

Is the bizarreness effect a special case of sentence reorganization?

SATOMI IMAI and CHARLES L. RICHMAN

Wake Forest University, Winston-Salem, North Carolina

The effects on memory of two types of bizarre sentences (rarely occurring, or atypical, and never occurring, or illogical) were investigated in comparison with those of common sentences. In Experiment 1, the effects of these three types of sentences, which were presented for 7 sec, were measured in mixed-list conditions. Evidence for the bizarreness effect (advantageous memorial effects of bizarre sentences over common ones) was found only with atypical and common sentences. In Experiment 2, the stimulus presentation time was 35 sec; free-recall and sentence-access performance were superior for the illogical sentences as opposed to the common and the atypical sentences. A proposal based on the assumption that subjects tend to spontaneously modify sentence structure is suggested.

Attempts to replicate the bizarreness effect—that is, superior performance in recall of bizarre, relative to common, plausible sentences—have not always been successful. Some researchers have suggested that bizarre sentences do not facilitate recall any more than common ones do (Kroll, Schepler, & Angin, 1986; Morris, 1978; Nappe & Wollen, 1973); yet several investigators have demonstrated the beneficial effects of bizarreness on recall (McDaniel & Einstein, 1986; Merry, 1980; Merry & Graham, 1978; Webber & Marshall, 1978). The equivocal findings regarding the bizarreness effect derive partly from the difficulty of manipulating independent variables (instructing subjects in the use of interactive, bizarre, or common images) and of controlling for other rival variables (effects of interaction, word attributes, and the use of imagery), as well as from methodological considerations (the bizarreness effect occurs in free-recall, mixed-list designs).

Apart from the mnemonic value of bizarre sentences, their theoretical importance has recently been emphasized with regard to imagery (Hirshman, Whelley, & Palij, 1989; Wollen & Margres, 1987). Differential effects of bizarre and common images on accessibility and recovery of constituents of memory traces, as well as on susceptibility to interference, are intriguing issues (Einstein, McDaniel, & Lackey, 1989), related to such concepts as the feeling of knowing, verbal and nonverbal codes, and the distinctiveness of encoding. Since investigators have developed theoretical arguments to account for the bizarreness effect and have attempted to integrate their data into memory models, a more precise definition of the nature of bizarreness is in order (e.g., see Wollen & Margres, 1987).

This article is based in part on a thesis submitted by S. Imai to the Department of Psychology, Wake Forest University, in partial fulfillment of the requirements for the Master's degree. We gratefully acknowledge the assistance and editorial comments of Robert C. Beck, Ron Cobb, Gilles O. Einstein, and Cathy Seta. This research was supported by the Wake Forest University Research and Publication Fund to C. L. Richman. Correspondence should be directed to Charles L. Richman, Department of Psychology, Wake Forest University, Winston-Salem, NC 27109.

There are many ways to classify external stimuli in terms of informational values; one variable that can be used for this purpose is frequency of occurrence (Imai & Richman, 1990). Some events occur in our experiences frequently, some rarely, and others never. We assume that the less frequently an event occurs, the stronger the violation of an exception. *Common* sentences thus designate events whose occurrence is highly probable, logical, and plausible. One can consider bizarre sentences, on the other hand, as being on a continuum that ranges from the representation of rare to that of impossible events. A rarely occurring event is implausible, yet it could occur (i.e., an atypical event). An impossible occurrence cannot occur; such an event is possible only in the world of the imagination (it is absurd, and maybe semantically senseless). For example, the sentence *the maid spilled ammonia on the table* represents a plausible, common event; *the maid licked the ammonia off the table*, an atypical, possible, yet seldom occurring event; and *the table dropped the maid out of the ammonia*, an illogical sentence that has zero probability of occurrence.

