Journal of Experimental Psychology:
Learning, Memory, and Cognition
2000, Vol. 26, No. 4, 1045–1056

0278-7393/00/$5.00 DOI: 10.1037//0278-7393.26.4.1045

Order Information and Retrieval Distinctiveness: Recall of Common Versus Bizarre Material

Mark A. McDaniel
University of New Mexico

Edward L. DeLosh and Paul S. Merritt
Colorado State University

The order-encoding hypothesis (E. L. DeLosh & M. A. McDaniel, 1996) assumes that serial-order information contributes to the retrieval of list items and that serial-order encoding is better for common items than bizarre items. In line with this account, Experiment 1 revealed better free recall and serial-order memory for common than for bizarre items in pure lists, and Experiment 2 showed that recall for bizarre items increased and the recall advantage of common items was eliminated when serial-order encoding for bizarre items was increased to the level of common items. However, inconsistent with a second assumption that bizarre-item advantages in mixed lists reflect better individual-item encoding for bizarre items, Experiments 3 and 4 showed that the bizarreness effect in mixed lists is eliminated when alternative retrieval strategies are encouraged. This set of findings is better explained by the differential-retrieval-process framework, which proposes that contextual factors (e.g., list composition) influence the extent to which various types of information are used at retrieval, with the bizarreness advantage in mixed lists dependent on a distinctiveness-based retrieval process.

The memory literature is replete with instances in which the effects of stimulus material on free recall are reversed or eliminated as a function of whether the manipulation that produces the effect is constructed within lists or between lists. (Note that within-lists manipulations involve mixed lists of items and between-lists manipulations involve lists that are composed entirely of each level of the particular material manipulation.) These effects include superior recall of bizarre sentences and of humorous sentences (relative to common sentences) in mixed but not unmixed lists (McDaniel & Einstein, 1986; McDaniel, Einstein, DeLosh, May, & Brady, 1995; Schmidt, 1994), superior recall of generated words and of words with perceptual interference (relative to intact words) in mixed but not unmixed lists (e.g., Mulligan, 1999; Serra & Nairne, 1993; Slamecka & Katsaiti, 1987), better recall of detailed pictures than of simple pictures in mixed but not unmixed lists (Zucco, Traversa, & Cornoldi, 1984), and greater recall of high-frequency words relative to low-frequency words in pure but not mixed lists (DeLosh & McDaniel, 1996; see Balota & Neely, 1980, for an exception with long mixed lists). These effects have stimulated considerable effort toward theoretical explanation, and the fruitfulness and complexity of these explanations have turned, in part, on their ability to account for the differential patterns that are observed as a function of between-list designs versus within-list designs, hereinafter referred to as list type (e.g., for the generation effect, word frequency effect, bizarreness effect, and humor effect). The bizarre-imagery effect is particularly interesting from this standpoint because there are several alternative theoretical explanations, and they are representative of broader theoretical ideas that are prominent in the current literature. The major objective of this article is to experimentally examine these theoretical explanations of the free-recall patterns associated with bizarre imagery. In doing so, this study helps inform the broader theories from which the explanations originate.

Order Encoding and the Bizarre-Imagery Effect

One potential explanation of the bizarre-imagery effect emerges from a theoretical account (based on initial work by Nairne, Riegler, & Serra, 1991) that attempts to explain the effects of list design within a single unifying framework (DeLosh & McDaniel, 1996). This framework centers on the observation that the encoding and recall of the serial order of events is important in daily functioning, as it allows people to fairly accurately remember the sequence of experienced events. On the basis of past studies, the assumption is that in the laboratory when participants are presented with a list of individual items to commit to memory, serial-order information is commonly encoded and can also contribute to the retrieval of list items even in free recall (cf. Mandler & Dean, 1969; Toglia & Kimble, 1976; Tzeng, Lee, & Wetzel, 1979). Serial-order information can be viewed as a type of relational information that helps to organize a list of unrelated items (DeLosh & McDaniel, 1996). Given that relational information serves to improve free recall (see Hunt & McDaniel, 1993, for a review), the degree to which serial-order information is encoded should influence free-recall levels.

Mark A. McDaniel, Department of Psychology, University of New Mexico; Edward L. DeLosh and Paul S. Merritt, Department of Psychology, Colorado State University.

Portions of Experiment 2 were presented at the April 1999 Meeting of the Southwestern Psychological Association in Albuquerque, New Mexico. We are grateful to Dan Burns and Steve Schmidt for helpful suggestions on this article. We thank Jim Worthen for discussions that stimulated preparation of this article.

Correspondence concerning this article should be addressed to Mark A. McDaniel, Department of Psychology, University of New Mexico, Albuquerque, New Mexico 87131. Electronic mail may be sent to mcdaniel@unm.edu.

The second critical assumption of the framework is that for typical or common items (such as sentences expressing common actions) serial-order information is ordinarily encoded, just as order information about daily events tends to be encoded. In contrast, more unusual items (such as sentences expressing bizarre actions) require attentional resources for processing and interpretation of the items (see Worthen, Garcia-Rivas, Green, & Vidas, 2000, for a review of this approach to the bizarreness effect). The attention devoted to making sense of the bizarre items typically detracts from encoding the order of the items. The prediction is that serial-order encoding of pure lists of common items should be greater than serial-order encoding of pure lists of bizarre items. For the situation when common and bizarre items are intermixed in one list, the order information for common items tends to be reduced (relative to pure lists) because in some cases these items follow bizarre items, which disrupt order encoding. By the same token, the order information for bizarre items tends to improve (relative to pure lists) because in some cases bizarre items follow common items, which will promote order encoding. In general, this dynamic tends to produce approximately equivalent order encoding for items in mixed lists. One initial study has verified these expectations for common versus bizarre sentences (McDaniel et al., 1995). As more general support for the order-encoding framework, other manipulations that yield relatively common and unusual items tend to also show the order-memory pattern outlined above (see Nairne et al., 1991, for read vs. fragmented presentations; see DeLosh & McDaniel, 1996, for high- vs. low-frequency words; see Mulligan, 1999, for words with no perceptual interference vs. words presented with perceptual interference; see Engelkamp & Dehn, 2000, for experimenter-performed actions vs. participant-performed actions).

The differential patterns of order encoding across pure and mixed lists is at the heart of DeLosh and McDaniel's (1996) account of why the effects of materials on free recall change as a function of list design. For more typical items (common sentences), recall should decrease from pure to mixed lists as order information becomes less well encoded. By contrast, for unusual items (bizarre sentences), recall should increase from pure to mixed lists as order information becomes better encoded. Further, because both relational and individual-item elaboration are needed for optimal recall (Einstein & Hunt, 1980; Hunt & McDaniel, 1993), for mixed lists of unrelated items, bizarre items (or any kind of item that stimulates more elaboration) should be better recalled than common items because bizarre items benefit from the encoding of both serial-order information and extensive individual-item information. More extensive individual-item encoding of bizarre items is posited by many theorists (e.g., Wollen & Margres, 1987; Worthen et al., in press). The order-encoding view also provides a ready explanation of why the bizarreness effect is eliminated or reversed in pure lists: Bizarre items have less order information to help guide recall than do common items, thereby countering their advantage for individual-item information.

