Bulletin of the Psychonomic Society
1990, 28 (3), 185-187

The bizarre sentence effect as a function of list length and complexity

CHARLES L. RICHMAN, JENNY DUNN, GREG KAHL, LISA SADLER, and KIM SIMMONS
Wake Forest University, Winston-Salem, North Carolina

Previous research has consistently shown that bizarre sentences are recalled better than common ones in mixed lists. It is unclear whether list length or sentence complexity affects these results. In the present study, we attempted to define the critical point at which the interacting conditions of list length and sentence complexity produce or fail to produce the bizarre sentence effect (BSE). We utilized a factorial design composed of two sentence-length and three mixed list-length conditions. The BSE was consistently found for each list length in the simple sentence condition but not in the complex sentence condition.

The bizarre sentence effect, the BSE, is defined by the superior recall of bizarre relative to normal (common) sentences. Controversy reigns, however, regarding the mnemonic value (Cox & Wollen, 1981; Kroll & Tu, 1988; McDaniel & Einstein, 1986), the theoretical import (Einstein, McDaniel, & Lackey, 1989; Hirshman, Whelley, & Palu, 1989; Wollen & Margres, 1987), and the empirical demonstrability (Einstein et al., 1989; Kroll & Tu, 1988) of the BSE. The most likely conditions under which the BSE occurs include sentence-recall or free-recall measures and mixed-list presentations (Wollen & Margres, 1987).

It is not clear, however, whether or not list length (number of sentences to be recalled) or sentence complexity have an effect on the BSE. For example, Einstein and his associates (Einstein & McDaniel, 1987; Einstein, McDaniel, & Cote, 1988) have consistently reported the BSE for short as well as long lists (sentences), whereas Kroll and Tu (1988) were unable to demonstrate the BSE with long lists. List length in the Kroll and Tu study (Experiments 1 and 2) was, as they suggested, confounded by their procedure of requiring subjects to rate the bizarreness of each sentence and/or by the subjects' spending too little time (5-7 sec) processing each sentence. McDaniel and Einstein (1986) suggested that rating bizarreness encourages semantic, not imaginary, processing, and Wollen and Margres (1987) asserted that bizarre image generation requires considerable time. Kroll and Tu (Experiment 3) reported the BSE when the sentence-presentation time was 10 sec and subjects rated sentences for vividness. Unfortunately, their tests only involved a short list (12 sentences). Under the ideal conditions of long presentation times and vividness rating procedures, Einstein and McDaniel (1987) reported the BSE, whereas Kroll and Tu (Experiment 5) found little evidence for the BSE with long lists. Thus, one of the primary purposes of the present study was to assess the BSE in conditions similar to those of Einstein and McDaniel and Kroll and Tu with short, medium, and long lists.

A second goal of the present study was to assess the effects of sentence complexity on the probability of obtaining the BSE. Hirshman et al. (1989) and McDaniel and Einstein (1989, Experiment 2) failed to find the BSE when they employed the complex sentences used by Kroll, Schepeler, and Angin (1986), but the BSE was evident with simpler sentences. No test of the possible interaction of sentence complexity and list length was provided in these two studies; thus, we assessed the possibility of a list length-complexity of sentence interaction. Here, we assumed that if the BSE were to be found with complex sentences, it would occur only with short lists.

Although the focus of this research was an attempt to resolve the empirical discrepancies found in the literature, we were also concerned with the theoretical implications regarding this resolution. The present study should provide evidence in support of or counterintuitive to the "distinctiveness hypothesis" and the "expectation-violation hypothesis." Einstein and McDaniel (1987) suggested that bizarre images are distinct relative to common ones because of the atypical relationships among the words within a bizarre sentence. When bizarre sentences are mixed with common ones, the former, in contrast to the latter, are distinct, resulting in the superior recall of the bizarre sentences via a facilitation of trace access (see Hirshman et al., 1989). Theoretically, the number of to-be-remembered bizarre and common sentences should not modify the probability of obtaining the BSE. Sentence complexity could have an effect on distinctiveness: adding two modifiers to each of three nouns per sentence may induce subjects to fragment these sentences. If the noun modifiers are logically related to the nouns, then these fragmented sentences (phrases) become normal, common, and less distinct, resulting in the reduction or elimination of the BSE. Hirshman et al. (1989) expectation-violation hypothesis is predicated on the assumption that bizarre

