Memory for common and bizarre stimuli: A storage–retrieval analysis

DAVID M. RIEFER and MARY L. LAMAY California State University, San Bernardino, California

We tested the hypothesis that common stimuli are stored in memory better than bizarre stimuli are. Subjects memorized a series of noun pairs embedded within 20 common or bizarre sentences. By using a between-list design, free and cued recall, and intentional-learning instructions, we were able to obtain a commonness effect (i.e., a recall advantage for the common sentences). Riefer and Rouder's (1992) multinomial processing-tree model for measuring storage and retrieval was applied to the data, which revealed that the recall advantage for common sentences was due to storage and not retrieval processes. We propose a two-factor theory: that common items are stored better in memory, but that bizarre items are retrieved better from memory. This storage—retrieval explanation does a good job of accounting for a number of findings associated with the bizarreness effect.

The bizarreness effect is the observation that bizarre or unusual stimuli are remembered better than common or ordinary stimuli (Lang, 1995; Weir & Richman, 1996; Worthen & Marshall, 1996). Memory researchers have been investigating this phenomenon for over 25 years; yet experimental support for the bizarreness effect over this time span has been somewhat mixed. Early experiments (e.g., Andreoff & Yarmey, 1976; Senter & Hoffman, 1976) failed to show any recall advantage of bizarre stimuli, and some studies actually showed a recall advantage for common stimuli, a phenomenon referred to as the commonness effect (e.g., Collyer, Jonides, & Bevan, 1972; Wollen, Weber, & Lowry, 1972). More recent studies have shown that, under the right circumstances, a reliable bizarreness effect can be obtained experimentally. Specifically, the effect is best achieved using free recall, mixedlist designs, incidental-learning instructions, and delayed recall (see Einstein & McDaniel, 1987, and McDaniel, Einstein, DeLosh, May, & Brady, 1995, for reviews).

However, this recent work also reveals that the bizarreness effect is a fragile and complex one, requiring a very specific set of experimental conditions. For example, McDaniel and Einstein (1986) showed that the effect is obtained when subjects evaluate the vividness of stimuli, but it will not occur when subjects rate the bizarreness or interactiveness of stimuli. In addition, Kroll and his associates (Kroll, Jaeger, & Dornfest, 1992; Kroll, Schepeler, & Angin, 1986; Kroll & Tu, 1988) have documented a number of studies failing to achieve the bizarreness effect, and Riefer and Rouder (1992) achieved the effect in only two of three replications of the same experiment.

This research was supported by NSF Grant SBR-9309667. The authors would like to thank Gil Einstein, Neal Kroll, and an anonymous reviewer for helpful comments on an earlier draft of the manuscript. Correspondence should be addressed to D. M. Riefer, Department of Psychology, California State University, San Bernardino, CA 92407 (e-mail: driefer@wiley.csusb.edu).

Thus, the bizarreness effect continues to be a somewhat elusive phenomenon.

In this article, we present a two-factor hypothesis that helps to explain why the bizarreness effect has been so hard to obtain in many previous studies. The hypothesis is based on the contrast between storage and retrieval, and it states that (1) common stimuli are stored in memory better than bizarre stimuli are and (2) bizarre stimuli are retrieved from memory better than common stimuli are. Thus, we are proposing an interaction between the storage and retrieval stages of memory that favors common stimuli for one stage (storage) but bizarre stimuli for the other (retrieval). Depending on the strength of the contribution of these two processes relative to each other, the two-factor hypothesis can account for bizarre-imagery effects that are strong, weak, or even nonexistent. In addition, as we will show in our Discussion section, the twofactor hypothesis can explain a wide range of findings associated with the bizarreness effect.

The first part of this hypothesis, that common stimuli receive better storage, is consistent with schema-based theories of memory (e.g., Brewer & Treyens, 1981). A great deal of research shows that memory is enhanced when we can meaningfully relate new information to what we already know (e.g., Bradshaw & Anderson, 1982; Bransford & Johnson, 1972). It is reasonable to assume that when such information is integrated into a preexisting schema, the result is better storage for that information within memory. Common or usual items are more likely to share features with an existing schema than are items that are bizarre or out of the ordinary.