Distinctiveness and the Bizarre-Imagery Effect

A plausible and prominent theoretical alternative for explaining the bizarre-imagery effect is that the mixed list itself produces a retrieval context that favors unusual items in recall, whereas unmixed lists do not. The idea here is that a list of items defines a retrieval set at the time of recall, and within this retrieval set distinct items will more likely be identified for recall than will nondistinct items (Knoedler, Hellwig, & Neath, 1999; McDaniel et al., 1995; Neath, 1991). Note that this distinctiveness hypothesis is a favored and long-standing notion in the memory literature (von Restorff, cited in Kofka, 1935; see Hunt, 1995, and Knoedler et al., 1999, for historical context), and accordingly represents a strong alternative to the order-encoding account. With regard to the present focus on the bizarre-imagery effect, for mixed lists, the encoded representations of bizarre items will be distinct relative to the representations of common items, thereby favoring recall of bizarre items (Einstein & McDaniel, 1987; McDaniel & Einstein, 1986). For unmixed lists, because all of the items in the retrieval set will be similar on the dimension of bizarreness (either all bizarre or all common), other dimensions of distinctiveness such as temporal distinctiveness will determine recall (Knoedler et al., 1999). Presuming that these other dimensions should be equally present in pure bizarre and pure common lists, there would be no bizarreness advantage in unmixed list designs.

In favor of the distinctiveness hypothesis, the preliminary evidence relating to the order-memory hypothesis in bizarre-imagery effects does not establish an unequivocal link between encoding of order information and recall. The one study to investigate order memory and bizarre-imagery effects (McDaniel et al., 1995, Experiments 4 and 5) revealed that order reconstruction for common sentences was reliably better than order reconstruction for bizarre sentences when pure lists were examined, but no difference was observed when mixed lists were examined. Importantly, however, for pure lists of bizarre material and mixed lists, the correlations between recall and the levels of order memory (as assessed on an order-reconstruction test) were low and not significant (Experiment 5). Further, Asch–Ebenholtz (1962) scores for input–output correspondence (an index that reflects the use of order information during recall) were not above chance levels for pure bizarre lists and mixed lists. In free recall, the usual pattern was obtained such that bizarre items were recalled significantly better than were common items in mixed but not unmixed lists. This set of results suggests that order information is not used to help guide retrieval for all types of lists, as presumed by the order-encoding view. Instead, the pattern is consistent with the idea that mixed lists produce a retrieval context in which bizarre sentences become distinctive, with distinctiveness guiding retrieval. The present experiments were designed to examine in greater detail the extent to which order-encoding dynamics, distinctiveness processes, or both contribute to the bizarreness effects in free recall.

Experiment 1

As just described, McDaniel et al. (1995; Experiment 5) reported that order information did not significantly relate to the retrieval process when lists contained bizarre material, thereby countering the order-encoding explanation. This explanation assumes that order information is used to guide recall regardless of list composition. However, order information could not guide recall if there were relatively poor order encoding for bizarre items to begin with. That is, it is possible that the amount of order information associated with bizarre material in McDaniel et al. was not adequate for effective use in guiding retrieval. Order memory was approximately 30% lower in their experiment than in a comparable study with high- and low-frequency words (DeLosh & McDaniel, 1996).

Therefore, before discounting the order-encoding explanation, we felt it was important to examine the bizarreness effect under conditions in which order memory was relatively good. To boost order memory, lists consisted of six items rather than the eight items used by McDaniel et al. (1995), and the initial instructions completely informed participants about the nature of the recall and order-reconstruction memory tasks, which was reiterated during the presentation of a practice list. If order information is used to guide retrieval regardless of the list composition (as the order-encoding framework suggests), then recall patterns should reflect the initial presentation order.

In contrast, if the bizarreness advantage found in recall of mixed lists is based on retrieval dynamics that depend on distinctiveness as determined by the items in the retrieval set (Knoedler et al., 1999; McDaniel et al., 1995), then degree of the serial-order encoding of the list items should be immaterial. Even with the encoding of serial-order information encouraged, recall of lists mixed with common and bizarre items should not be guided by order information.

Method

Participants and design. Participants were 72 Purdue University undergraduates who completed the experiment in partial fulfillment of an introductory psychology course. Item type (bizarre vs. common) was varied between and within subjects using Erlebacher's (1977) method, with 24 participants randomly assigned to each of the three conditions (pure lists of common sentences, pure lists of bizarre sentences, and mixed lists). Each participant was tested individually in a 30-min session using an IBM-compatible computer.

Materials. The materials consisted of capitalized noun triplets embedded in simple sentences. A total of 30 triplets were used (24 triplets adapted from McDaniel & Einstein, 1986, plus 6 new triplets). Bizarre and common sentences were constructed for each noun triplet by varying the relationship between nouns. For example, one common sentence was "The MINISTER read the BIBLE after DINNER," and the corresponding bizarre sentence was "The MINISTER ate the BIBLE during DINNER." Five pure common lists and five pure bizarre lists were formed by randomly assigning six sentences of the appropriate type to each list. Five mixed lists were constructed by randomly assigning three common and three bizarre sentences to each list. Then a second set of mixed lists was created by replacing common and bizarre sentences of the first set with their counterparts. Half of the participants in the mixed-list condition received Set 1, and half received Set 2. As a result, the common and bizarre versions of each noun triplet appeared equally often. Within each list, sentences were randomly assigned to one of the six serial positions.

Procedure. In the initial instructions, both the recall and the order-reconstruction tests were described in detail. In addition, a practice list and practice tests were given. Thus, participants knew that they would be expected to reproduce the order of item presentation prior to the encoding of any of the critical lists. A total of five experimental lists were presented (with each list representative of the condition to which a participant was assigned: pure common, pure bizarre, or mixed lists). The six sentences in a list were shown one at a time on a computer monitor for 7 s, with a warning tone preceding each presentation. Participants were instructed to form an interactive image of the event described by each sentence and to hold that image "in their head" for the entire time that the sentence remained on the screen. Following each sentence, participants were given 3 s to rate the vividness of their mental image on a scale of 1 to 5, using a computer keyboard to input their choice.

After imaging all six items constituting a list, participants completed math problems for 30 s (distractor activity), followed by recall and then order reconstruction. For the free-recall test, 2 min were provided for participants to write target nouns on a response sheet. For the order-reconstruction test, the complete list of six sentences was shown on the computer monitor, but in a different (random) order than that of the study phase. These sentences were labeled A–F. Using a response sheet with response spaces representing each serial position (labeled first through last), participants were told to reproduce the original presentation order of the sentences. To do this, they placed the sentence labels in the response blanks in accordance with their memory for the order of the sentences.

Results

The rejection level for all analyses in this article was .05. All means are displayed in Table 1.

Table 1
Order Reconstruction, Free Recall, and Input–Output Correspondence as a Function of Bizarreness in Experiments 1–4

	Reconstruction		Free recall		I–O correspondence	
Experiment and condition	C	B	C	B	C	B
Experiment 1						
Pure lists	.62	.55	.67	.62	.66	.58
Mixed lists	.52	.54	.59	.69	(.58)	(.58)
Experiment 2						
Standard instructions	.66	.53	.67	.59	.72	.61
Serial instructions	.64	.61	.65	.67	.71	.68
Experiment 3						
Standard instructions	—	—	.49	.57	(.51)	(.51)
Serial instructions	—	—	.58	.55	(.65)	(.65)
Experiment 4						
Uncategorized lists	—	—	.43	.51	(.50)	(.50)
Categorized lists	—	—	.60	.58	(.50)	(.50)

Note. Experiment 2 used pure lists, whereas Experiments 3 and 4 used mixed lists. Experiments 3 and 4 did not include order reconstruction tests. The input–output correspondence values given in parentheses pertain to mixed lists as a whole. C = common sentences; B = bizarre sentences; I–O = input–output.