Researchers interested in bizarre imagery have combined atypical and illogical bizarre sentences in an arbitrary manner. Some investigators are inclined to use atypical sentences (e.g., McDaniel & Einstein, 1986). Others use sentences such as *shotgun measure building* or *table spoon blast lemonade* (Collyer, Jonides, & Bevan, 1972), which we consider an extreme on the bizarreness continuum.

Imai and Richman (1990) instructed three independent groups of subjects to rate 45 sentences on bizarreness, surprise, or the likelihood that the event could actually occur. Illogical sentences were judged to be more bizarre, surprising, and unlikely to occur in the real world than atypical ones; and atypical sentences were judged to be more bizarre, surprising, and unlikely to occur than common sentences.

EXPERIMENT 1

In Experiment 1, we examined the effects of common versus atypical and common versus illogical sentences on

three measures of recall performance. We intended to assess the bizarreness effect by defining bizarreness on the probability-of-occurrence dimension.

Method

Design. Forty-eight subjects were randomly assigned to three groups. The materials for each group consisted of (1) six atypical (A) and six illogical (I) sentences (Group A-I), (2) six atypical (A) and six common (C) sentences (Group A-C), and (3) six illogical (I) and six common (C) sentences (Group I-C). The following are examples of the common, atypical, and illogical sentences used:

1. (C) The GOLDFISH was swimming in the BOWL next to the SOFA.
 (A) The GOLDFISH was eating out of the BOWL on the SOFA.
 (I) The SOFA was eating the BOWL under the GOLDFISH.
2. (C) The MAID spilled AMMONIA on the TABLE.
 (A) The MAID licked the AMMONIA off the TABLE.
 (I) The TABLE dropped the MAID out of the AMMONIA.
3. (C) The CAPTAIN yelled at his CREW from the MAST.
 (A) The CAPTAIN tied the CREW to the MAST.
 (I) The MAST looked at the CAPTAIN through the CREW.

The sentences in each list were arranged in one of two randomized orders, with the restriction that no more than two types of sentences be juxtaposed in the sequence. Since primacy and recency effects of recall require control in immediate testing situations, the serial order positions of each type of sentence were counterbalanced between the different types of sentences. The subjects were tested individually.

Procedure. Each subject was instructed that the purpose of the study was to assess individual differences in the creating of images. The subject sat in front of a computer terminal, which presented 12 sentences on the screen at a fixed rate of 1 per 7 sec. The three nouns in each sentence appeared in uppercase. The subject was asked to create clear (vivid) and interacting images from each sentence and then to rate the clearness (vividness) of the images on a scale of 1 = *unclear images* (poorly formed nonvivid images) to 5 = *extremely clear images* (vivid) by pressing the appropriate number key on the keyboard, within 4 sec. Two practice sentences were given prior to the experimental sentences. Upon completion of presenting the 12 sentences, each subject performed a 3-min distraction task (counting backwards from 300 by 3s). The subject was then given 4 min to write down the nouns that had appeared in uppercase.

Results

Vividness ratings. In the A-C condition, the effect of sentence type was significant [$t(15) = 8.40, p < .001$]. Images from common sentences were rated more vivid than those from atypical sentences. In the I-C condition, images from common sentences were rated more vivid

than those from the illogical sentences [$t(15) = 10.93, p < .001$]. In the A-I condition, there was also an effect of sentence type [$t(15) = 8.32, p < .001$]. Images from atypical sentences were rated more vivid than ones from the illogical sentences (see Table 1).

Percent of nouns recalled. *T* tests were performed separately on the free-recall, sentence-access, and items-per-sentence scores for each of the three conditions. In the A-C condition, there was a marginally significant effect [$t(15) = 1.75, p < .10$]. Recall from atypical sentences tended to be better than recall from common sentences. In the A-I and I-C conditions, sentence type was nonsignificant ($ts < 1.00$).