The second part of the hypothesis, that bizarre stimuli are retrieved from memory better than common stimuli are, has already been verified in a study by Riefer and Rouder (1992). Specifically, Riefer and Rouder developed a multinomial processing-tree model that is capable of separately measuring storage and retrieval factors in the bizarreness effect. When they applied the model to a series

of memory experiments on bizarre and common stimuli, they discovered that bizarre items were retrieved better, but not stored better, than common items were. McDaniel et al. (1995) have theorized that bizarre stimuli are retrieved better due to their distinctiveness in memory.

The specific purpose of the present study was to test experimentally the first part of the two-factor hypothesis (i.e., that common stimuli are stored better within memory than bizarre stimuli). We conducted a memory experiment involving both bizarre and common items and applied the Riefer–Rouder model to measure the relative contribution of storage and retrieval to the results. However, in order for the Riefer–Rouder model to establish that a memory phenomenon is due to storage or retrieval factors, it is first necessary to establish that phenomenon empirically. For example, it was crucial for Riefer and Rouder (1992) to obtain a bizarreness effect in their experiments before the model could verify that the effect was due to retrieval factors.

In the present study, it was therefore necessary to manipulate variables in such a way as to obtain a recall advantage for common items. This proved to be an interesting challenge, because, as far as we can determine, no study has ever been conducted with the specific purpose of obtaining a commonness effect. The few studies that have achieved such an effect were actually conducted with the purpose of observing a bizarreness effect, but they obtained a commonness effect contrary to expectations. Thus, while researchers have a good blueprint for how to produce a bizarreness effect, it is not totally clear how to manipulate variables to produce a commonness effect.

In the experiment reported here, we attempted to obtain a commonness effect by taking the variables known to produce a bizarreness effect and manipulating them in the opposite manner. The result was an experiment similar in many respects to early experiments on the bizarreness effect. Thus, we presented subjects with unmixed lists containing all bizarre or all common items, instead of using a mixed-list design. We measured cued recall in addition to the more traditional free recall. We asked subjects to memorize all items using intentional-learning instructions, without having subjects rate the items first for vividness. We also used fairly immediate recall of the to-be-remembered items. Our hope was that these specific manipulations would be successful in producing a recall advantage for common stimuli. If so, then the Riefer-Rouder model would be able to determine whether the locus of this effect was due to storage factors, retrieval factors, or both.

METHOD

Subjects

The subjects consisted of 75 undergraduate students from the University of California at Irvine. Each subject received extra course credit for volunteering to be in the experiment.

Materials

The subjects were presented with 20 sentences. Each sentence contained a noun pair (e.g., BANKER-NEWSPAPER), which described either a

common relationship ("The BANKER folded the NEWSPAPER") or a bizarre relationship between the items ("The BANKER floated on the NEWSPAPER"). As seen in the examples, the noun pairs were capitalized within each sentence. Most of the sentences were modified from sentences used by McDaniel and Einstein (1986) and Riefer and Rouder (1992), and the full set of sentences can be found in the Appendix.

Procedure

The subjects were run in six groups ranging from 10 to 14 people. Sentences were presented on slides, one at a time, at a rate of 10 sec per sentence. Order of presentation was determined randomly for each group. The subjects saw sentences that were either all common or all bizarre, with 38 subjects receiving common sentences and 37 subjects receiving bizarre sentences.

After presentation of the sentences, the subjects were given a 3-min distractor task in which they circled small differences between pairs of nearly identical pictures. The purpose of this distractor task was to help eliminate recency effects. Following this, the subjects were given 3 min to freely recall all noun pairs onto a blank sheet of paper. The subjects were instructed to recall both items of each noun pair and, if they could not remember both items, to write down individual items. A 3-min cued recall task immediately followed free recall. For this task, the subjects were presented with the first nouns of each pair, randomly ordered on a single sheet of paper. The subjects were instructed to recall the second noun from each pair.