Order reconstruction. Order-reconstruction scores were submitted to a 2 × 2 analysis of variance (ANOVA) with sentence type and list type as independent variables (using Erlebacher's [1977] method). This analysis yielded a marginally significant interaction, $F(1, 66) = 3.42$, $MSE = 0.02$, $p = .07$, but no main effects ($Fs < 2.15$). Tukey's highly significant difference (HSD) tests were performed to examine the expectations outlined in the introduction. Order memory was significantly greater for pure lists of common sentences than for pure lists of bizarre sentences, whereas in mixed lists, order memory did not differ across sentence types. Nine participants in the mixed-list group showed an advantage for common sentences, 11 showed an advantage for bizarre sentences, and there were 4 ties. Tukey comparisons indicated that the order memory associated with common sentences was lower in mixed lists than in pure lists. Unexpectedly, from the DeLosh and McDaniel (1996) framework, the order memory associated with bizarre sentences did not improve in mixed lists.

Free recall. Scores for the proportion of nouns recalled were submitted to a 2 × 2 (Sentence Type × List Type) ANOVA using Erlebacher's (1977) method. This analysis yielded an interaction between sentence type and list type, $F(1, 43) = 7.42$, $MSE = 0.02$. Words from common sentences were recalled better than words from bizarre sentences when pure lists were compared, whereas the opposite held true for mixed lists (see Table 1 for means). Two participants in the mixed-list group showed a recall advantage for common sentences, 17 exhibited an advantage for bizarre sentences, and there were 3 ties. The observed crossover interaction resulted from dynamics whereby the recall of common sentences decreased, but the recall of bizarre sentences increased in mixed lists relative to pure lists.

Input–output correspondence. Asch–Ebenholtz (1962) scores were computed to assess input–output correspondence. The first word recalled from each sentence was used to determine a particular sentence's output order. That is, if other words were recalled from the same sentence, they were ignored in computing the Asch–Ebenholtz score. A one-way between-subjects ANOVA of these scores revealed a significant effect of list type, $F(2, 69) = 4.40$, $MSE = 0.06$. Tukey HSD tests indicated that input–output correspondence was greater for pure common lists (.6) than for pure bizarre lists (.58) or mixed lists (.58). Only performance in the pure common list condition was significantly above that expected by chance, $t(23) = 5.12$, $SEM = 0.03$.

Discussion

The results of the current experiment are only partially consistent with the predictions of the DeLosh and McDaniel (1996) order-encoding hypothesis. As expected, pure lists of common sentences produced better order memory than did pure lists of bizarre sentences, and there was no difference across sentence types in the case of mixed lists. In similar fashion, common lists yielded greater input–output correspondence than did bizarre lists. Accompanying these differences in memory for serial order and the use of serial order in recall, words were recalled from common sentences more often than from bizarre sentences when pure lists were compared.

In mixed lists, the advantage for bizarre sentences over common sentences was obtained, and this advantage was associated with an increase in bizarre-item recall for mixed relative to unmixed lists. These expected patterns in free recall did not seem to be closely related to order memory, however. First, the order memory associated with bizarre sentences did not improve in mixed relative to pure lists; indeed, it nominally decreased. In the order-encoding view, such an increase is predicted and is assumed to underlie the increase in bizarre-item recall. Of course, this null result by itself is not overly strong counterevidence against the order-encoding view.

Second, however, the input–output correspondences for mixed lists were not significantly above chance, suggesting that order information was not used to help guide recall. This result was obtained even with relatively good serial-order retention. We observed order-reconstruction scores of .54 and .52 for bizarre and common sentences, respectively, in mixed lists, compared with McDaniel et al.'s (1995) reported scores of .29 and .28 in Experiment 4 and .39 and .34 in Experiment 5. Thus, the argument that order memory was not robust enough to be used in recall is much less plausible in the present experiment. Indeed, because order-memory testing followed free recall, order-memory accuracy could have been somewhat disrupted by output order in free recall. If so, then the observed order-memory scores may underestimate the amount of order information present at the time that free recall was tested. Despite the possibility that order memory during free recall was even higher than measured, free recall for mixed lists reflected little, if any, use of order information.

The current results (see also McDaniel et al., 1995) suggest a more integrative approach that combines features of both the order-encoding and the distinctiveness views (for purposes of exposition, we label this approach the *differential-retrieval-process* view). Our proposal (cf. McDaniel et al., 1995) is that regardless of the amount of order information initially encoded, list composition influences the extent to which order information is actually used at retrieval. The idea is that order information is routinely incorporated into the retrieval process for pure lists of common items, whereas the use of order information—even if encoded—is discouraged when bizarre items are present (in either pure or mixed lists). Furthermore, in free recall of mixed lists of bizarre and common items, item distinctiveness typically becomes a more salient cue for item identification (or retrieval; e.g., Knoedler et al., 1999), with bizarre items being distinct relative to the other (common) items that compose the retrieval set delimited by contextual cues (Hunt & McDaniel, 1993; McDaniel et al., 1995). Note that the idea here is that distinctiveness is not due to more involved or elaborate processing of bizarre items, but instead emerges as a function of the retrieval context.

The present findings support this differential-retrieval-process view in that input–output correspondence was not significantly above chance for lists containing bizarre items, despite the increase in the overall level of order encoding relative to Experiment 5 of McDaniel et al. (1995). The differential-retrieval-process view is also attractive because it provides an account for why common items were better remembered than were bizarre items in unmixed lists (a result obtained in McDaniel et al., 1995, as well). Current views of the bizarreness effect that either assume more involved or effortful processing of bizarre items (Wollen & Margres, 1987), or assume that retrieval is based solely on distinctiveness relative to the retrieval context (i.e., the distinctiveness account; Knoedler et al., 1999; McDaniel & Einstein, 1986), or both (e.g., Worthen, Marshall, & Cox, 1998), do not provide an explanation for an advantage in recall for common items. Although promising, clearly, the foregoing account is a post hoc accommodation of the data. To evaluate its fruitfulness, we conducted the next three experiments to more analytically test the differential-retrieval-process account of the bizarreness effects.

Experiment 2

In this experiment, we focused on pure lists of common items and bizarre items. The critical feature was the inclusion of a condition that strongly encouraged participants to encode order information and use that information to guide retrieval (by explicitly instructing participants to memorize the serial order of the presented material and use this order information during recall). Our intent in the serial-instruction groups was to produce better encoding of order information directly without also varying list composition and to force the use of order information in recall for both common and bizarre items. One limitation of previous studies supporting the general order-encoding framework is that the encoding of order information was not directly manipulated. Instead, order memory was observed to vary with list composition, and the observed variation in order memory was assumed to be related to recall levels. In the current experiment, we reasoned that by comparing recall of the serial-instruction groups with recall for groups given standard encoding and recall instructions, we could evaluate whether order information can play a role in improving recall. More specifically, according to both the differential-retrieval-process view and the order-encoding view, common items are better recalled from pure lists than are bizarre items because recall of common items relies in part on the encoding and use of order memory, whereas recall for bizarre items typically does not. (Again, the distinctiveness view alone offers no explanation for the expected superior free recall of common items.) If this approach has merit, then explicit instructions to encode and use serial-order information should improve both order memory and recall for bizarre items relative to the standard-instructions condition. Further, such improvement would not necessarily be expected for common items.