This research was supported by the Wake Forest University Research and Publication Fund. We thank Ron Cobb, Gilles Einstein, and Satomi Imai for their helpful comments. Requests for reprints may be addressed to Charles L. Richman, Department of Psychology, Box 7778, Reynolda Station, Wake Forest University, Winston-Salem, NC 27109.

sentences initiate emotional reactions that protect the items from interference. Recall for bizarre sentences should be inversely related to list length (number of sentences) because the emotional response cue associated with the nouns in bizarre sentences are overloaded. List length should have no effect on common-sentence recall. Adding modifiers to each noun should augment the subject's emotional reaction to bizarre sentences, thus eliminating the BSE.

METHOD

Subjects

The subjects were 78 introductory psychology students who received course credit for their participation. The subjects were tested in groups of 4, and the experimental sessions lasted approximately 30 min.

Materials

The to-be-remembered items consisted of two sets of 12 noun triplets each. Each noun triplet was presented in four types of sentences: simple common, simple bizarre, complex common, and complex bizarre. The two sets of simple sentences were derived from McDaniel and Einstein (1986). These 24 sentences were lengthened by inserting two preceding adjectives per noun, which produced our complex sentences. Sentences were presented using an overhead projector and transparencies, with each transparency containing one typed and centered sentence. The subjects recorded all responses on previously distributed answer sheets. A simple mathematics task was used as an intervening task.

Design

A factorial design composed of two sentence-complexity and three list-length conditions was utilized. The presentation of the material was counterbalanced so that, across subjects, each noun triplet was presented equally often in common and bizarre contexts. Sentences were presented in short, simple form and in long, modified form within 8, 16, or 24 sentences. In the 12/12 conditions, all sentences were presented to each subject. In both the 4/4 and the 8/8 conditions, the sentences were distributed among these two groups in an approximate, equal manner.

Procedure

The subjects were tested in groups of 4, with each intact group being randomly assigned to one of the conditions. The subjects were told that the purpose of the experiment was to investigate differences between individuals in imaging ability. For each sentence, they were asked to form a mental image of the sentence that included the three uppercase nouns. After viewing each sentence for 10 sec, the subjects were asked to rate the vividness of that image on a 5-point scale, with 5 being very vivid and 1 being not very vivid. Each trial began with a sentence centered on the overhead projector screen. The subjects were asked not to divert their attention from the screen for the full 10 sec, after which they were to record their vividness rating on the scale presented to them. After the last experimental trial, the subjects were required to work on a set of simple arithmetic problems. The subjects received more than enough of the problems to fill the 4-min retention interval. The subjects were then asked to write down as many of the uppercase nouns from the sentences as possible. They were allowed 5 min for this free-recall task. Before they were dismissed, the subjects were debriefed as to the true nature of the experiment.

RESULTS

In this study, three measures of memory were assessed: percentage of capitalized nouns recalled, percentage of sentences accessed (if at least one noun per sentence was recalled), and items per sentence (if one noun in a sentence was recalled, was at least one other of the nouns in the sentence recalled?). Each of these measures was subjected to a 2 (simple vs. complex sentences) × 3 (8 vs. 12 vs. 24 sentences) × 2 (bizarre vs. common sentences) analysis of variance. The former two measures were between-subject variables and the latter one was a within-subject factor.

The nouns recalled and the sentences-accessed measures revealed similar results. Free recall and sentence access in simple sentences were superior to complex ones [Fs(1,72) = 32.00 and 10.17, respectively, ps < .001]. There was also an inverse relationship between the percentage of nouns recalled/sentences accessed and list length [Fs(2,72) = 34.58/54.23, ps < .001]. The percent-recall scores for 8, 16, and 24 sentences was .59, .48, and .30, respectively, with each score differing at the .01 level of significance when assessed by post hoc HSD tests. The sentences-accessed measure revealed similar results: .69, .55, and .35, ps < .01, for 8, 16, and 24 sentences, respectively.

