

Further constraints on the bizarreness effect: Elaboration at encoding

BRIDGET ROBINSON-RIEGLER
Purdue University, West Lafayette, Indiana

and

MARK A. McDANIEL
University of New Mexico, Albuquerque, New Mexico

In this study, we attempted to determine why the mnemonic benefit of bizarreness is not found with the use of complex sentences (i.e., those containing additional modifiers of nouns) as stimuli. Several explanations were investigated, including the idea that complexity reduces the imageability of the sentence and the idea that complexity itself is mnemonically beneficial. The results of four experiments favored the latter explanation. We suggest that the cues associated with the complexity of the sentence provide more effective or salient retrieval cues than do those associated with sentence bizarreness. Consequently, the mnemonic benefit of bizarreness appears to occur only with relatively impoverished encoding contexts (e.g., simple, unelaborated sentences).

Interest in the mnemonic effectiveness of bizarre imagery has existed for centuries, dating as far back as the ancient Greeks. The assumption that image bizarreness enhances memory continues to be common among many professional mnemonists, and bizarre imagery is often recommended as a memory aid (Lorayne & Lucas, 1974; Yates, 1966). Yet empirical evidence supports the mnemonic benefit of bizarre imagery only under certain conditions (McDaniel & Einstein, 1986; Pra Baldi, de Beni, Cornoldi, & Cavedon, 1985). One critical factor in determining the mnemonic effectiveness of bizarre imagery is the method used to test memory. Studies have shown positive effects of bizarre imagery in free recall, but not in cued recall or recognition (Cornoldi, Cavedon, de Beni, & Pra Baldi, 1988; McDaniel & Einstein, 1986; Pra Baldi et al., 1985; Wollen & Cox, 1981a, 1981b). Another important factor in determining the bizarreness effect is list structure (mixed lists, in which both bizarre and common sentences are presented, vs. unmixed lists, in which all sentences presented are bizarre or common). The majority of studies done with unmixed lists (Collyer, Jonides, & Bevan, 1972; Cox & Wollen, 1981; Marshall, Nau, & Chandler, 1980) have not resulted in superior retention for bizarre as opposed to common items (see Einstein, McDaniel, & Lackey, 1989, for an important exception). In contrast, the majority of studies done with mixed lists (McDaniel & Einstein, 1986; O'Brien & Wolford, 1982; Pra Baldi et al., 1985) have resulted in enhanced free recall for bizarre as opposed to common items.

The aforementioned constraints have been the focal point of a number of theoretical explanations of the bizarreness effect. McDaniel and Einstein (1986) have focused on the distinctiveness of bizarre items when such items are presented in the context of common items. Hirshman, Whelley, and Palij (1989) have focused on a surprise response associated with bizarre items which serves as an additional cue when retrieval is challenging (e.g., free recall). Finally, Wollen and Marges (1987) have focused on the imagery generation process, which is assumed to be different for common items than for bizarre items. These frameworks have been directed at explaining why mixed-list designs and free recall are necessary for obtaining the bizarreness effect. One recent finding is troubling for these perspectives, however. Kroll, Schepeler, and Angin (1986), using complex sentences as stimuli (i.e., those containing additional modifiers of nouns), did not find a recall advantage for bizarre items. This was true despite the use of both a mixed-list design and free recall testing.

McDaniel and Einstein (1989) found that eliminating the adjectives from these complex sentences produced the bizarreness effect. Likewise, adding adjectives to the simple sentences employed in most other studies on bizarre imagery failed to produce a significant bizarreness effect. These results directly demonstrate that complexity can eliminate the advantage for bizarre items, but they do not offer insight into why the bizarreness effect is not obtained with the use of complex sentences. Moreover, the current theoretical accounts of the bizarreness effect (mentioned above) do not anticipate this additional constraint, nor do they offer a clear post hoc interpretation of such a limitation.