Importantly, the differential-retrieval-process view and the order-encoding view make competing predictions for the level of recall for bizarre items relative to common items in the serial-instruction condition. The differential-retrieval-process view predicts that free recall will not differ across bizarre and common items (in the serial-instruction condition). According to this account, observed recall advantages for bizarre items reflect distinctiveness arising from the mixed-list retrieval context. Thus, the serial instructions, by eliminating the pure-list difference in order memory, would eliminate the common-item advantage in recall but would still not produce a bizarre-item advantage (because pure lists do not provide a context in which bizarre items are distinctive). In contrast, in the order-encoding view, the bizarreness effect is in part due to more involved, elaborate, or effortful encoding of bizarre sentences (Wollen & Margres, 1987; see also Worthen et al., 2000) that confers an advantage in recall when the use of order (relational) information is relatively equated across bizarre and common items, as it usually is in mixed lists (DeLosh & McDaniel, 1996). The straightforward prediction from this view is that when order information is as well encoded for bizarre as for common material, and when that order information is used at retrieval (serial-instruction condition), bizarre material should now show a significant advantage in free recall relative to common material in pure-list situations. The following experiment directly tested these competing predictions.

Method

Participants and design. Eighty Purdue University undergraduates completed the experiment in partial fulfillment of an introductory psychology course. The two independent variables of interest were sentence type (common vs. bizarre) and encoding condition (standard vs. serial encoding). These variables were manipulated using a 2 × 2 between-subjects design, with 20 participants randomly assigned to each of the factorial conditions. For this experiment, we limited our examination to pure lists. Each participant was tested individually in a 30-min session using an IBM-compatible computer.

Materials and procedure. The current experiment used the same stimulus materials that were used in the pure-list conditions of Experiment 1. For the standard-instruction condition, all procedures were identical to those of the previous experiment. However, for the serial-instruction condition, several procedural components were altered in an attempt to encourage the encoding of order information (regardless of sentence type). First, the initial task instructions emphasized that the critical feature of the experiment was memory for serial order. Second, participants were explicitly instructed to memorize the presentation order of sentences during encoding. Third, to reinforce the importance of remembering presentation order, we instructed participants to recall items in serial order. That is, participants were given serial-recall instructions rather than free-recall instructions (although we scored performance as though it were a free-recall task). The serial-instruction condition was identical to the standard-instruction condition in all other respects.

Results

Order reconstruction. Scores representing the mean proportion correct in order reconstruction were submitted to a 2 × 2 between-subjects ANOVA with instructions and sentence type as the independent variables. A significant effect of sentence type was obtained such that order memory was better for common sentences (.65) than for bizarre sentences (.57), $F(1, 76) = 6.04$, $MSE = 0.02$. A marginally significant interaction was also observed, $F(1, 76) = 2.86$, $MSE = 0.02$, $p = .09$. Tukey's HSD tests revealed that order reconstruction was more accurate for common sentences than for bizarre sentences in the free-recall condition, with no difference in the serial-recall condition. It is also noteworthy that serial-order instructions significantly boosted order reconstruction for lists of bizarre sentences (relative to standard free-recall instructions) but did not improve order reconstruction for lists of common sentences (see Table 1 for means).

Free recall. Next, a 2 × 2 ANOVA was conducted using the proportions of items correctly recalled. This analysis yielded a reliable interaction between instructions and sentence type, $F(1, 76) = 4.23$, $MSE = 0.01$. Tukey's post hoc comparisons showed that there was a free-recall advantage for common sentences in the free-recall condition but no difference in the serial-recall condition. This reflects a dynamic whereby serial instructions significantly improved recall for bizarre sentences but did not improve recall for common sentences (see Table 1 for means).

Input–output correspondence. A 2 × 2 ANOVA was also conducted using Asch–Ebenholtz (1962) scores for input–output correspondence. This analysis revealed a main effect of sentence type such that presentation order was better preserved in the recall protocols of common sentences (.72) than in the recall protocols of bizarre sentences (.65), $F(1, 76) = 5.94$, $MSE = 0.02$. The interaction between instructions and sentence type was not significant ($p = .22$). It is nonetheless noteworthy that the common-item advantage in input–output correspondence was driven primarily by differences in the free-recall condition. Tukey's post hoc comparisons indicated that input–output correspondence was significantly greater for common sentences than for bizarre sentences in the free-recall condition, but this difference was reduced and was not significant in the serial-recall condition (see Table 1 for means).

Discussion

In the present experiment, one group of participants was instructed to attend to and use serial-order information for recall of bizarre material. In this group, order memory for bizarre material improved relative to the group wherein order encoding was not required. Under these conditions in which order memory was improved without concomitant changes in list composition or materials, recall of bizarre material showed significant improvement (relative to recall of bizarre material in the standard encoding condition). Thus, we have obtained strong evidence that the encoding and use of order information can be beneficial to recall performance. This finding fits well with the observation that serial-order information can be an important attribute in long-term episodic memory performance (Burns, 1996; DeLosh & McDaniel, 1996; Mandler & Dean, 1969; Nairne, 1990; Nairne et al., 1991; Toglia & Kimble, 1976). We suggest that with lists of unrelated items, serial-order information can serve to relate list items, thereby providing an organizational structure to guide retrieval. This suggestion fleshes out the relational-individual-item approach to recall of events (lists) that have no semantic, taxonomic, or other relational information to structure retrieval (cf. Hunt & McDaniel, 1993).

It is also noteworthy that the serial instructions did not boost order memory for common sentences. This supports the assumption of the order-encoding framework that order information is routinely encoded for pure lists of common material when standard instructions are used (DeLosh & McDaniel, 1996). Nor was recall for common sentences improved in the serial-instructions group. Therefore, it is not the case that serial-recall instructions, per se, improve memory (possibly because of greater effort exerted during recall). The pattern across common and bizarre materials indicates that the encoding of serial-order information and the use of that information relates to recall levels. This result supports a central assumption of the order-encoding framework, an assumption also incorporated in the differential-retrieval-process account.

Finally, the pure-list recall advantage for common sentences was found again under standard-recall instructions but was eliminated with serial instructions. We did not, however, observe a significant pure-list recall advantage for bizarre material, a result consistent with the differential-retrieval-process account but not with the order-encoding framework (DeLosh & McDaniel, 1996; Nairne et al., 1991). The order-encoding framework presumes that bizarre material enjoys enhanced individual-item memory relative to common material, which should confer a recall advantage when order memory is equivalent for bizarre and common material. In contrast, the differential-retrieval-process account assumes that the bizarreness advantage is obtained because bizarre items provide distinctiveness within a retrieval set consisting of bizarre and common items (mixed lists), and this distinctiveness is used to guide retrieval (Knoedler et al., 1999; McDaniel et al., 1995). With only bizarre items in the retrieval set (pure lists), there would be no distinctiveness for bizarre items and consequently no advantage for those items at retrieval.

The differential-retrieval-process approach also generates a strong and unique prediction regarding the boundary conditions for the robust bizarreness effect in mixed lists. This formulation implies that if retrieval in free recall were guided by dimensions other than distinctiveness, then the bizarre items would have no advantage over common items, and the bizarreness effect in mixed lists would consequently be completely eliminated in recall. Experiments 3 and 4 were conducted to test this prediction.