RESULTS

Empirical Analysis

Table 1 presents the proportion of bizarre and common noun pairs correctly recalled in both the free recall and the cued recall tests. Free recall was measured as the number of noun pairs in which at least one item was freely recalled. Cued recall was measured as the number of correct responses to the second noun when the first was given as a cue (cf. Riefer & Rouder, 1992).

The main effect of sentence type was statistically significant, with the subjects recalling more common noun pairs than bizarre noun pairs for both free and cued recall $[F(1,73) = 12.46, MS_e = 17.53, p < .01]$. In addition, the subjects recalled more items during cued recall than during free recall $[F(1,73) = 202.35, MS_e = 4.21, p < .01]$. As Table 1 shows, there was a slight commonness effect for free recall, with an even greater advantage for common items in cued recall. However, this interaction between item type and type of recall failed to reach statistical significance $[F(1,73) = 2.69, MS_e = 4.21, p > .05]$.

Model Analysis

The experiment was successful in obtaining a commonness effect. However, the empirical recall advantage

Table 1
Recall Statistics for Bizarre and Common Noun Pairs

Sentence	Free Recall	Cued Recall	N_1	N_2	N_3	N_4	N_5	N_6
Bizarre	.46	.67	294	13	186	3	36	208
Common	.55	.82	391	4	223	2	14	126

Note— N_1 , both items freely recalled, correct cued recall; N_2 , one and only one item freely recalled, correct cued recall; N_3 , neither item freely recalled, correct cued recall; N_4 , both items freely recalled, incorrect cued recall; N_5 , one and only one item freely recalled, incorrect cued recall; N_6 , neither item freely recalled, incorrect cued recall.

for common items in itself does not reveal whether this effect was due to storage or retrieval factors. To determine this, Riefer and Rouder's (1992) model was applied to the data. The model is a multinomial processing-tree model as described by Riefer and Batchelder (1988), and the tree structure for the model is presented in Figure 1.

The model uses the contrast between free and cued recall to measure storage and retrieval processes and catalogues response frequencies into six data events: E_1 , both items freely recalled, correct cued recall; E_2 , one and only one item freely recalled, correct cued recall; E_3 , neither item freely recalled, correct cued recall; E_4 , both items freely recalled, incorrect cued recall; E_5 , one and only one item freely recalled, incorrect cued recall; E_6 , neither item freely recalled, incorrect cued recall, N_i is the frequency of occurrence for event E_i , aggregated across subjects and items. The N_i frequencies for the present experiment are presented in Table 1.

The Riefer-Rouder model assumes that the six recall events result from the following cognitive processes: a, probability of forming and storing a stimulus-response association; r_1 , probability of retrieving a stored association during free recall; r_2 , probability of retrieving a stored association during cued recall; s_1 , probability of recalling an unretrieved associate as a singleton; s_2 , probability of recalling a nonassociated item as a singleton. The model contains five parameters, where a and r_1 are the parameters that measure storage and retrieval processes, respectively.

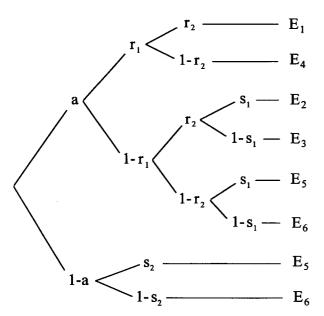


Figure 1. Riefer and Rouder's multinomial processing-tree model for measuring storage and retrieval. Interpretation of the model's parameters and the E_i events is explained in the text. From "A Multinomial Modeling Analysis of the Mnemonic Benefits of Bizarre Imagery," by D. M. Riefer and J. N. Rouder, 1992, *Memory & Cognition*, 20, p. 603. Copyright 1992 by the Psychonomic Society, Inc.