The main objective of the present study was to attempt to determine why complexity eliminates the ad-

Correspondence should be addressed to B. Robinson-Riegler, Augsburg College, Psychology Department, 2211 Riverside Ave., Minneapolis, MN 55454.

vantage of bizarre imagery. In so doing, we hoped to sharpen the current accounts of the bizarreness effect and develop a clearer understanding of its robustness and importance. We consider two general classes of explanations for the failure to find a bizarreness effect for complex sentences. The first is that complexity disrupts the imaginal processing of the bizarre sentences, processing that may be associated with their mnemonic advantage (cf. McDaniel, Anderson, Einstein, & O'Halloran, 1989; McDaniel & Einstein, 1986). In Experiment 1, we addressed the possibility that the original set of complex sentences (from Kroll et al., 1986) might have reduced sentence imageability because of the inclusion of abstract adjectives to describe each of the target nouns. A more interesting variant of this general idea is that the complexity per se disrupts the imageability of the sentence by making it more difficult to maintain a clear image of each object detailed within the sentence. This possibility was examined in Experiments 2 and 3.

The second general class of explanations appeals to the idea that the complexity of the sentence itself provides a mnemonic benefit. Complex sentences contain additional adjectives to describe nouns. These modifiers may create more elaboration of the target nouns for both the bizarre and the common sentences. If this elaboration confers a mnemonic benefit on the common sentences, which are usually recalled poorly, then the mnemonic advantage of image bizarreness could be obscured. This possibility will be considered in Experiment 4.

EXPERIMENT 1

The bizarreness effect seems to depend on the use of imaginal processing at encoding. In several studies, a recall advantage for bizarre sentences has been found only when an imagery orienting task has been used (Cornoldi et al., 1988; McDaniel et al., 1989; McDaniel & Einstein, 1986). McDaniel and Einstein (1989) observed that some of the adjectives used in Kroll et al.'s (1986) complex sentences tended to be abstract in nature (i.e., "The *lazy arrogant* waiter serves a narrow white necktie") and therefore possibly more difficult to image. Consequently, the observance of a bizarreness effect may have been obscured by the inability to form an image of the target sentences.

In addition, Cornoldi et al. (1988) have made a distinction between *unusual* items (e.g., a waiter serving a tie), which are strange but possible, and *truly bizarre* items (e.g., a waiter driving a tie), which are strange and impossible. These authors argue that some of the failures to find a bizarreness effect in the literature could have resulted from the inclusion of unusual items in the bizarre imagery conditions. Examination of the sentences used by Kroll et al. (1986) revealed that some of the sentences included instances of unusual items, thereby reducing the possible benefits of bizarre imagery on subsequent recall.

To test whether the features just mentioned may have been the basis for the absence of a bizarreness effect

with complex sentences (Kroll et al., 1986), in the present experiment we contrasted Kroll et al.'s original sentence set with a revised set that included only concrete modifiers and truly bizarre interactions. If adjective concreteness and/or the degree of bizarreness were playing a role in modulating the bizarreness effect, a bizarreness effect would be expected for the revised sentences but not for the original sentences used by Kroll et al.

Method

Subjects. Thirty-two introductory psychology students at Purdue University participated to fulfill a course requirement.

Design and Materials. The design was a two-factor mixed factorial, with sentence set (original vs. revised sentences) as the between-subjects factor and sentence type (bizarre vs. common) as the within-subjects factor. Sixteen subjects were randomly assigned to each of the between-subjects conditions.

Two lists consisting of 12 sentences each (6 common and 6 bizarre sentences per list) were constructed for each experimental condition (see the Appendix). One list contained sentences selected from those used by Kroll et al. (1986; see their appendix). The sentences chosen were those for which both the bizarre and the common versions contained adjectives considered by the authors to be abstract and/or for which the bizarre version was unusual rather than bizarre. For the second list, these sentences were revised by changing the abstract adjectives to ones that were considered to be more concrete and by changing the interactions in the bizarre sentences from unusual to bizarre.

Each sentence contained two target nouns. Two versions of each list were constructed so that each target noun pair would appear once in a bizarre context and once in a common context. There were identical adjectives in both the common and the bizarre versions of each sentence containing the same target nouns. The bizarre and common contexts were achieved by altering the verb to produce the appropriate interaction. The revised sentences were equated in length to the original sentences. The original sentences averaged 10.71 words in length, and the revised sentences averaged 10.33 words in length.

Procedure. The subjects were tested according to an incidental learning procedure. They were told that the purpose of the experiment was to investigate individual differences in imagery ability and use. The subjects were asked to form a mental image of the scene suggested by each sentence as the sentence appeared on the computer screen. They were told to be sure to include in this mental scene the two nouns presented in capital letters (the target nouns). After forming an image from reading each sentence, they were asked to rate the vividness of that image on a 5-point scale, with 1 representing a *clear, vivid image* and 5 representing an *unclear, poor image*.

