectively), $F(1, 22) = 6.93$, $MSE = 0.016$, and no significant effect of transposition [$M = 0.34$ and $M = 0.29$, for the transposed and read exemplars, respectively; $F(1, 22) = 1.70$, $MSE = 0.016$; null model probability of 0.69]. Thus, there was a significant benefit to generation from category-sorting but not from anagram-transposition.

We also analyzed whether there was a transposition effect or a sorting effect on recognition as a proportion of production. Although there were just 17 observations,⁵ a $2 \times 2 \times 2$ mixed ANOVA with encoding task and exemplar frequency as the within-subjects factors and elaboration type as the between-subjects factor revealed a significant benefit for elaborated exemplars, $F(1, 15) = 5.13$, $MSE = 0.100$. To confirm the equivalent benefit to recognition from both types of elaboration, we assessed the elaboration effect separately for the transposition ($M = 0.48$ for read exemplars and $M = 0.79$ for transposed exemplars) and sorting ($M = 0.56$ for read exemplars and $M = 0.76$ for sorted exemplars) conditions, and both effects were significant, $F(1, 11) = 12.91$, $MSE = 0.042$, and $F(1, 11) = 14.28$, $MSE = 0.016$, respectively.

Discussion

We replicated the significant transposition effect in cued recall for high-frequency exemplars. And the results following anagram-transposition were again consistent with the generate–recognize model for these exemplars. In particular,

⁵ There were 24 participants in the category production+recognition condition, but just 8 anagram-transposition participants and 9 category-sorting participants generated all four types of word at test.

category-cued recall was not significantly better than category production, and category production + recognition was not significantly better than category-cued recall. Finally, the anagram-transposition effect in the recognition phase, but not the generation phase, of category production + recognition indicated that anagram-transposition impacts recall at the recognition phase in a generate–recognize model.

The new result was that category-sorting also significantly improved cued recall. But in contrast to anagram-transposition, category-sorting also significantly improved category production. Moreover, the results in the category-sorting condition were also consistent with the generate–recognize model for high-frequency exemplars, in that category-cued recall was not significantly better than category production and category production+recognition was not significantly better than category-cued recall. Thus, we can localize the relational processing elicited by category-sorting to the generation phase of recall in a generate–recognize model. Thus, continuing from the perspective that category-sorting enhances relational processing and anagram-transposition enhances item-specific processing, these patterns suggest that relational processing influences the generation phase of recall and item-specific processing selectively influences the recognition phase of recall in a generate–recognize model.

A result that potentially complicates the above interpretation is the significant category-sorting effect on recognition in the category production + recognition task. One possibility is that category-sorting evoked some item-specific processing. However, we believe it unlikely that any additional encoding of item-specific information was evoked by category-sorting, because we presented the category exemplars in a blocked order to make the relational information salient. Instead, it seems more likely that the recognition phase that followed the generation phase was affected by the fluency of generating category exemplars following category-sorting. Further evidence is the fact that there was not the usual significant benefit for low-frequency exemplars in recognition. Note that such an effect of category-sorting on recognition would not be expected in a more typical recognition paradigm in which recognition is not preceded by generation of these items. Nor would such an effect be expected in the recognition phase of category-cued recall, because the revised generate–recognize model already incorporates the idea that fluent generation can obviate an independent recognition process (Jacoby & Hollingshead, 1990). Finally, note that the effect of anagram-transposition on recognition could not be attributed to the fluency of generation, because there is no evidence that anagram-transposition affects generation, as there was no significant item-specific processing effect on category production in any experiment.

The results for low-frequency exemplars continue to be puzzling. In this experiment, the results for low-frequency exemplars were actually consistent with the generate–recognize

model following anagram-transposition but inconsistent with the model following category-sorting.

General discussion

The generate–recognize model and the relational–item-specific distinction are separate and influential approaches to explaining recall. We combined them in the present study. Our goal was to localize the influence of relational and item-specific information in the generate–recognize model. To the extent that this model can account for recall, the effect of these different types of information can be localized to the different phases of recall.

We varied item-specific information (Experiments 1–3) by asking participants to read words and generate words from letter-transposed anagrams. This anagram-transposition task improved category-cued recall (as compared with reading) but did not impact category production. Moreover, the pattern of performance across category production, category-cued recall, and category production + recognition conformed to the predictions of the generate–recognize model for high-frequency (Experiments 2 and 3) and medium-frequency (Experiment 1) category exemplars. Thus, we were able to localize the item-specific processing associated with anagram-transposition to a phase in the generate–recognize model for these exemplars. The recognition phase, but not the generation phase, was affected, implying that the influence on recognition was responsible for the benefit to category-cued recall.