As can be seen from the model in Figure 1, the probability of each of the six data events can be represented as a function of the model's parameters:

$$P(E_1) = ar_1r_2$$

$$P(E_2) = a(1-r_1)r_2s_1$$

$$P(E_3) = a(1-r_1)r_2(1-s_1)$$

$$P(E_4) = ar_1(1-r_2)$$

$$P(E_5) = a(1-r_1)(1-r_2)s_1 + (1-a)s_2$$

$$P(E_6) = a(1-r_1)(1-r_2)(1-s_1) + (1-a)(1-s_2).$$

On the basis of these equations, Riefer and Rouder (1992) derived closed-form solutions for the model's parameters in terms of the N_i statistics:

$$a = [(N_1 + N_2 + N_3)(N_1 + N_4)]/NN_1$$

$$r_1 = N_1/(N_1 + N_2 + N_3)$$

$$r_2 = N_1/(N_1 + N_4)$$

$$s_1 = N_2/(N_2 + N_3)$$

$$s_2 = (N_1N_5 - N_2N_4)/[N_1(N_5 + N_6) - N_4(N_2 + N_3)],$$

where $N = \sum N_i$. The estimates of the model's parameters based on the above equations are presented in Table 2. Also shown are the 95% confidence intervals for each parameter.

Riefer and Rouder (1992) presented two versions of their model. Case I is an unrestricted version containing all five parameters. Case II operates under the assumption that $s_1 = s_2$, which frees up 1 degree of freedom to test the goodness of fit for the model. This fit can be evaluated using the loglikelihood ratio statistic G^2 , which is asymptotically distributed as a chi-square variable (see Riefer & Batchelder, 1988). When Case II of the model was applied to the two data sets in Table 1, this version of the model fit poorly $[G^2(2) = 20.26$; critical value = 5.99]. For this reason, Case I of the model was used to derive the parameter estimates presented in Table 2.²

As can be seen in Table 2, the values of r_2 are relatively high, which indicates that the storage of a noun pair generally resulted in correct cued recall for that pair. However, retrieval effects in the model are better represented by parameter r_1 . A computer program by Hu (1994) was used to determine if differences between parameter estimates were statistically reliable. This analysis revealed that common and bizarre noun pairs did not significantly differ on their retrievability, as measured by r_1 [$G^2(1) = 1.53, p > .05$]. In contrast, common noun pairs were stored significantly better than were bizarre noun pairs, as measured by parameter a [$G^2(1) = 40.47, p < .01$].

The only other parameter to reach statistical significance was s_1 , which was significantly higher for the bizarre stimuli $[G^2(1) = 6.53, p < .05]$. This is an interesting result, because Riefer and Rouder (1992) also found that bizarre items were recalled better as singletons than were common items. This indicates that the singleton advantage for bizarre items may be fairly robust, given that it has been observed under both mixed- and unmixed-list

Table 2
Parameter Estimates From the Riefer-Rouder Model

Sentence	а	$r_{\rm l}$	<i>r</i> ₂	s_1	s_2
Bizarre	.67	.60	.99	.07	.15
	(.6467)	(.5663)	(.98-1.0)	(.0409)	(.1119)
Common	.82	.63	.99	.02	.10
	(.7984)	(.6066)	(.98-1.0)	(.0103)	(.0614)

Note—Numbers in parentheses are 95% confidence intervals. a, probability of forming and storing a stimulus-response association; r_1 , probability of retrieving a stored association during free recall; r_2 , probability of retrieving a stored association during cued recall; s_1 , probability of recalling an unretrieved associate as a singleton; s_2 , probability of recalling a nonassociated item as a singleton.

conditions. If the bizarreness of stimuli can reliably be shown to improve single-item access, such a result may have important implications for theories on the bizarreness effect.

DISCUSSION

Memory for bizarre and common stimuli can be explained in terms of a two-factor hypothesis involving storage and retrieval. Results of the present experiment indicate that common stimuli benefit from superior storage processes in memory, whereas Riefer and Rouder (1992) showed that bizarre stimuli exhibit better retrieval from memory. This two-factor hypothesis helps to explain why the bizarreness effect has historically been weak or nonexistent in many previous studies. Bizarre stimuli may benefit from superior retrieval, but this can be offset to some degree by the fact that common stimuli are stored in memory better than bizarre stimuli are. The combined, empirical effect of these two opposing processes may produce a significant bizarreness effect, but it could also produce a commonness effect or no effect at all, depending on the relative contributions of storage and retrieval processes in the experiment.