The 12 sentences were presented randomly, with the restriction that no more than 3 sentences of one type (common or bizarre) could appear consecutively. Each sentence appeared at the center of a CRT screen for 11 sec, during which the subjects were to form an image of the sentence. At the end of the 11-sec interval, the vividness rating scale appeared on the screen. The subjects were given 4 sec to rate the vividness of their image before the next sentence was presented.

After the subjects had completed the imaging task on all 12 sentences, they were given the Individual Differences Questionnaire on Verbal and Imaginal Ability (Paivio & Harshman, 1983). We expected that if the absence of a bizarreness effect with Kroll et al.'s (1986) sentences was related to difficulty in imagery, then the effect might emerge for "high imagers" (defined as people with a regular use of imagery and proficient imagery ability). Paivio and Harshman (1983) presented the questionnaire as consisting of six scales, from which we chose four: good verbal expression and fluency; habitual use of imagery; self-reported read-

ing difficulties; and vividness of dreams, daydreams, and imagination. The subjects were given 5 min to complete the questionnaire and were then given a free recall test in which they were asked to write down as many of the capitalized nouns from the sentences (the two target nouns) as they could remember.

Results and Discussion

For all the tests reported herein, the rejection level was set at .05. The proportion of target nouns recalled (see Figure 1) was subjected to a 2×2 mixed analysis of variance (ANOVA). There was a marginally significant main effect of sentence type [$F(1,30) = 2.93$, $MS_e = .02$, $p < .10$], with more nouns recalled from bizarre sentences than from common sentences.

The two-way interaction with sentence set was not significant [$F(1,30) = 2.13$, $MS_e = .02$, $p < .15$], nor was there a main effect of sentence set ($F < 1$). However, visual inspection of Figure 1 seems to indicate that nouns from bizarre sentences were recalled better than nouns from common sentences for the original sentences but not the revised sentences. A one-tailed t test

revealed that the bizarreness effect was significant for the original sentences [$t(15) = 1.91$] but was not significant for the revised sentences [$t(15) = 1.25$].

The emergence of a significant bizarreness effect with the use of the original sentences from Kroll et al. (1986) contradicts Kroll et al.'s finding of no such effect. This contradiction, in combination with the lack of a bizarreness effect for the revised sentences, indicates that the occurrence of a bizarreness effect with the use of complex sentences seems tenuous at best. The erratic nature of this bizarreness effect contrasts sharply with the consistency of the effect found when one employs simple sentences. Why this is the case remains a question.

The average vividness ratings (see Figure 1) were also subjected to a 2×2 mixed ANOVA. The two-way interaction with sentence set was not significant. In addition, the revised sentences were not rated as significantly more vivid than the original sentences.

Subsidiary analysis. Given that imaginal encoding may be a critical factor in the occurrence of a bizarre imagery effect, it is possible that the effect could be mediated by individual differences in imagery ability. That is, only people who are considered "high imagers" would significantly benefit from the use of bizarre imagery with complex sentences. In order to test this idea, the Individual Differences Questionnaire on Verbal and Imaginal Ability (Paivio & Harshman, 1983) was scored and the two scales dealing with imagery were added together to produce one imagery score for each subject. Subjects above the median were classified as "high imagers" and subjects at or below the median were classified as "low imagers." A $2 \times 2 \times 2$ (sentence type \times ability \times bizarreness) mixed ANOVA of the proportion of target nouns recalled (see Figure 1) showed no significant main effects or interactions (all F s < 1), indicating that individual differences in imagery ability were not masking a bizarreness effect.

The absence of a bizarreness effect with the revised sentences, in combination with the fact that the revised sentences were not rated as more vivid than Kroll et al.'s (1986) sentences, suggests that the imageability of the adjectives and the lack of truly bizarre interactions are not the bases of the puzzling nature of the effect of bizarreness on complex sentences.