We varied relational information (Experiment 3) by asking participants to read words and sort words into taxonomic categories. This category-sorting task improved both category-cued recall and category production (as compared with reading). Moreover, the pattern of performance across category production, category-cued recall, and category production + recognition conformed to the predictions of the generate–recognize model for high-frequency category exemplars. Thus, we were able to localize the relational processing associated with category-sorting to a phase in the generate–recognize model for these exemplars. Unlike with anagram-transposition, with category-sorting the generation phase was affected, implying that the influence on generation was responsible for the benefit to category-cued recall.

In contrast to the results for high-frequency exemplars, it is less clear what strategy is supporting retrieval of low-frequency exemplars. The fact that category-cued recall was better than category production for low-frequency exemplars in some conditions might indicate that participants did not use a generate–recognize strategy to recall these exemplars and, instead, used a more direct retrieval strategy. Alternatively, it might be the case that participants did use a generate–recognize strategy but that the nature or extent of the generation process differed across the two retrieval tasks. In particular, because

each retrieval cue (category label) permitted more than just a single completion, in contrast to the Jacoby and Hollingshead (1990) single-completion word stem cues, perhaps participants generated exemplars until they felt that they had recognized and, thus, recalled a suitable number of them. Although we cannot decide between these possibilities in the present work, we note that both Hamann (1994) and Halamish (2009) used single completion conceptual cues and found, as in the present study and Gynn and McDaniel (1999), that low-frequency exemplars were recalled better than would be expected with a generate–recognize strategy. At this point, we cannot rule out the possibility that different processes support retrieval for high- and low-frequency exemplars. A modeling approach such as that advocated by Brainerd and Reyna (2010) could help to resolve this issue in future work.

Our results (at least for the medium- and high-frequency exemplars, for which the generate–recognize model did apply) localizing relational and item-specific processing to the different phases of recall can be interpreted in the context of current views on the mnemonic benefit of distinctiveness (e.g., Burns, 2006; Mulligan, 2006; Smith, 2006). On these views, the mnemonic benefit of distinctiveness is thought to arise either as a consequence of item-specific processing alone or in the context of relational processing. Our results are consistent with these views, as our “distinctiveness” manipulation influenced only the measures that were thought to reflect item-specific processing (recognition and category-cued recall) and not the measure that was thought to reflect purely relational processing (category production). Moreover, the item-specific processing effect was smaller for low-frequency exemplars, which presumably afforded less relational processing, than for high-frequency exemplars.

Also, with our localizing relational processing to the generation phase and item-specific processing to the recognition phase in the generate–recognize model, any mnemonic benefits of distinctiveness would be viewed as arising at the recognition phase of recall. An interesting prediction from this view is that mnemonic benefits of distinctiveness might not then be expected when the generation phase proceeds fluently and, thus, obviates the need for a recognition check.

Alternative interpretation of the anagram-transposition generation effect

In the context of the generate–recognize model, we offer an alternative interpretation of the dissociation between category production and category-cued recall as a function of the anagram-transposition task. Mulligan (2002) suggested that the dissociation implied that category production and category-cued recall, both conceptual tests, were differentially sensitive to relational versus item-specific processing

(Hunt & McDaniel, 1993). Our results are not inconsistent with this interpretation, but in the spirit of stimulating new work on this topic, we suggest another possibility.

Specifically, to the extent that anagram-transposition facilitates the encoding of nonsemantic, perceptual information (Flory & Pring, 1995; Kinoshita, 1989; Roediger & McDermott, 1993), an alternative interpretation is that an anagram-transposition effect in category-cued recall implies that category-cued recall is not a process-pure conceptual test and, as such, is susceptible to perceptual influences. How could recall that is cued and guided by category information, which is clearly semantic, involve perceptual processing? To the extent that category-cued recall is accomplished by a generate–recognize strategy, perceptual influences could affect category-cued recall, as evidenced by the fact that the anagram-transposition task affected the recognition component of category-cued recall. In at least some prominent models of recognition memory, a recognition decision can be based on a feeling of familiarity for an item, and the feeling of familiarity reflects the integration of the item’s perceptual features (Atkinson & Juola, 1974; Mandler, 1980). Thus, an alternative interpretation of the dissociation between category production and category-cued recall in terms of the effects of anagram-transposition is that category-cued recall, but not category production, can rely on perceptual information (i.e., at the recognition phase of recall).