Experiment 3

In this experiment, we manipulated the type of item (bizarre vs. common) within lists (mixed lists), and, as in Experiment 2, we required participants in one condition to encode serial-order information and use that information to guide recall. There is no published study of which we are aware that has failed to find the bizarreness effect in mixed lists with the current experimental materials (except for one condition in McDaniel & Einstein, 1986, for which participants were required to rate the unusualness of the sentences), which reflects the provocative nature of the prediction that the bizarreness effects should be eliminated in the serial-instruction condition.

By our understanding, other accounts of the bizarreness effect would not expect the bizarreness effect to be eliminated because of instructions to encode and use order information. Indeed, the order-encoding account assumes that when order encoding for common and bizarre items is comparable, then the bizarreness effect will be obtained because of the additional item elaboration presumed to be attracted by bizarre items (DeLosh & McDaniel, 1996). Thus, the serial-instruction condition should be especially conducive to obtaining a bizarreness effect because it ensures that bizarre items are not deficient in order information. More generally, all views of the bizarre-imagery effect that assume that the effect is due at least in part to more elaborate or effortful encoding of the bizarre items (e.g., Hirshman, Whelley, & Pulij, 1989; Wollen & Margres, 1987) would seem to anticipate at least some advantage for bizarre items in recall. Also, views that assume that both effortful encoding and the distinctiveness of bizarre items contribute to the bizarreness effect (similar to Worthen et al.'s, 1998, hybrid account) do not necessarily anticipate elimination of the bizarreness effect in a serial-recall condition. The retrieval assumptions are sketchy in these views, but if one assumes that distinctiveness is used to help precisely recover or construct items that are partially identified through relational information like serial order (cf. Hunt & McDaniel, 1993), then basing recall on serial order should not preempt the advantage of bizarreness. On these views, then, there is no a priori basis for expecting the effect to be eliminated in mixed lists when serial-order information is incorporated into recall.

Method

Participants and design. Thirty-two Purdue University undergraduates completed the experiment in partial fulfillment of an introductory psychology course requirement. As in Experiment 2, the two independent variables of interest were sentence type (common vs. bizarre) and task instructions (standard vs. serial recall). Task instructions were manipulated between subjects with 16 participants randomly assigned to each condition, but sentence type was manipulated within subjects (i.e., mixed lists were used). Each participant was tested individually on a personal computer in a 30-min session.

Materials and procedure. The materials consisted of 24 of the 30 bizarre and common sentences used in the first two experiments. Three lists of eight sentences were constructed using the method described in Experiment 1, each list consisting of four bizarre sentences and four common sentences. Two sets of mixed lists were used, with half of the participants in each instruction condition receiving Set 1 and half receiving Set 2. The experiment followed the same general procedure that was used in the previous experiments, except reconstruction tests were not included, and participants were given 2.5 min for recall instead of 2 min (because lists were eight items in length instead of six). The instruction manipulation was implemented in the same fashion described in Experiment 2.

Results

Free recall. Scores representing the mean proportion of words correctly recalled were submitted to a 2 × 2 (Instructions × Sentence Type) mixed ANOVA. This analysis yielded a significant interaction between sentence type and list type, $F(1, 30) = 3.96$, $MSE = 0.01$. Tukey's post hoc comparisons showed that in the free-recall condition, recall was significantly better for bizarre sentences than for common sentences. Twelve participants showed a recall advantage for bizarre sentences, and 4 showed an advantage for common sentences. In the serial-recall condition, however, recall did not significantly differ as a function of sentence type. In this case, 6 participants better recalled bizarre items and 7 better recalled common items, and there were 3 ties (see Table 1 for means).

Input–output correspondence. Next we analyzed input–output correspondence as a function of task instructions. Because the current experiment used mixed lists, it was not possible to compute separate Asch–Ebenholtz (1962) scores for bizarre and common items so as to examine input–output correspondence as a function of sentence type. Our analysis was therefore limited to a one-way between-subjects ANOVA on Asch–Ebenholtz scores with task instructions as the sole independent variable. This analysis revealed a significant effect of task instructions, $F(1, 30) = 9.89$, $MSE = 0.02$, such that recall protocols better preserved presentation order in the serial-recall condition than in the free-recall condition (see Table 1 for means). Only performance in the serial-recall condition was significantly greater than chance, $t(15) = 4.13$, $SEM = 0.04$.

Discussion

The results confirmed the prediction of the differential-retrieval-process account that the usual bizarreness advantage in recall of mixed lists of common and bizarre items (replicated again in the standard-recall condition) would be eliminated if participants were encouraged to use a retrieval strategy that did not rely on distinctiveness. When participants were required to retrieve items on the basis of serial order, common items were recalled nominally better than were bizarre items. The rarity of this pattern is underscored by Einstein and McDaniel's (1987) review that indicated that with mixed lists, bizarre items were recalled better than were common items in 8 of 10 experiments (one of the experiments from McDaniel & Einstein, 1986, that failed to find the effect was noted in the introduction). Since then, Hirshman et al. (1989) and McDaniel et al. (1995) have reported nine additional experiments with mixed lists (composed of equal numbers of bizarre and common sentences like those used herein), eight of which revealed a significant bizarreness effect in free recall. (In the one experiment failing to find a bizarreness effect [Hirshman et al., Experiment 6], all sentences were presented in an unusual fashion with the word "bizarre" or "normal" inserted in each sentence.) Clearly, superior recall of bizarre items from mixed lists is a robust phenomenon. In line with this empirical fact, as developed in the introduction, no theoretical account of the bizarreness effect of which we are aware would necessarily anticipate that the bizarreness advantage would be eliminated in the serial-instruction condition. Thus, the differential-retrieval-process framework is especially attractive because it uniquely anticipated the unusual pattern of recall for mixed lists of bizarre and common items that was obtained in the present experiment.

Accounts of the bizarreness effect that posit increased item elaboration for bizarre items (see the introduction) cannot be completely ruled out, however. One post hoc interpretation of the current result is that the focus on serial-order information during encoding prevented the additional item elaboration that would usually be prompted by bizarre items.[1] If so, then the recall advantage for bizarre items would be eliminated. Experiment 4 further examined these ideas.

Experiment 4

We attempted to gain converging support for the idea that the bizarreness effect in mixed lists primarily reflects a retrieval process guided by the distinctiveness of the target items (with distinctiveness determined by the retrieval set). By this view, serial recall is not required to eliminate the bizarreness advantage. Encouraging other free-recall strategies that rely on information other than distinctiveness should also eliminate the bizarreness advantage. Accordingly, in this experiment, some lists were constructed such that the target nouns in the sentences were categorically related. Categorical information is thought to be used in guiding free recall (e.g., Einstein & Hunt, 1980; Hunt & McDaniel, 1993; Raaijmakers & Shiffrin, 1980) and has been found to be prominent in recall relative to other sources of information that might otherwise guide recall (e.g., see Nairne et al., 1991). Thus, the expectation from the differential-retrieval-process view is that the bizarreness effect will be eliminated for the categorized lists but not for uncategorized lists using the same materials (cf. Wollen & Margres, 1987).