It should be noted, however, that this two-factor hypothesis has been tested only under a very specific set of experimental conditionsnamely, the retrieval advantage for bizarre stimuli using a mixed-list design (Riefer & Rouder, 1992) and the storage advantage for common stimuli using an unmixed-list design (the present study). What has yet to be determined is whether these specific list conditions are the only ones capable of producing these effects. If there is something inherent within stimuli that makes common items easier to store, or bizarre items easier to retrieve, then it should be possible to observe both of these effects in the same experimental design. However, this has yet to be demonstrated in any controlled study, and, therefore, a logical extension of the present research would be to investigate how mixing list items determines the relative contributions of storage and retrieval in the bizarreness effect. A possible framework for this type of study comes from an experiment by Worthen and Marshall (1996), in which the ratio of common to bizarre items was systematically varied within memorized lists.

In general, the two-factor hypothesis has the potential to guide and inform theoretical development on the bizarreness effect. Storage and retrieval are important processes in human memory, and it is reasonable to assume that they can differ between bizarre and common items in many situations. For example, a great deal of research has explored how the bizarreness effect is influenced by different experimental variables. A storage—retrieval analysis can help determine why a certain variable has a particular effect, which in turn can aid in the development of theoretical explanations for the bizarreness effect. We illustrate this with a few examples:

Free versus cued recall. It is well established that the bizarreness effect is harder to obtain using cued recall relative to using free recall (Riefer & Rouder, 1992). This is consistent with the idea that presenting subjects with a cue minimizes or eliminates retrieval. Because the

bizarreness effect depends on retrieval processes, anything that reduces the demands of retrieval should neutralize the effect.

Sentence complexity. A number of studies (e.g., McDaniel & Einstein, 1989; Richman, Dunn, Kahl, Sadler, & Simmons, 1990) have demonstrated that the bizarreness effect fails to occur when sentences are complex (e.g., containing extra adjectives). Robinson-Riegler and McDaniel (1994) have proposed that the surplus information in complex sentences helps memory by providing extra retrieval cues, but that these cues are more salient and hence predominate over the cues provided by bizarreness. This is a retrieval-based explanation that is consistent with the two-factor hypothesis. It assumes that bizarre items are retrieved better in simple sentences, but that this retrieval advantage is eliminated in complex sentences because the cues associated with complexity are equally effective for common and bizarre sentences.

Presentation rate. If bizarre stimuli are stored more poorly in memory than are common stimuli, then it is reasonable to suspect that bizarre items might derive a greater benefit from experimental manipulations that improve storage. One variable that should have a strong effect on encoding and storage is the presentation rate for stimuli. Giving extra study time for items may help overcome the storage disadvantage for bizarre items and, assuming that bizarre items retain their retrieval advantage, could help strengthen the bizarreness effect. As it turns out, there is evidence that the bizarreness effect can be aided by longer study times. Kline and Groninger (1991) found that, with longer presentation rates, complex bizarre sentences were recalled better than were common sentences. In addition, laccino and Sowa (1989) obtained a bizarreness effect for immediate recall when study time was subject-paced, but not when it was experimenter-paced.

Level of association. Another variable that should have a strong effect on storage capacity is the level of association between items within a sentence. For example, Batchelder and Riefer (1986) used a pairclustering model to show that high-associate category pairs were stored better, but not retrieved better, than were low-associate pairs. In the present paradigm, it is reasonable to assume that high association between noun pairs helps in the formation and storage of the stimulus-response association in memory (as measured by the storage parameter a in the Riefer-Rouder model). Lang (1995) has demonstrated that association level can, in fact, be an important factor in the bizarreness effect; specifically, she found that high-associate sentences produced a bizarreness effect but that weak-associate sentences resulted in a commonness effect. The two-factor hypothesis can explain this finding by assuming that high levels of association, like extra study time, enhance the storage of bizarre stimuli relative to common stimuli (which presumably are already stored well). Using high-associate items may serve to reduce the storage advantage for common stimuli, thus strengthening the bizarreness effect.