EXPERIMENT 2

Beech and Allport (1978) suggest that during the incorporation of new objects into an image, items previously encoded begin to fade as the new objects are assimilated into the image. With complex sentences, many details must be incorporated into the image to reflect the specific descriptions produced by the adjectives. It is possible that during the integration of these details, the subjects' normal imaging strategy allows objects and details previously encoded to deteriorate as new details are added into the scene (thus preventing a complete and integrated image of the sentence's referent). Furthermore, Wollen and Magres (1987) have suggested that

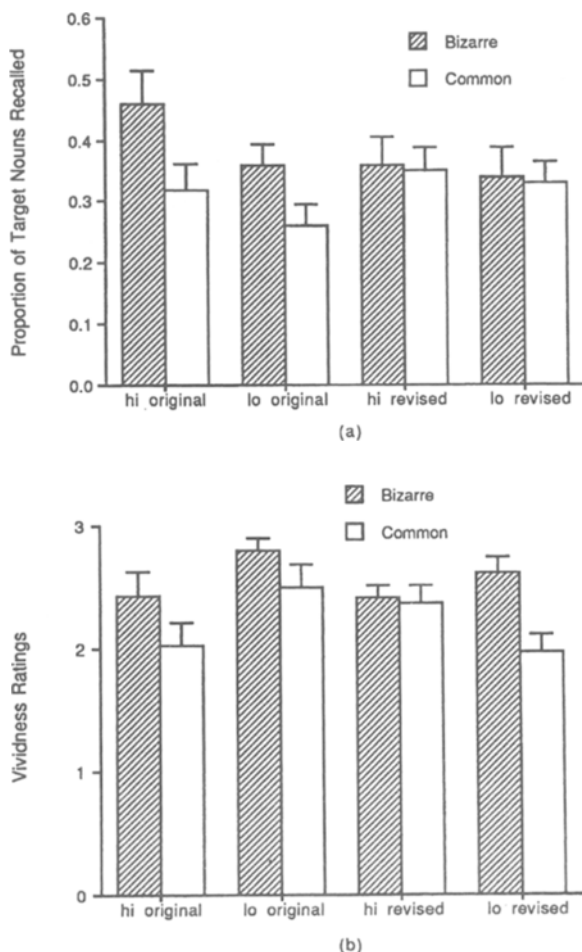


Figure 1. (a) The proportion of target nouns recalled and (b) the vividness ratings in Experiment 1, as a function of sentence type (bizarre or common), sentence set (original or revised), and imagery ability (hi or lo).

bizarre images take more effort to form than common images do. The complexity of the sentence may exacerbate this difficulty. In addition, studies in which image formation times for complex items have been compared with those for simple items have shown that to form an image of a complex item takes longer than to form an image of a simple item (Beech & Allport, 1978; Hubbard & Stoeckig, 1988; Kosslyn, Reiser, Farah, & Fliegel, 1983), and Kline and Groninger (1991) have suggested that a bizarreness effect can be obtained with complex sentences when sufficient processing time is given. Simply put, the time allotted for image formation in Experiment 1 might have been insufficient for some subjects to complete the image suggested by these more complex sentences, and/or subjects may not have been able to form adequate images from these sentences spontaneously (especially the bizarre sentences). If so, the benefits of attempting bizarre image formation may have been eliminated (cf. Kroll & Tu, 1988).

Therefore, in Experiment 2, sentence processing was subject paced in order to ensure sufficient time for forming an image from a complex sentence. Moreover, we provided our subjects with specific imagery instructions in order to help prevent the possible deterioration of previously imaged details in the sentence. In accordance with Kosslyn and his colleagues (Kosslyn, Cave, Provost, & von Gierke, 1988; Kosslyn et al., 1983), who have suggested that image formation processes are sequential, subjects were instructed to form the images in a series of cumulative stages. The intent was to try to maximize, as much as possible, the opportunity for subjects to form and maintain an image from reading the complex sentences (especially the bizarre sentences).

Method

Subjects. Thirty-two introductory psychology students at Purdue University participated to fulfill a course requirement.

Design and Materials. The design was a two-factor mixed factorial, with imagery instructions (no guidance vs. experimenter guidance) as the between-subjects factor and sentence type (bizarre vs. common) as the within-subjects factor. Sixteen subjects were randomly assigned to the two between-subjects conditions.

Given that the focus of the present investigation was to delineate the nature of the interaction between complexity and bizarreness and not to investigate the influence of the particular sentences used, we combined both sets of materials in Experiment 2. Half the subjects in each condition received Kroll et al.'s (1986) sentences, and the other half received the revised sentences.