It is important to note that participants were given encoding instructions for the categorized-list condition that were identical to those given in the uncategorized-list (control) condition. Further,

[1] We thank Dan Burns and Steve Schmidt for pointing out this possibility.

at encoding, participants were required to rate the bizarreness of each sentence so that the participants' attention was focused on the individual sentences, as well as on the unusual nature of the bizarre sentences. In short, the encoding task should favor increased item encoding for bizarre sentences regardless of list construction (categorized and uncategorized lists). Therefore, on accounts embracing the idea that the bizarreness effect is mediated by enhanced item encoding of bizarre material, either because of the unusual nature of the bizarre material, per se, (e.g., DeLosh & McDaniel's [1996] order-encoding framework; Worthen et al.'s [1998] hybrid model) or because of the relative "strangeness" of the sentences in the presence of common sentences (Wollen & Margres, 1987), the clear prediction is that the bizarreness effect should be obtained for both the categorized and uncategorized mixed (bizarre and common) lists. One other feature of this experiment is that longer lists were used than in Experiment 3, thereby producing lists that are more comparable in length with those typically found in the literature.

Method

Participants and design. Participants were 32 Colorado State University undergraduates who completed the experiment in partial fulfillment of an introductory psychology course requirement. List type (categorized vs. uncategorized) was varied between subjects. Sixteen participants were randomly assigned to each of the conditions. Each participant was tested individually on a personal computer in a 30-min session.

Materials. The materials were simple sentences with 2 nouns and 1 verb. Thirty-six nouns were selected from the Battig and Montague (1969) category norms, composed of the three most frequent responses from 12 categories. These nouns served as the subjects of the sentences and were the target words that participants were asked to recall. One bizarre and one common sentence were constructed from each of the 36 nouns by varying the relationship between the words in the sentence. For example, one common sentence was "The oak was by the river," and the corresponding bizarre sentence was "The oak swam up the river." Three mixed lists were constructed by assigning six common and six bizarre sentences to each list. The lists in the categorized-list condition were comprised of three sentences from four categories (with two categories comprised of bizarre sentences and two categories comprised of common sentences), whereas the lists in the uncategorized-list condition contained one sentence from each category (see the Appendix for the categorized lists). A second set of mixed lists was created by replacing common and bizarre sentences of the first set with their counterparts. Half of the participants received Set 1, and half received Set 2. As a result, the common and bizarre versions of each sentence appeared equally often.

Procedure. In the initial instructions, participants were told that they would be shown a series of sentences and that they should form and maintain a mental image of the scenario described by each sentence while it was being presented. After each sentence, participants rated the bizarreness of the sentence on a 5-point scale, with 1 representing a *very bizarre image* and 5 representing a *common or usual image.* Three lists of 12 items were given. The remainder of the procedure followed that used in the standard-instruction conditions of the previous experiments, except that participants were given 1 min to recall the target nouns.

Results

Bizarreness ratings. To verify the assumption that the differential bizarreness of the sentences was noticed and encoded equally well for both the categorized- and uncategorized-list conditions, a 2 × 2 (List Type × Sentence Type) mixed ANOVA was conducted on participants' bizarreness ratings. As expected, bizarre sentences (M = 1.66) were judged to be more bizarre than were common sentences (M = 4.32), $F(1, 30) = 529.25$, $MSE = 0.22$. Note that this pattern did not differ for uncategorized (Ms = 1.66 and 4.28) versus categorized lists (Ms = 1.65 and 4.36), $F < 1$.

Free recall. Next, we scored the mean proportion of target nouns (i.e., the subjects of the sentences) recalled by each participant.[2] These scores were submitted to a 2 × 2 (List Type × Sentence Type) mixed ANOVA. This analysis revealed a significant effect of list type such that recall was better for categorized lists (M = .59) than for uncategorized lists (M = .47), $F(1, 30) = 5.73$, $MSE = 0.04$. In addition, a significant interaction between list type and sentence type was observed, $F(1, 30) = 4.52$, $MSE = 0.01$. Tukey's post hoc comparisons showed that in the uncategorized-list condition, recall was significantly better for bizarre sentences than for common sentences. Eleven participants showed a recall advantage for bizarre sentences, 3 showed an advantage for common sentences, and there were 2 ties. In the categorized-list condition, however, recall did not significantly differ as a function of sentence type. In this case, 7 participants better recalled bizarre items, 6 better recalled common items, and there were 3 ties (see Table 1 for means).

Category clustering. To determine whether participants used a category-based retrieval strategy when categorized lists were used, we measured the prevalence of category clustering in participants' recall protocols. Specifically, we computed the adjusted ratio of clustering (ARC) index for each participant in the categorized-list condition (cf. Roenker, Thompson, & Brown, 1971). Note that an ARC score of 0.00 represents chance clustering, and a score of 1.00 reflects perfect clustering. The mean ARC score for the categorized-list condition was .67, with scores ranging from .18 to 1.00. This mean is significantly greater than that expected by chance, $t(15) = 8.76$, $SEM = 0.08$.

Input–output correspondence. Finally, we analyzed input–output correspondence as a function of the type of list (categorized vs. uncategorized). As with Experiment 3, it was not possible to compute separate Asch–Ebenholtz (1962) scores for bizarre and common sentences because of the present focus on mixed lists. Thus, we conducted a one-way between-subjects ANOVA on Asch–Ebenholtz scores with the type of list as the only independent variable. The effect of list type was not significant, $F(1, 30) = 0.01$, $MSE = 0.02$, and neither mean score was significantly greater than that expected by chance (see Table 1 for means).

Discussion

For uncategorized lists, the typical bizarreness advantage in free recall was obtained with new materials. This finding further rein-

[2] Although participants sometimes recalled the other nouns of the sentences, we felt it most appropriate to score the proportion of subject nouns that were recalled, because it is these nouns that participants were explicitly instructed to recall, and in the categorized-list condition, it is these words that belonged to categories. A similar pattern of results was obtained when all nouns were scored and analyzed: A marginally significant (p = .07) interaction between list type and item type was observed such that bizarre items were recalled better than were common items for uncategorized lists, with no significant difference in recall for categorized lists.

forces the robust nature of the effect in mixed lists. The literature has tended to replicate the mixed-list recall effect using only a few standard sets of materials, perhaps raising doubts about its generality (cf. Kroll, Schepeler, & Angin, 1986). This experiment demonstrates that the effect is not restricted to a small set of widely used stimuli.

More importantly for the present concerns, the elimination of the bizarreness effect in the categorized list further supports the differential-retrieval-process view. The encoding instructions required participants to attend to and form images of the individual sentences, and they directed the participants' attention to the differential bizarreness of bizarre and common sentences. Such encoding procedures presumably enhance individual-item encoding of bizarre items (Wollen & Margres, 1987). Further, the encoding instructions appeared to be followed equally in both uncategorized and categorized lists inasmuch as the pattern of bizarreness ratings did not differ across lists. Yet, the bizarreness effect did not emerge for categorized lists, lists that presumably encouraged a retrieval scheme other than one based on item distinctiveness. More specifically, the high category-clustering scores support the assumption that participants organized their recall of the categorized list according to categorical information.