Sentence access. Sentence access was recorded if a subject recalled at least one noun from the sentence. In the A-C condition, there was a significant effect [$t(15) = 2.23, p < .05$]. More atypical sentences were accessed than common sentences. There were no significant effects in the I-C or the A-I list conditions ($ts < 1.00$). (See Table 1.)

Items per sentence. The items-per-sentence measure assesses recovery of constituents of the sentence—that is, whether a subject is more likely to recall more than one noun in the sentences if a noun is recalled. *T* tests performed separately for each condition resulted in statistically nonsignificant effects for all conditions ($ts < 1.00$).

Discussion

In Experiment 1, we examined how different types of sentences affected memory performance. When atypical and common sentences were presented in a mixed list, more atypical sentence nouns tended to be recalled than common sentence nouns. The superior recall of atypical sentence nouns was due to the fact that more atypical sentences were retrieved than common sentences.

The number of subjects who reported forming vivid images of illogical sentences was significantly lower than that for the other two types of sentences. It is quite possible that the 7-sec presentation time was inadequate for subjects to form images of the illogical sentences; this may have been the factor mediating the inferior recall of illogical sentences. McDaniel and Einstein (1986) found the bizarreness effect with a 7-sec presentation time, but their bizarre sentences were atypical.

If bizarre sentences are on a continuum from least to most bizarre and McDaniel and Einstein's (1986) stimuli are only moderately bizarre, then subjects should require even more time to generate images based on illogical sentences.

Table 1
Vividness Ratings and Mean Proportion of Items Recalled, Mean Number of Sentences Accessed, and Mean Number of Items Per Sentence Recalled as a Function of Type of Sentence for 7-sec, Experiment 1, and 35-sec, Experiment 2, Presentations

A-C				I-C				A-I															
Vividness		Recall		Sentence Access		Items		Vividness		Recall		Sentence Access		Items									
7	35	7	35	7	35	7	35	7	35	7	35	7	35	7	35								
Common																							
4.10	4.20	.35	.35	.45	.49	2.37	2.20	4.30	4.40	.42	.45	.52	.50	2.35	2.50								
Atypical																							
2.70	2.20	.45	.46	.59	.60	2.20	2.21					3.30	4.00	.40	.49	.52	.58	2.23	2.64				
Illogical																							
								1.80	1.60	.42	.61	.57	.73	2.14	2.50	1.90	2.40	.32	.60	.50	.71	1.88	2.51

Note—The "7" and "35" column headings refer to the 7-sec and 35-sec sentence-presentation times in Experiment 1 and Experiment 2, respectively.

Pra Baldi, De Beni, Cornoldi, and Cavedon (1985) assessed the effects of semantically unpredictable versus common sentences on uncued recall. Their unpredictable sentences were simpler than our illogical sentences (only two nouns per sentence) and most were more plausible than ours (e.g., *the refrigerator lights the table*; an open refrigerator door can indeed light a kitchen table). Pra Baldi et al. did not find the bizarreness effect with an uncued complete sentence recall measure but did report the effect when the total number of nouns was assessed. Pra Baldi et al.'s results suggest that if bizarre sentences fall within the range of our atypical and illogical sentences and a short intersentence interval is employed, the bizarreness effect is augmented. Hirshman et al. (1989) suggest that highly bizarre sentences decrease the probability of elaborative encoding. However, providing subjects with sufficient encoding time may allow the subjects time to develop the elaborative encodings required to produce the bizarreness effect. Consistent with this analysis is Wollen and Margres's (1987) suggestion that subjects require more time to form bizarre, relative to common, images. Furthermore, Kroll and Tu (1988) found a borderline bizarreness effect with atypical versus common sentences when subjects were given 5–7 sec to process each sentence; 10 sec, however, produced a reliable effect. We therefore reasoned that the present procedure of allowing subjects only 7 sec to generate images may have been an inadequate test of the bizarreness effect with illogical sentences.