Procedure. The subjects were tested according to the same incidental learning procedure as that in Experiment 1. The subjects in the no-guidance condition were treated identically to those in Experiment 1. In the experimenter-guided condition, subjects were told that the sentences they were about to use to form images were rather long and complex, and that the imaging would therefore be fairly difficult. In order to assist the subjects in forming their images, the subjects were asked to form their images from the sentences in a series of stages. They were given an example sentence: "The large majestic eagle swoops down to land on the rotting wooden fence." First they were asked to form an image of the subject of the sentence (the large majestic eagle). Once they had a clear image of the subject, they were asked to image the subject performing the action of the sentence (the large majestic eagle

swooping down to land). Once they had a clear image of the subject performing the action of the sentence, they were asked to incorporate into the scene the object of the sentence (the large majestic eagle swooping down to land on the rotting wooden fence).

In both conditions, the subjects were given as much time as they needed to form the image, and they were instructed to hit the return key on completion. They were then asked to rate the vividness of the image on the same 5-point scale as that used in Experiment 1, taking as much time as they needed. After they entered their vividness ratings, the next sentence was presented. The time required to encode each sentence and form an image from it was recorded.

After completing the imagery task, the subjects were given the Purdue Spatial Visualization of Rotations Test. This test was given in another attempt to identify "high imagers." The subjects were given 5 min to work on the test, after which they were given a free recall test in which they were asked to write down as many of the capitalized nouns from the sentences as they could remember. Upon completion of the free recall test, the subjects were instructed to finish the mental rotation test.

Results and Discussion

The rejection level was again set at .05. The proportions of target nouns recalled (see Figure 2) were subjected to a 2×2 mixed factor ANOVA. This analysis revealed no main effects or interactions (all F 's < 1). This pattern indicates that the guided imagery instructions did not facilitate the occurrence of the bizarreness effect. Given the significant bizarreness effect found in Experiment 1 for the original sentence set, the two sentence sets were analyzed separately for the existence of a bizarreness effect. In contrast to the results of Experiment 1, a bizarreness effect was not found for the original sentence set ($M_{\text{biz}} = .53$, $M_{\text{com}} = .49$) or the revised sentence set ($M_{\text{biz}} = .35$, $M_{\text{com}} = .39$). In accordance with the means, the ANOVA revealed no main effects [$F(1,30) = 0.06$, $MS_e = .047$] or interactions [$F(1,30) = .47$, $MS_e = .047$].

The average vividness ratings and average image formation times (see Figure 2) were each subjected to a 2×2 mixed factor ANOVA. The analysis of the average vividness ratings showed the typical finding that bizarre images are rated as significantly less vivid than common images [$F(1,30) = 37.47$, $MS_e = .38$]. The two-way interaction with imagery instructions was not significant. In addition, images formed in the guided imagery condition were not rated as significantly more vivid than those in the no-guidance condition.

The analysis of the reaction time data showed that the overall reaction time to form an image was significantly longer in the guided imagery condition than in the no-guidance condition [$F(1,30) = 4.67$, $MS_e = 23,288,443$]. In addition, bizarre images took significantly longer to form than the common images [$F(1,30) = 26.25$, $MS_e = 2,370,553.41$]. There was also a significant interaction between these two variables [$F(1,30) = 8.34$, $MS_e = 2,370,553.41$]. Examination of the means reveals that the increased time in the guided imagery condition was most pronounced for the bizarre sentences. This result indicates that the imagery guidance instructions may have produced more extensive processing of the bizarre