Past work has shown that free recall is better when individual-item elaboration is encouraged in the context of a categorized list rather than an uncategorized list (McDaniel, Einstein, & Lollis, 1988). Thus, on the basis of the view that the typical bizarreness effects in mixed lists are based on enhanced encoding of bizarre items, the categorical information used for recall of the categorized lists would be expected to further boost bizarre-item recall, even more so than common-item recall. That is, on the basis of this view, a more robust bizarreness effect might be expected for the categorized lists. Clearly, this did not occur. A less striking finding that would still be consistent with the view just mentioned is that improved recall of the categorized lists (as was found) would maintain the bizarreness advantage observed in the uncategorized lists. This pattern did not emerge either. Instead, the categorized lists improved recall of common items more so than recall of bizarre items.

Finally, as evidenced by chance levels of input–output correspondences, order information did not appear to be used to guide retrieval of mixed lists of bizarre and common items, a finding consistent with the previous experiments. In sum, the pattern of results supports the view that the bizarreness effect is mediated by a retrieval process that relies on distinctiveness.

General Discussion

These results converge on the general idea that the complex pattern associated with the bizarreness effect reflects retrieval dynamics that vary depending on type of material (common vs. bizarre) and list composition (including mixed vs. unmixed and categorically related vs. unrelated). Before discussing this differential-retrieval-process view, we discuss the shortcomings of existing accounts of the bizarreness effect in light of the present findings.

One promising account of the different recall patterns found for mixed versus unmixed lists for a number of different stimulus manipulations is the order-encoding framework. This account of the bizarreness effect in particular assumes that for unrelated lists of items, serial-order information is used to help guide retrieval in a free-recall task. In pure lists, common items are recalled as well as or better than bizarre items because of differential encoding of order information favoring common items (i.e., order information is routinely encoded for common material but is disrupted for bizarre items). In mixed lists, order memory is equivalent across bizarre and common material; thus, the richer encoding of individual-item information for bizarre items is reflected as a corresponding advantage in free recall (DeLosh & McDaniel, 1996). The attractiveness of the order-encoding account is that it can explain the free-recall advantage for common material in pure lists (found in Experiments 1 and 2), which distinctiveness accounts cannot readily explain (e.g., McDaniel & Einstein, 1986).

The current patterns generally reinforce the assumption and the previous findings (McDaniel et al., 1995) that order information is differentially encoded as a function of item type in conjunction with list composition. Further, Experiment 2 supported the assumption of the order-encoding hypothesis that the encoding of order information for common items is fairly well accomplished through standard learning processes (at least for short, pure lists). When participants were explicitly instructed to commit the serial order of the items to memory, there was no improvement in order memory relative to those participants who were given standard instructions to learn the material. Experiment 2 also indicated that the reduced order encoding for bizarre items could be overcome if participants are instructed to memorize serial-order information.

More important, however, was the finding that serial-order information did not appear to be a consistent component of retrieval as assumed by the order-encoding account. On the one hand, under standard free-recall instructions, pure lists of common material produced output orders in recall that did preserve the original input order significantly above chance levels. On the other hand, lists with bizarre items generally did not display use of serial-order information in recall, unless the instructions required participants to recall the items in their original order (Experiments 2 and 3). The importance of the current result is that serial-order information apparently was not used to guide retrieval in free recall of the bizarre items even with presentation procedures that produced relatively high levels of order encoding for bizarre materials in Experiments 1 and 2 (on average, over half of the bizarre sentences were placed in their correct location on the order-reconstruction test). This finding establishes limitations to the basic premise of the order-encoding framework that serial-order information is used to guide retrieval of unrelated lists of items. Order information has been found to be associated with recall of low- and high-frequency words (DeLosh & McDaniel, 1996), however, so the order-encoding account may hold for some effects.

The important implication for present purposes is that the variations in order encoding across bizarre and common materials cannot completely account for the bizarre imagery patterns in free recall because the encoded order information does not appear to be involved in free recall of bizarre sentences. As discussed above, the order-encoding account still provides an understanding of the free-recall advantage of common sentences in pure lists. The outstanding issue is how bizarre sentences gain a recall advantage in mixed lists. One prominent and intuitively appealing idea outlined earlier is that bizarre items are recalled better than are common items by virtue of their enhanced encoding (because of greater attention, effort, or elaboration). This idea, however, does not easily explain the nearly equivalent recall for bizarre and common items in pure lists and the nominally higher recall for common items in mixed lists when order information is encoded and used in recall (Experiments 2 and 3). In these situations, a bizarreness advantage would be expected because the equivalent use of order information for bizarre and common sentences should allow the presumed additional elaboration of bizarre items to produce an advantage in recall for bizarre relative to common items. One might argue, however, that the serial-encoding instructions in Experiments 2 and 3 attenuated the individual-item elaboration normally attracted by bizarre items. Experiment 4 ruled out this possibility by requiring only the typical imagery-encoding task. In this standard-encoding situation, when relational information was available in the list (categorized-list condition), common items were again recalled nominally better than were bizarre items in a mixed list. If greater individual-item elaboration of bizarre items (which might especially occur in the presence of common items at encoding) were mediating the bizarreness advantage, then bizarre items should be especially well recalled in the related lists because they would enjoy the benefit of both relational information and item information to support recall (see Hunt & McDaniel, 1993). Yet, in direct contrast to this expectation and to the oft-reported findings of bizarreness effects in mixed lists, free recall of common items was slightly better than that of bizarre items.

The set of results reported in this study converges on the differential-retrieval-process framework that combines assumptions of the order-encoding view (for explaining the pure-list advantage of common sentences) and notions regarding distinctiveness. One key feature of this view is the assumption that the bizarreness advantage in free recall rests on dynamics operative at retrieval (see McDaniel et al., 1995; Riefer & Rouder, 1992; Waddill & McDaniel, 1998). The idea is that recall involves delimiting a search set (e.g., a list presented in a particular spatial–temporal context) and that within this set, item distinctiveness can be a dominant dimension determining recall (Knoedler et al., 1999; McDaniel et al., 1995). In mixed but not pure lists, bizarre items would clearly be distinct so that bizarre items would ordinarily be favored. However, the present framework also assumes that if information other than distinctiveness is provided to guide recall, then the bizarre items will lose their advantage. This idea is consistent with existing findings showing that when recall is cued, then the bizarreness effect in mixed lists is eliminated (Cornoldi & de Beni, 1999; Nappe & Wollen, 1973; Pra Baldi, de Beni, Cornoldi, & Cavedon, 1985; Wollen & Cox, 1981).

It is important that this study further illuminated the role of distinctiveness as a favored but not ubiquitous element in free recall. When serial order was required for participants' retrieval strategies (Experiment 3) and when categorical information was available to guide recall (Experiment 4), bizarre items were recalled nominally less well than were common items in mixed lists, lists in which bizarre items are considered distinctive (McDaniel et al., 1995; Schmidt, 1991). That is, greater distinctiveness of an item within a set of to-be-retrieved items does not compel better recovery of that item (cf. Knoedler et al., 1999). Retrieval strategies can emphasize other dimensions of information (e.g., serial order or categorical information in the present study), and doing so can render the bizarreness of an item nonfunctional in guiding recall.

To close on a more general point, our results suggest that a complete understanding of free recall includes the notion that mixed lists augment a distinctiveness dimension (e.g., the atypical nature of some of the list items) that is ordinarily relied on for retrieving particular items. Further, this distinctiveness dimension is not preeminent in recall, as use of alternative retrieval strategies can negate and possibly reverse the potential advantage enjoyed by distinctive items. Also, free recall of pure lists forces reliance on other dimensions (e.g., Knoedler et al., 1999; Neath, 1993), with common lists of items profiting from order information, at least in shorter lists.