EXPERIMENT 2

To test the hypothesis that highly bizarre sentences require more processing time than common or atypical ones do, we used illogical, atypical, and common sentences in mixed lists with 35-sec sentence presentations. Unmixed-list groups were not used, since this procedure consistently fails to produce the bizarreness effect with atypical sentences (see, e.g., Cobb & Richman, 1990; McDaniel & Einstein, 1989) and illogical ones (e.g., Cobb & Richman, 1990; Imai & Richman, 1990).

Method

Subjects. Thirty-four undergraduate students at Wake Forest University participated in this study as a partial fulfillment of their introduction to psychology course requirements.

Design and Procedure. The subjects were divided into three mixed-list conditions (Group A–C), 12 subjects; Group I–C, 11 subjects; and Group A–I, 11 subjects). The materials, presentation orders, and general procedures were the same as those in Experiment 1, with the exception that each stimulus sentence was presented on a screen via an overhead projector for 35 sec. Two or 3 subjects were tested at a time.

Results and Discussion

Table 1 shows a marginally significant trend for Group A–C in the predicted direction for both the free-recall and sentence-access measures [$t(11) = 1.44$ and 1.40 , respectively, $ps < .10$, one-tailed tests]. Atypical sentence recall tended to be better than common sentence recall. The absolute differences found here between atypical and common sentence recall are similar to those reported by others (Einstein et al., 1990; Kroll & Tu, 1988; McDaniel & Einstein, 1986). Our marginally significant effects were most likely due to small sample sizes.

The most intriguing findings occurred in Group I–C. Both the free-recall and the sentence-access data showed a reliable bizarreness effect [$t(10) = 1.98$ and 2.10 , respectively, $ps < .05$].

The Group A–I data showed that subjects tended to recall and access illogical sentences better than the atypical ones [$t(10) = 1.38$ and 1.44 , respectively, $ps < .10$, albeit one-

tailed tests]. The items-per-sentence results were statistically nonsignificant ($ts < 1.00$ for all group comparisons).

Vividness ratings. The vividness ratings within the three groups were quite similar to those in Experiment 1. As Table 1 suggests, the more logical the sentence structure, the higher the vividness ratings. Common sentences were rated as more vivid than atypical and illogical ones [for Group A–C, $t(11) = 5.62$, and for Group I–C, $t(10) = 16.00$, $ps < .05$]. Furthermore, the atypical sentences were rated more vivid than the illogical sentences in Group A–I [$t(10) = 7.35$, $p < .05$].

Comparisons between 7- and 35-sec presentations. As Table 1 indicates, recall performances were similar for common sentences at both 7- and 35-sec presentations ($ts < 1.00$). Furthermore, the recall performance also remained the same for the atypical sentences, regardless of presentation time ($ts < 1.00$). However, in both Group I–C and Group A–I the noun-recall performance of illogical sentences was inferior at 7-sec relative to 35-sec presentations [$t(25) = 2.54$ and 3.71 , $ps < .05$, respectively]. The sentence-access data derived from the illogical sentences produced superior recall at 35 sec as opposed to 7 sec; for example, in Groups I–C and A–I, $t(25) = 2.50$ and 3.08 , respectively ($ps < .05$). Facilitation of the items-per-sentence data via the extension of the presentation time was also evident for the illogical sentences [i.e., $t(25) = 2.07$ and 3.01 , $ps < .05$, for Groups I–C and A–I, respectively]. In summary, the noun-recall, sentence-access, and items-per-sentence measures were facilitated by lengthening the presentation times from 7 to 35 sec for the illogical sentences but not for the common or the atypical ones.