The most important finding of the present study was the statistically significant complexity × type of sentence interaction for the nouns-recalled and sentences-accessed measures [Fs(1,72) = 5.96 and 14.88, respectively, ps < .05]. Table 1 presents the mean percent-recall and sentences-accessed scores derived from this interaction. Identical results were found for both memorial measures. Bizarre simple sentence recall/access was superior to common simple, common complex, and bizarre complex sentence recall/access (ps < .05). The common simple versus common complex sentence recall/access comparison and the bizarre complex versus common complex recall/access comparison did not approach statistical significance (ps > .10).

The items-per-sentence measure showed that simple sentences were superior to complex sentences [Ms = 2.75 and 2.51, respectively, $F(1,72) = 58.49$, $p < .001$], and common sentences were superior to bizarre ones [Ms = 2.68 and 2.58, respectively, $F(1,72) = 4.54$, $p < .05$]. No other main effects or interactions were statistically significant for the items-per-sentence, nouns-recalled, or sentences-accessed measures (ps > .10).

Vividness ratings were obtained for each sentence; however, recording and subject errors prohibited analyses of these data.

DISCUSSION

The present research provides evidence that the BSE occurs in mixed list lengths ranging from 8 to 24 sentences, for both free-recall and sentences-accessed measures. There was no suggestion of the BSE with

Table 1
Percentage of Nouns Recalled (Sentences Accessed) as a Function of Sentence Complexity for Bizarre and Common Sentences

	Type of Sentences				
	Bizarre			Common	
	M	*SD*		*M*	*SD*
Simple Sentences	.59	.64	<	.48	.51
Complex Sentences	.35	.45	=	.40	.52

complex sentences of any length. The items-per-sentence measure showed that more common-sentence nouns were recalled than bizarre ones.

These data conform to predictions generated from the distinctiveness hypothesis (e.g., Einstein & McDaniel, 1987). According to this model, distorted relationships among objects are more distinctive, and therefore better remembered, than normal, nondistinctive relationships, and increasing the number of to-be-remembered sentences should have little effect on distinctiveness (Hirshman et al., 1989; McDaniel & Einstein, 1986).

The present list-length data are counterintuitive to Hirshman et al.'s (1989) expectation-violation model. Hirshman and Whelley predict that the puzzlement response associated with bizarre sentences should habituate, resulting in the elimination of the BSE. (Our data reveal the BSE with list lengths from 8 to 24 sentences.)

In marked contrast to the list-length manipulation, sentence complexity eliminated the BSE at all list lengths (i.e., no BSE was observed when two adjectives were added to each of the three nouns per sentence). Although speculative, it is assumed that adding two modifiers to each noun served to reduce the distinctiveness of the three phrases that composed each sentence. (For example, the bizarre simple sentence, "the biscuits screamed when the oven jumped out of the window," taken as a complete sentence is absurd.) However, the similar complex sentence broken down into three phrases is perfectly logical (e.g., "the hot, fluffy biscuit," "smoked, cracked oven," and "bright, glass window"). We assume that the distinctive coding of a sentence requires a limited number of words per sentence. Increasing the number of words per sentence beyond this point encourages subjects to break the sentence down into phrases. If these phrases are logical (the noun modifiers make sense), they are considered by the subject as being common and nondistinct. These findings are also consistent with the expectation-violation hypothesis. Adding adjectives to each noun per sentence serves to increase the logic of each phrase and reduce the puzzlement and surprise of the component parts of the sentence. The lack of a BSE with complex sentences may be due to the subjects' having too little time to employ semantic and/or imaginal elaboration processes (e.g., Kroll et al., 1986; Wollen & Margres, 1987). In a recent unpublished study, Imai and Richman (1989) found equal recall for common ("The maid licked the ammonia off the table") versus illogical ("The table dropped the maid out of the ammonia") sentences with 7-sec sentence presentations, but the BSE was found when presentation time was increased to 35 sec. It should be noted that common-sentence recall was not affected by lengthening the presentation time, whereas illogical sentence recall improved substantially by the increase in presentation time.