References

Asch, S. E., & Ebenholtz, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *Journal of Psychology, 54,* 3–31.

Balota, D. A., & Neely, J. H. (1980). Test-expectancy and word frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 576–587.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs, 80*(3, Pt. 2).

Burns, D. J. (1996). The item-order distinction and the generation effect: The importance of order information in long-term memory. *American Journal of Psychology, 109,* 567–580.

Cornoldi, C., & de Beni, R. (1999). *How bizarre is bizarreness effect: Metacognitive beliefs and retrieval effects.* Unpublished manuscript, University of Padova, Padova, Italy.

DeLosh, E. L., & McDaniel, M. A. (1996). The role of order information in free recall: Application to the word-frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1136–1146.

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.

Einstein, G. O., & McDaniel, M. A. (1987). Distinctiveness and mnemonic benefits 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.

Engelkamp, J., & Dehn, D. M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 671–682.

Erlebacher, A. (1977). Design and analysis of experiments contrasting the within- and between-subjects manipulation of the independent variable. *Psychological Bulletin, 84,* 212–219.

Hirshman, E., Whelley, M. M., & Pulij, M. (1989). An investigation of paradoxical memory effects. *Journal of Memory and Language, 28,* 594–609.

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review, 2,* 105–112.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language, 32,* 421–445.

Knoedler, A. J., Hellwig, K. A., & Neath, I. (1999). The shift from recency to primary with increasing delay. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 474–487.

Kofka, K. (1935). Principles of Gestalt psychology. New York: Harcourt Brace.

Kroll, N. E. A., Schepeler, E. M., & Angin, K. T. (1986). Bizarre imagery: The misremembered mnemonic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 42–53.

Mandler, G., & Dean, P. J. (1969). Seriation: Development of serial order in free recall. *Journal of Experimental Psychology, 81,* 207–215.

McDaniel, M. A., & Einstein, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 54–65.

McDaniel, M. A., Einstein, G. O., DeLosh, E. L., May, D., & Brady, P. (1995). The bizarreness effect: It's not surprising, it's complex. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 422–435.

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. (1999). The effect of perceptual interference at encoding on organization and order: Investigating the roles of item-specific and relational information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 54–69.

Nairne, J. S. (1990). Similarity and long-term memory for order. *Journal of Memory and Language, 29,* 733–746.

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 702–709.

Nappe, G. W., & Wollen, K. A. (1973). Effects of instruction to form common and bizarre mental images on retention. *Journal of Experimental Psychology, 100,* 6–8.

Neath, I. (1991). *Context and distinctiveness in memory: Evidence for a two-stage process of recall.* Unpublished doctoral dissertation, Yale University.

Neath, I. (1993). Contextual and distinctive processes and the serial position function. *Journal of Memory and Language, 32,* 820–840.

Pra Baldi, A., de Beni, R., Cornoldi, C., & Cavedon, A. (1985). Some conditions for the occurrence of the bizarreness effect in free recall. *British Journal of Psychology, 76,* 427–436.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 207–262). New York: Academic Press.

Riefer, D. M., & Rouder, J. N. (1992). A multinominal modeling analysis of the mnemonic benefits of bizarre imagery. *Memory & Cognition, 20,* 601–611.

Roenker, D. L., Thompson, C. R., & Brown, S. C. (1971). A comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76,* 45–48.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition, 19,* 523–542.

Schmidt, S. R. (1994). Effects of humor on sentence memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 953–967.

Serra, M., & Nairne, J. S. (1993). Design controversies and the generation effect: Support for an item-order hypothesis. *Memory & Cognition, 21,* 34–40.

Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language, 26,* 589–607.

Toglia, M. P., & Kimble, G. A. (1976). Recall and use of serial position information. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 431–445.

Tzeng, O. J. L., Lee, A. T., & Wetzel, C. D. (1979). Temporal coding in verbal information processing. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 52–64.

Waddill, P. J., & McDaniel, M. A. (1998). Distinctiveness effects in recall: Differential processing or privileged retrieval? *Memory & Cognition, 26,* 108–120.

Wollen, K. A., & Cox, S. (1981). Sentence cuing and the effectiveness of bizarre imagery. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 386–392.

Wollen, K. A., & Margres, M. G. (1987). Bizarreness and the imagery multiprocess model. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 103–128). New York: Springer-Verlag.

Worthen, J. B., Garcia-Rivas, G., Green, C. R., & Vidas, R. A. (2000). Tests of a cognitive resource allocation account of the bizarreness effect. *Journal of General Psychology,* **127,** 117–144.

Worthen, J. B., Marshall, P. H., & Cox, K. B. (1998). List length and the bizarreness effect: Support for a hybrid model. *Psychological Research, 61,* 147–156.

Zucco, G., Traversa, A. B., & Cornoldi, C. (1984). Il ruolo di dettagli non essenziali e della loro predicibilita contestuale nella rievocazione di nemi di figure [The role of nonessential details and of their contextual predictivity in the recall of names of pictures]. *Ricerche di Psicologia, 4,* 44–58.

(Appendix follows)

Appendix

Categorized Sentences Used in Experiment 4

Set 1	Set 2
The dog was riding the bicycle.	The dog was chasing the bicycle.
The cat was pouring the syrup.	The cat knocked over the syrup.
The horse was building a fence.	The horse jumped over the fence.
The flute was played by the student.	The flute was chasing the student.
The drum was on the box.	The drum crawled out of the box.
The trumpet was held by the clown.	The trumpet was biting the clown.
The knife was on the counter.	The knife sat down at the counter.
The spoon fell behind the refrigerator.	The spoon jumped over the refrigerator.
The fork was on the napkin.	The fork was throwing the napkin.
The chair leaped out of the window.	The chair was next to the window.
The bed was making a quilt.	The bed was covered with a quilt.
The desk was reading a book.	The desk was piled with books.
The apple fell next to the gate.	The apple unlocked the gate.
The orange rolled out of the bag.	The orange jumped into the bag.
The grapes were dropped in the street.	The grapes were laughing in the street.
The shirt was vacuuming the rug.	The shirt was lying on the rug.
The pants escaped from the closet.	The pants were hanging in the closet.
The socks were taking a shower.	The socks were drying in the shower.
The doctor was holding the chart.	The doctor was kissing the chart.
The lawyer was leaning against the podium.	The lawyer was dancing on the podium.
The teacher came into the room.	The teacher exploded inside the room.
The oak swam up the river.	The oak was by the river.
The pine was jogging up the path.	The pine was next to the path.
The willow ran through the park.	The willow was in the park.
The car was parked by the swings.	The car was playing on the swings.
The bus drove up to the theater.	The bus sat down in the theater.
The train went through the field.	The train was crying in the field.
The axe was stuck in the tree.	The axe was talking to the tree.
The drill cut through the brick.	The drill fell in love with the brick.
The saw was left on the bench.	The saw ran off with the bench.
The tulip was yelling at the vase.	The tulip was in the vase.
The daisy was kicking the barrel.	The daisy was growing by the barrel.
The rose was eating the pillow.	The rose was placed on the pillow.
The fly set fire to the barn.	The fly flew through the barn.
The ant wrote in the cement.	The ant was crossing the cement.
The bee was throwing a rock.	The bee landed on the rock.

Received July 20, 1999
Revision received February 1, 2000
Accepted February 3, 2000 ■