GENERAL DISCUSSION

The present results confirmed McDaniel and Einstein's (1989) finding that the bizarreness effect (superior recall of atypical, as opposed to common, sentences) tends to occur with brief presentation times. The bizarreness effect was not observed here when illogical sentences were mixed with common ones presented for 7 sec. This result is again similar to the results of McDaniel and Einstein (1989), who found little evidence for the bizarreness effect when complex bizarre sentences were mixed with common ones with short presentations. In the present study, however, 35-sec presentations produced the bizarreness effect when illogical and common sentences were mixed. The latter finding was due to the facilitation of noun-recall, sentence-access, and items-per-sentence scores of illogical sentences when presentations were increased from 7 to 35 sec, with no change in performance observed for the common or atypical sentences.

The present results were similar to those of Cornoldi, Cavedon, De Beni, and Pra Baldi (1988, Experiment 1). Cornoldi et al. gave subjects 30 sec to generate their own atypical, illogical, or common "images" from target nouns and found that noun recall was superior when subjects were asked to generate illogical relative to atypical or common "images." Noun recall was equivalent in the atypical versus common "image" condition. Note that subjects in the Cornoldi et al. study were not only instructed to generate images but also required to write sentences describing their images. Thus, Cornoldi et al.'s findings may have been due to image generation and/or sentence reorganization.

Cornoldi et al. (1988, Experiment 3) provides support for the notion that the processes involved in creating an image and in sentence construction or reorganization may have similar effects. Cornoldi et al. found a bizarreness main effect (illogical sentence recall was superior to common sentence recall, $p < .005$) but little evidence in support of a type of sentence (illogical vs. common) \times type of processing (image vs. sentence construction) interaction ($p > .10$).

Wollen and Margres's (1987) IMP model suggests that recall of atypical sentences is superior to recall of common sentences because subjects find the former more interesting and because atypical sentences demand more effort to elaborate and/or require longer schematic image searches than common sentences do. We agree with regard to illogical sentences as well. Illogical sentences require subjects to expend more effort and to take a longer period of time to search for relevant schematic images than atypical ones do. In the present study, recall performance for both common and atypical sentences was unaffected by extending the processing time. However, performance gains were noted when subjects were given 35 sec relative to 7 sec to process illogical sentences. The IMP model predicts that subjects should require an extended amount of time to complete their schematic image search when they are presented with illogical sentences.

With a concern for sacrificing explanatory power for parsimony, we cautiously offer an alternative interpretation of the bizarreness effect—a semantic analysis, without an appeal to imagery. Pylyshyn (1973) has argued that mental imagery experiments do not require a mental imagery interpretation. We suggest here that images may be generated by sentences, but that these images may have little direct effect on recall performance. The processes involved in generating images, propositions, or linguistic networks that are related to memory—not the image, proposition, or linguistic outcome itself—improve memory (Marschark, Richman, Yuille, & Hunt, 1987). Cobb and Richman (1990) have reported the bizarreness effect for subjects who were instructed to create images from sentences as well as subjects who were not instructed to image. Furthermore, they found nonsignificant correlations between their subjects' report of images during retrieval and the bizarreness effect.

The illogical sentences employed in the present study could best be described as sentence anagrams. Semantic relationships were violated in every sentence by pairing inappropriate noun-noun and noun-verb sequences. The illogical sentences and the phrases within each sentence were assumed to be more distinct (phrase-specific) and less integrated than the atypical ones. Distinctiveness is usually defined relatively, the uniqueness of a stimulus being determined by the relationship of that stimulus to other items within its learning context (Einstein & McDaniel, 1987; Jacoby & Craik, 1979).

To create meaningful sentences out of illogical ones, subjects may have to rearrange the entire sentence structure and/or replace the verbs and propositions with new ones. Here subjects were required to generate phrase relationships that took considerable time and cognitive effort. The generation of "new" sentences may also increase the probability of encoding variability that facilitates vertical retrieval. Traditionally, encoding variability referred to distributed repetitions of items that promote the encoding of the items in more than one context. As the separation between the to-be-remembered items increases, the probability of differences in the encoding context increases (Glass & Holyoak, 1986).