Although the two-factor hypothesis is capable of accounting for a number of basic findings, it is important to note that the above theoretical discussion remains largely speculative. No model analysis has ever been conducted to determine directly whether presentation rate or association level affects the storage of common and bizarre stimuli in the manner described above or whether sentence complexity serves to equalize the retrievability of bizarre and common items. Indeed, it is quite plausible that the complexity of sentences may effect their encodability and storage in memory, as well as their retrievability.

However, all of the above hypotheses can easily be tested using multinomial models of storage and retrieval, such as the Riefer-Rouder model.³ One of our goals has been to outline how this model can be used as a helpful tool in research on bizarre and common stimuli. In addition to the variables mentioned above, other variables that have been shown to influence the bizarreness effect include delayed recall (e.g., Iaccino & Sowa, 1989; O'Brien & Wolford, 1982), list length (Richman et al., 1990), and implicit versus explicit memory tasks (Nicolas & Marchal, 1996). It would be straightforward to use the Riefer-Rouder model to explore the influence of these variables on storage and retrieval. The model could also prove informative in exploring new issues concerning bizarre and common stimuli, such as the differences between pictorial versus written stimuli, or self-generated versus experimenter-produced imagery. Another useful extension of the present research would be to explore more thoroughly the validity of the model as a method for mea-

suring storage and retrieval. This could be accomplished by manipulating basic experimental variables to see if they have a selective influence on one process or the other. Many of the variables mentioned above can be expected, on logical or theoretical grounds, to have their main effect on storage (e.g., presentation rate, association level) or retrieval (e.g., cued recall, sentence complexity).

As a final illustration of how models of storage and retrieval can be used to explore storage and retrieval factors in the bizarreness effect, we focus on two experimental variables that might play an important role in producing the effect, especially in terms of their influence on storage processes. The first variable is the contrast between incidental and intentional learning. Most experiments on the bizarreness effect (e.g., Lang, 1995; Robinson-Riegler & McDaniel, 1994; Worthen & Marshall, 1996) have used an incidental learning task, in which subjects rate stimuli on some dimension (such as vividness) but are not informed of the final memory test. However, a few studies have obtained the bizarreness effect under conditions where subjects are given explicit instructions to memorize (e.g., Nicolas & Marchal, 1996; Wollen & Cox, 1981). This variable has not been systematically explored in many studies on bizarre imagery, but it is reasonable to speculate that it may have its main effect on the storage of information. Subjects under intentionallearning instructions are more likely to memorize by using various mnemonic techniques, which should mainly enhance the encoding and storage of items. Any possible storage differences between incidental and intentional learning, and how they may impact the bizarreness effect, are issues that can be easily explored using the Riefer-Rouder model.

A second potentially important variable in the bizarreness effect is association level. Lang's (1995) research on this issue is noteworthy because previous studies of bizarre imagery have not attempted to measure or control the level of association between items. Many studies on the bizarreness effect, including the present one, have used stimuli adopted from McDaniel and Einstein (1986). Yet a close examination of their sentences, as well as ours, reveals stimulus items that range from weakly associated (e.g., cat-coffee, lamp-book) to strongly associated (e.g., minister-bible, horse-hay). A more thorough examination of the role of association level, and especially its effects on storage and retrieval processes, may provide valuable insights into the nature of the bizarreness effect.

REFERENCES

- ANDREOFF, G. R., & YARMEY, A. D. (1976). Bizarre imagery and associative learning: A confirmation. Perceptual & Motor Skills, 43, 143-148.
 BATCHELDER, W. H., & RIEFER, D. M. (1986). The statistical analysis of a model for storage and retrieval processes in human memory. British Journal of Mathematical & Statistical Psychology, 39, 129-149.
- Bradshaw, G. L., & Anderson, J. R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning & Verbal Behavior*, 21, 165-177.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning & Verbal Behavior, 11, 717-726.
- Brewer, W. F., & Treyens, J. C. (1981). Role of schemata for places. *Cognitive Psychology*, **13**, 207-230.
- COLLYER, S. C., JONIDES, J., & BEVAN, W. (1972). Images as memory aids: Is bizarreness helpful? *American Journal of Psychology*, **85**, 31-38.
- EINSTEIN, G. O., & McDaniel, M. A. (1987). Distinctiveness and the mnemonic benefit of bizarre imagery. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 78-102). New York: Springer-Verlag.
- Hu, X. (1994). Statistical inference program for multinomial binary tree models (Version 2.0) [Computer program]. Irvine: University of California
- IACCINO, J. F., & SOWA, S. J. (1989). Bizarre imagery in paired-associate learning: An effective mnemonic aid with mixed context, delayed testing, and self-paced conditions. *Perceptual & Motor Skills*, 68, 307-316.
- KLINE, S., & GRONINGER, L. D. (1991). The imagery bizarreness effect