Along these lines, other reported dissociations may also reflect the fact that some prototypical conceptual explicit tests involve perceptual processes. For example, the picture superiority effect is obtained on the explicit tests of category-cued recall and associate-cued recall, but not on the parallel implicit tests of category production and word association (Weldon & Coyote, 1996). Weldon and Coyote invoked Nelson’s (1979) explanation of the picture superiority effect to explain their results, according to which the picture superiority effect arises because the sensory codes of pictures are more distinctive than are the sensory codes of words. On explicit conceptual tests, this distinctiveness is useful for discriminating among items (e.g., studied vs. nonstudied), but on implicit conceptual tests, it is not necessary to discriminate among studied versus nonstudied items, and so the more distinctive sensory codes of pictures are not useful.

A related interpretation could be offered for other dissociations between explicit and implicit memory tests. For example, the concreteness effect is obtained in free recall and on the explicit general knowledge test, but not on the parallel implicit general knowledge test (Hamilton & Rajaram, 2001). Concrete words are believed to induce the encoding of perceptual information about the word’s referent (Clark & Paivio, 1987; Marschark & Hunt, 1989), and the concreteness effect may arise because the representations of concrete words are more perceptually distinctive than are the representations of abstract words (Marschark & Hunt, 1989).

Similarly, the primacy effect is obtained on the explicit test of category-cued recall, but not on the parallel implicit test of category production (Mulligan & Stone, 1999). Several models of the primacy effect suggest that the effect is related to the temporal distinctiveness of items presented in initial list positions (Glenberg, 1987; Glenberg & Swanson, 1986; Knoedler et al., 1999; Nairne, Neath, Serra & Byun, 1997), with temporal information presumably involving perceptually-based components.

Conclusion

Recall is probably best characterized as an opportunistic test in which individuals adopt whatever information and strategies may be useful for retrieval. Our results indicate that individuals adopt a generate–recognize strategy that is useful when there are high-frequency exemplars in the retrieval set. But this strategy may not be useful for low-frequency exemplars, and thus individuals may also adopt a more direct retrieval strategy (Weldon & Colston, 1995) for these exemplars. This is somewhat speculative at this point for the reason indicated earlier, but the fact that, in some cases, low-frequency exemplars are recalled better than would be expected with a generate–recognize strategy leaves this open as a possibility for future work.

References

- Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, 79, 97–123.
- Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz & R. C. Atkinson (Eds.), *Contemporary developments in mathematical psychology: I. Learning, memory and thinking* (pp. 243–293). Oxford: W. H. Freeman.
- Bahrick, H. P. (1970). Two-phase model for prompted recall. *Psychological Review*, 77, 215–222.
- Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, 80, 1–46.
- Bellezza, F. S. (2003). Evaluation of six multinomial models of conscious and unconscious processes with the recall–recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 779–796.
- Brainerd, C. J., & Reyna, V. F. (2010). Recollective and nonrecollective recall. *Journal of Memory and Language*, 63, 425–445.
- Burns, D. J. (2006). Assessing distinctiveness: Measures of item-specific and relational processing. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory*. New York: Oxford University Press.
- Clark, J. M., & Paivio, A. P. (1987). A dual coding perspective on encoding processes. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 5–32). New York: Springer-Verlag.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268–294.
- Einstein, G. O., & Hunt, R. R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 588–598.
- Flory, P., & Pring, L. (1995). The effects of data-driven and conceptually driven generation of study items on direct and indirect measures of memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 48A, 153–165.
- Glenberg, A. M. (1987). Temporal context and recency. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and learning: The Ebbinghaus Centennial Conference* (pp. 173–190). Hillsdale: Erlbaum.
- Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 3–15.
- Guynn, M. J., & McDaniel, M. A. (1999). Generate-sometimes recognize, sometimes not. *Journal of Memory and Language*, 41, 398–415.
- Halamish, V. (2009). *Quality control in memory retrieval and reporting: “Front-end” and “back-end” components*. Israel: Unpublished doctoral dissertation, University of Haifa.
- Hamann, S. B. (1994). Comparing implicit and explicit performance in perceptual and conceptual tasks: Applicability of a generation/recognition model (Doctoral dissertation, University of Toronto, 1994). *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 54(9-B), 4949.
- Hamilton, M., & Rajaram, S. (2001). The concreteness effect in implicit and explicit memory tests. *Journal of Memory and Language*, 44, 96–117.
- Hogan, R. M., & Kintsch, W. (1971). Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning and Verbal Behavior*, 10, 562–567.
- Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, 20, 497–514.
- Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, 32, 421–445.
- Hunt, R. R., & Smith, R. E. (1996). Accessing the particular from the general: The power of distinctiveness in the context of organization. *Memory & Cognition*, 24, 217–225.
- Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user’s guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 3–26.
- Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, 134, 131–148.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 3, 306–340.
- Jacoby, L. L., & Hollingshead, A. (1990). Toward a generate/recognition model of performance on direct and indirect tests of memory. *Journal of Memory and Language*, 29, 433–454.
- Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 391–422). Hillsdale: Erlbaum.
- Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In S.

- Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 383–400). New York: Guilford.
- Jones, G. V. (1978). Recognition failure and dual mechanisms in recall. *Psychological Review*, *85*, 464–469.
- Kinoshita, S. (1989). Generation enhances semantic processing? The role of distinctiveness in the generation effect. *Memory & Cognition*, *17*, 563–571.
- Kintsch, W. (1968). Recognition and free recall of organized lists. *Journal of Experimental Psychology*, *78*, 481–487.
- Kintsch, W. (1970). Models of free recall and recognition. In D. A. Norman (Ed.), *Models of human memory* (pp. 331–373). New York: Academic.
- Kintsch, W. (1978). More on recognition failure of recallable words: Implications for generation-recognition models. *Psychological Review*, *85*, 470–473.
- Knoedler, A. J., Hellwig, K. A., & Neath, I. (1999). The shift from recency to primacy with increasing delay. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 474–487.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271.
- Mandler, G., Pearlstone, Z., & Koopmans, H. S. (1969). Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, *8*, 410–423.
- Marschark, M., & Hunt, R. R. (1989). A reexamination of the role of imagery in learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 710–720.
- Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*, 679–690.
- McDaniel, M. A., DeLosh, E. L., & Merritt, P. S. (2000). Order information and retrieval distinctiveness: Recall of common versus bizarre material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1045–1056.
- McDaniel, M. A., Einstein, G. O., & Lollis, T. (1988). Qualitative and quantitative considerations in encoding difficulty effects. *Memory & Cognition*, *16*, 8–14.
- Mulligan, N. W. (1996). The effects of perceptual interference at encoding on implicit memory, explicit memory, and memory for source. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1067–1087.
- Mulligan, N. W. (2002). The effects of generation on conceptual implicit memory. *Journal of Memory and Language*, *47*, 327–342.
- Mulligan, N. W. (2006). Conceptual implicit memory and the item-specific – relational distinction. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory*. New York: Oxford University Press.
- Mulligan, N. W., Guyer, P. S., & Beland, A. (1999). The effects of levels-of-processing and organization on conceptual implicit memory in the category exemplar production test. *Memory & Cognition*, *27*, 633–647.
- Mulligan, N. W., & Stone, M. (1999). Attention and conceptual priming: Limits on the effects of divided attention in the category-exemplar production task. *Journal of Memory and Language*, *41*, 253–280.
- Nairne, J. S., Neath, I., Serra, M., & Byun, E. (1997). Positional distinctiveness and ratio rule in free recall. *Journal of Memory and Language*, *37*, 155–166.
- Nairne, J. S., Pusean, C., & Widner, R. L. (1985). Representation in the mental lexicon: Implications for theories of the generation effect. *Memory & Cognition*, *13*, 183–191.
- Nelson, D. L. (1979). Remembering pictures and words: Appearance, significance, and name. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 45–67). Hillsdale: Erlbaum.
- Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134.
- Raaijmakers, J. G., & Shiffrin, R. M. (1992). Models for recall and recognition. *Annual Review of Psychology*, *43*, 205–234.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.
- Rappold, V. A., & Hashtroudi, S. (1991). Does organization improve priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 103–114.
- Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam: Elsevier.
- Smith, R. E. (2006). Adult age differences in episodic memory: Item-specific, relational and distinctive processing. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory*. New York: Oxford University Press.
- Smith, R. E., & Hunt, R. R. (2000). The effects of distinctiveness require reinstatement of organization: The importance of intentional memory instructions. *Journal of Memory and Language*, *43*, 431–446.
- Tulving, E. (1962). Subjective organization in free recall of “unrelated” words. *Psychological Review*, *69*, 344–354.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*, 381–391.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14*, 779–804.
- Watkins, M. J., & Gardiner, J. M. (1979). An appreciation of generate-recognize theory of recall. *Journal of Verbal Learning and Verbal Behavior*, *18*, 687–704.
- Weldon, M. S., & Colston, H. L. (1995). Dissociating the generation stage in implicit and explicit memory tests: Incidental production can differ from strategic access. *Psychonomic Bulletin & Review*, *2*, 381–386.
- Weldon, M. S., & Coyote, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 670–686.