The items-per-sentence results showed a small but reliable advantage for common, relative to bizarre, sentences, suggesting an increase in accessibility to common, relative to bizarre, images. Here, the relational value of nouns within a common sentence aids recall. These results are consistent with imagery models of the bizarreness effect (e.g., Wollen & Margres, 1987). Within a bizarre sentence, image-generated images are not linked to the same degree as they are in common sentences. Thus, if subjects recall one image in a common, relative to a bizarre, sentence, they are more likely to recall other aspects of an integrated image generated by nouns within that sentence. The items-per-sentence measure is an assessment of image access in which the retrieval of the components of the imagined sentence may, as was shown here, produce an advantage for common, relative to bizarre, images (Pra Baldi, deBeni, Cornoldi, & Cevedon, 1985; Wollen & Cox, 1981; see also Einstein et al., 1989).

Recently, Einstein et al. (1989) suggested the use of Nelson's (1979) theoretical framework to explain the relationship between the BSE and distinctiveness (see also Hunt & Elliott, 1980). Distinctiveness is inversely related to the number of shared features between a perceived event and information in memory. Bizarre images share fewer relationships with information in memory than do common images; thus, the former are more distinct than the latter. Nelson assumed that the unique encoding of distinct events reduces the likelihood of interference derived from other encodings. Within the context of the present mixed lists, there are fewer shared relationships between the nouns in the bizarre sentences relative to those in common sentences. Furthermore, the bizarre simple sentences are less similar to information in long-term memory than are the events depicted in the simple common sentences. Adding appropriate/logical modifiers to each noun in the complex bizarre sentences (e.g., "the active, baby chimp" is a meaningful phrase) serves to reduce the distinctiveness of sentence fragments, resulting in poorer memory than for simple sentences. It is important to note that the noun-recall and sentences-accessed measures in bizarre, but not common, sentences were augmented by complexity, suggesting that the change in recall of bizarre sentences was due to an increase in commonness resulting from complexity.

REFERENCES

Cox, S. D., Wollen, K. A. (1981). Bizarreness and recall. *Bulletin of the Psychonomic Society*, **18**, 244-245.

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-Verlag.

Einstein, G. O., McDaniel, M. A., & Cote, N. (1988, November). *Factors controlling the bizarreness effect*. Paper presented at the annual meeting of the Psychonomic Society, Chicago.

Einstein, G. O., McDaniel, M. A., & Lackey, S. (1989). Bizarre imagery, interference, and distinctiveness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 137-146.

Hirshman, E., Whelley, M. M., & Palu, M. (1989). An investigation of paradoxical effects. *Journal of Memory & Language*, **28**, 594-609.

Hunt, R. R., & Elliott, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects upon retention. *Journal of Experimental Psychology: General*, **109**, 49-74.

Imai, S., & Richman, C. L. (1989). *Bizarreness: An initial approach to definition*. Unpublished manuscript. Wake Forest University, Winston-Salem, NC.

Kroll, N. E., Schepeler, 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.

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.

Nelson, D. L. (1979). Remembering pictures and words: Appearance, significance and name. In L. S. Cermak & F. I. M. Craik (Eds.), *Level of processing in human memory* (pp. 45-76). Hillsdale, NJ: Erlbaum.

Pra Baldi, A., deBeni, 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.

Wollen, K. A., & Cox, S. (1981). The bizarreness effect in a multitrial intentional learning task. *Bulletin of the Psychonomic Society*, **18**, 296-298.

Wollen, K. A., & Margres, M. G. (1987). Bizarreness and the imagery multiprocess modes. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 103-127). New York: Springer-Verlag.

(Manuscript received September 9, 1989.)