- as a function of sentence complexity and presentation time. Bulletin of the Psychonomic Society, 29, 25-27.
- Kroll, N. E. A., JAEGER, G., & DORNFEST, R. (1992). Metamemory for the bizarre. *Journal of Mental Imagery*, 16, 173-190.
- KROLL, N. E. A., SCHEPELER, E. M., & ANGIN, K. T. (1986). Bizarre imagery: The misremembered mnemonic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 12, 42-53.
- KROLL, N. E. A., & Tu, S.-F. (1988). The bizarre mnemonic. Psychological Research, 50, 28-37.
- LANG, V. A. (1995). Relative association, interactiveness, and the bizarre imagery effect. American Journal of Psychology, 108, 13-35.
- McDaniel, M. A., & Einstein, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 12, 54-65.
- McDaniel, M. A., & Einstein, G. O. (1989). Sentence complexity eliminates the mnemonic advantage of bizarre imagery. *Bulletin of the Psychonomic Society*, 27, 117-120.
- McDaniel, M. A., Einstein, G. O., DeLosh, E. L., May, C. P., & Brady, P. (1995). The bizarreness effect: It's not surprising, it's complex. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21, 422-435.
- NICOLAS, S., & MARCHAL, A. (1996). Picture bizarreness effect and word association. Current Psychology of Cognition, 15, 629-643.
- O'BRIEN, E. J., & WOLFORD, C. R. (1982). Effects of delay in testing on retention of plausible versus bizarre mental images. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 8, 148-152.
- RICHMAN, C. L., DUNN, J., KAHL, G., SADLER, L., & SIMMONS, K. (1990). The bizarre sentence effect as a function of list length and complexity. Bulletin of the Psychonomic Society, 28, 185-187.
- RIEFER, D. M., & BATCHELDER, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95, 318-339.
- RIEFER, D. M., & ROUDER, J. N. (1992). A multinomial modeling analysis of the mnemonic benefits of bizarre imagery. *Memory & Cognition*, 20, 601-611.
- ROBINSON-RIEGLER, B., & McDaniel, M. A. (1994). Further constraints on the bizarreness effect: Elaboration at encoding. *Memory & Cognition*, 22, 702-712.
- ROUDER, J. N., & BATCHELDER, W. H. (in press). Multinomial models for measuring storage and retrieval processes in paired-associate learning. In C. Dowling, F. Roberts, & P. Theuns (Eds.), *Progress in mathematical psychology*. Hillsdale, NJ: Erlbaum.
- SENTER, R. J., & HOFFMAN, R. R. (1976). Bizarreness as a nonessential variable in mnemonic imagery: A confirmation. *Bulletin of the Psychonomic Society*, 7, 163-164.
- WEIR, D., & RICHMAN, C. L. (1996). Subject-generated bizarreness: Imagery or semantic processing. American Journal of Psychology, 109, 173-185.
- WOLLEN, K. A., & Cox, S. D. (1981). The bizarreness effect in a multitrial intentional learning task. *Bulletin of the Psychonomic Society*, 18, 296-298.
- WOLLEN, K. A., WEBER, A., & LOWRY, D. H. (1972). Bizarreness versus interaction of mental images as determinants of learning. Cognitive Psychology, 3, 518-523.
- WORTHEN, J. B., & MARSHALL, P. H. (1996). Intralist and extralist sources of distinctiveness and the bizarreness effect: The importance of contrast. *American Journal of Psychology*, **109**, 239-263.