Encoding variability in the present context refers to subjects' initially processing the nominal phrases contained in each sentence and then processing the functional phrases via stimulus substitution and elaborative strategies in order to create meaning. Subjects in a 35-sec time period may have generated several meaningful phrases from each nominal phrase, thereby increasing the network of codes associated with each sentence. When subjects replace verbs and rearrange the nouns within a sentence in order to create a proper sentence, they are initiating integrative, relational processes that serve to facilitate horizontal retrieval. The findings of Marshall, Nau, and Chandler (1980) suggest that a positive relationship exists between the number of transformations within a sentence and recall performance. We further suggest that this relationship may occur when subjects spontaneously improve the integration of a sentence.

It is of some interest that the vividness ratings for common atypical and illogical sentences were not affected by time. This suggests that when subjects are given a mixed list, they assess the vividness, bizarreness, and frequency of occurrence (and perhaps meaningfulness) of the nominal and not the modified sentence.

REFERENCES

- COBB, R., & RICHMAN, C. L. (1990, June). *Is imagery a special case of semantic processing?* Paper presented at the 2nd annual meeting of the American Psychological Society, Dallas, TX.
- COLLYER, S. C., JONIDES, J., & BEVAN, W. (1972). Images as memory aids: Is bizarreness helpful? *American Journal of Psychology*, *85*, 31-38.
- CORNOLDI, C., CAVEDON, A., DE BENI, R., & PRA BALDI, A. (1988). The influence of the nature of material and of mental operations on the occurrence of the bizarreness effect. *Quarterly Journal of Experimental Psychology*, *40A*, 73-85.
- EINSTEIN, G. O., & MCDANIEL, M. A. (1987). Distinctiveness and the 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.
- EINSTEIN, G. O., MCDANIEL, M. A., & LACKEY, S. (1989). Bizarre imagery, interference, and distinctiveness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 137-146.
- GLASS, A. L., & HOLYOAK, K. J. (1986). *Cognition* (2nd ed.). New York: Random House.
- HIRSHMAN, E., & WHELLEY, M. M., & PALIJ, M. (1989). An investigation of paradoxical memory effects. *Journal of Memory & Language*, *28*, 594-609.
- IMAI, S., & RICHMAN, C. L. (1990). *Bizarreness: An initial approach to definition*. Paper presented at the meeting of the Midwestern Psychological Association Meeting, Chicago, IL.
- JACOBY, L. L., & CRAIK, F. I. M. (1979). *Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery processing in human memory*. Hillsdale, NJ: Erlbaum.
- KROLL, N. E., SCHEPLER, E. M., & ANGIN, K. T. (1986). Bizarre imagery: The misremembered mnemonic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *12*, 42-54.
- KROLL, N. E., & TU, S. F. (1988). The bizarre mnemonic. *Psychological Research*, *50*, 28-37.
- MARSCHARK, M., RICHMAN, C. L., YUILLE, J., & HUNT, R. R. (1987). The role of imagery in memory: On shared and distinctive information. *Psychological Bulletin*, *102*, 28-41.
- MARSHALL, P. H., NAU, K., & CHANDLER, C. K. (1980). A functional analysis of common and bizarre visual mediators. *Bulletin of the Psychonomic Society*, *15*, 375-377.
- 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.
- MERRY, R. (1980). Image bizarreness in incidental learning. *Psychological Reports*, *46*, 427-430.
- MERRY, R., & GRAHAM, N. C. (1978). Imagery bizarreness in children's recall of sentences. *British Journal of Psychology*, *69*, 315-321.
- MORRIS, P. E. (1978). Sense and nonsense in traditional mnemonics. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 155-163). New York: Academic Press.
- 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.
- 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.
- PYLYSHYN, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*, 1-24.
- WEBBER, S. M., & MARSHALL, P. H. (1978). Bizarreness effects in imagery as a function of processing level and delay. *Journal of Mental Imagery*, *2*, 291-300.
- 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. 103-127). New York: Springer.