NOTES

1. The remaining parameters are not central to the measurement of storage and retrieval, and are therefore "nuisance" parameters. Parameter r_2 , like r_1 , also reflects retrieval processes—specifically, the probability that the second noun will successfully be retrieved given the first noun as a cue, given that the two nouns are associated within memory. As Riefer and Rouder (1992) and Rouder and Batchelder (in press) have pointed out, when the interval between study and recall is relatively short (as is the case in the present experiment), the estimate of r_2 should be very close to 1 and, thus, not very informative. The parameters s_1 and

 s_2 represent the recall of individual items as singletons, provided that the noun pair is not associated or that it fails to be retrieved. These parameters can be assumed to reflect both storage and retrieval processes (similar to parameter u in Batchelder & Riefer's, 1986, pair-clustering model).

2. Rouder and Batchelder (in press) have developed and tested an alternative version of the Riefer-Rouder model. Specifically, they present a family of models that can be used to measure storage and retrieval in the same free-then-cued paradigm used in this study. Their models are based on the same general principles as the Riefer-Rouder model, but they make different assumptions about secondary processes, such as cued-recall retrieval and singleton recall. The Rouder-Batchelder models, when applied to the data from the present experiment, fit the data significantly better than does the Riefer-Rouder model, while producing the same pattern of results. Specifically, common items were stored significantly bet-

ter than were bizarre items, bizarre items were recalled better as singletons (their parameter s), with no significant effects for any other parameters. Because both models reach essentially the same conclusions, only the results from the Riefer-Rouder model are presented here in detail.

3. Rouder and Batchelder's (in press) model is also a possible choice as a method for measuring storage and retrieval. To date, all data sets that have been analyzed using the Riefer-Rouder and Rouder-Batchelder versions have produced the same basic pattern of results in terms of the models' main parameters. This is not surprising since both versions share many of the same underlying assumptions. However, the superior fit of the Rouder-Batchelder model to the present data set suggests that it may be a more parsimonious and valid methodology for measuring storage and retrieval processes. Further research comparing these models needs to be conducted to determine if this is in fact the case.

APPENDIX Bizarre and Common Sentences Used in the Experiment

Common	Bizarre			
The BANKER folded the NEWSPAPER.	The BANKER floated on the NEWSPAPER.			
The BOY fed the CAT.	The BOY juggled the CAT.			
The CAR drove past the FENCE.	The CAR was petting the FENCE.			
The COCKROACH appeared on the STOVE.	The COCKROACH moved the STOVE.			
The DOCTOR read the JOURNAL.	The DOCTOR burned the JOURNAL.			
The DOG chased the BICYCLE.	The DOG rode the BICYCLE.			
The GIRL kissed the DOLL.	The GIRL boiled the DOLL.			
The GOLDFISH was swimming in the BOWL.	The GOLDFISH was eating out of the BOWL.			
The HORSE ate the HAY.	The HORSE smoked the HAY.			
The LAMP shined on the BOOK.	The LAMP read the BOOK.			
The MAID spilled the AMMONIA.	The MAID drank the AMMONIA.			
The NURSE worked at the COMPUTER.	The NURSE danced on the COMPUTER.			
The PERFORMER sang for the AUDIENCE.	The PERFORMER ate the AUDIENCE.			
The PLANT rested on top of the TELEVISION.	The PLANT screamed at the TELEVISION.			
The SHOES were placed by the MILK.	The SHOES were filled with MILK.			
The SNOWFLAKE fell on the MOUNTAIN.	The SNOWFLAKE climbed the MOUNTAIN.			
The SOLDIER was sitting in the CHAIR.	The SOLDIER was wearing the CHAIR.			
The SPIDER crawled on the SIDEWALK.	The SPIDER watered the SIDEWALK.			
The SUN set in the WEST.	The sun did backflips in the west.			
The WOLF howled at the MOON.	The WOLF lassoed the MOON.			

(Manuscript received January 22, 1997; revision accepted for publication October 14, 1997.)