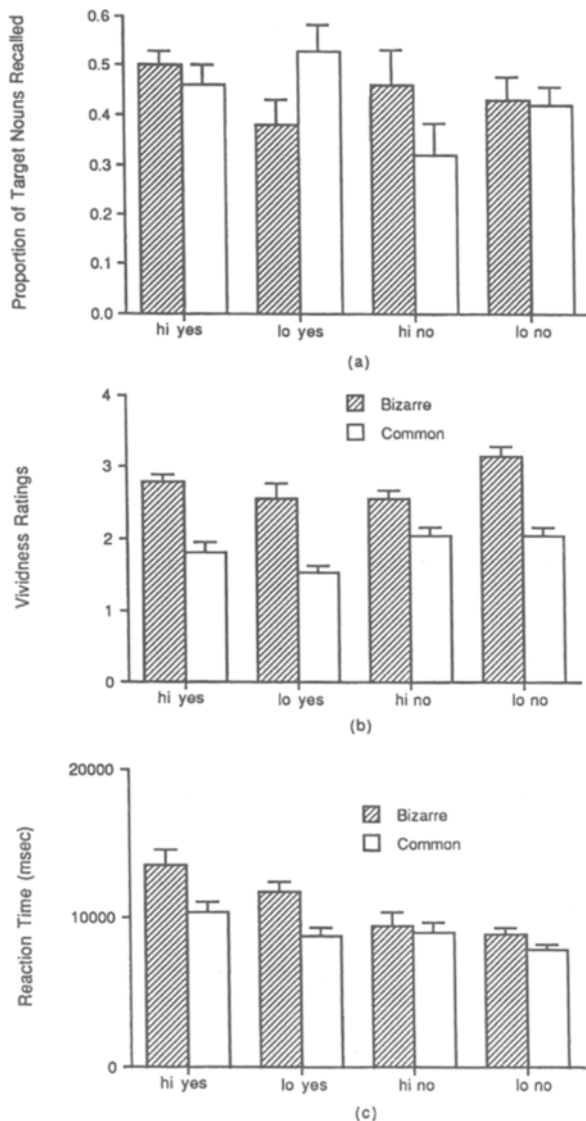


Figure 2. (a) The proportion of target nouns recalled, (b) the vividness ratings, and (c) reaction time in Experiment 2, as a function of sentence type (bizarre or common), guided imagery instructions (yes or no), and imagery ability (hi or lo).

sentences, as these instructions were designed to do. Because it is easier to form images from reading common sentences, common sentences would not necessarily require guidance in imaging (cf. Wollen & Magres, 1987). Nevertheless, even under guided imagery instructions, a bizarreness effect was not found.

Subsidiary analysis. In order to further investigate the possibility that failure to adequately image is responsible for the absence of bizarreness effects obtained with the use of complex sentences, we analyzed recall as a function of imagery ability, assessed through the use of a mental rotation test. Subjects above the median were considered to be "high imagers" and subjects at or below the median were considered "low imagers" (see Figure 2). A $2 \times 2 \times 2$ (instructions \times ability \times bizarreness)

mixed factor ANOVA on the proportion of target nouns recalled revealed no significant main effects or interactions (all F s < 1). These results indicate that the absence of a bizarreness effect was not mediated by individual differences in imagery ability (as defined by performance on a mental rotation test).

Even though subjects were presumably given enough time and guidance for forming an image from reading a complex sentence, a bizarreness effect did not emerge. There is some preliminary evidence in the literature, however, that the allotted processing times were not generous enough. Accordingly, we conducted another experiment in which we required all subjects to spend an extensive amount of time imaging each sentence.

EXPERIMENT 3

The results of the previous two experiments seem to indicate that the failure to find a bizarreness effect is not an outcome of the complexity disrupting the imaginal processing that may confer a mnemonic advantage upon bizarre sentences. However, other researchers (Kline & Groninger, 1991; Richman, personal communication¹) have succeeded in finding the bizarreness effect when the presentation time of complex sentences has been fixed at relatively long intervals (e.g., 30 sec). In Experiment 2, in which subject-paced presentation intervals were used, image formation time never exceeded 15 sec (less than half the time used when a bizarreness effect has been found). Therefore, in Experiment 3, each sentence was presented for 30 sec.

Method

Subjects. Sixteen introductory psychology students at Purdue University participated to fulfill a course requirement.

Design and Materials. Each subject was presented with 12 sentences containing 6 bizarre and 6 common interactions. Each sentence was presented on one page of a booklet containing all 12 sentences. Half the subjects in each condition received the Kroll et al. sentences, and the other half received the revised sentences.

Procedure. The subjects were told that they would be given 30 sec to form an image from reading the sentence presented on each page of the booklet. They were asked to maintain this image for as much of the 30-sec interval as possible. At the end of the 30 sec, they were asked to turn the page and rate the vividness of the image on the same 5-point scale as that used in Experiment 1. They were given 4 sec to rate the sentence. At the end of the 4 sec, they were asked to turn the page and form an image from reading the next sentence.

Results and Discussion

The rejection level was again set at .05. The average vividness ratings and the proportion of target nouns recalled were each subjected to a separate one-way ANOVA (see Figure 3). The analysis of the proportion of target nouns recalled showed no effect of sentence type [$F(1,15) = .47$, $MS_e = .022$]. The analysis of the average vividness ratings showed that common images were rated as significantly less vivid than bizarre images [$F(1,15) = 19.95$, $MS_e = .257$]. This result is directly opposed to the results of previous studies on the bizarre-

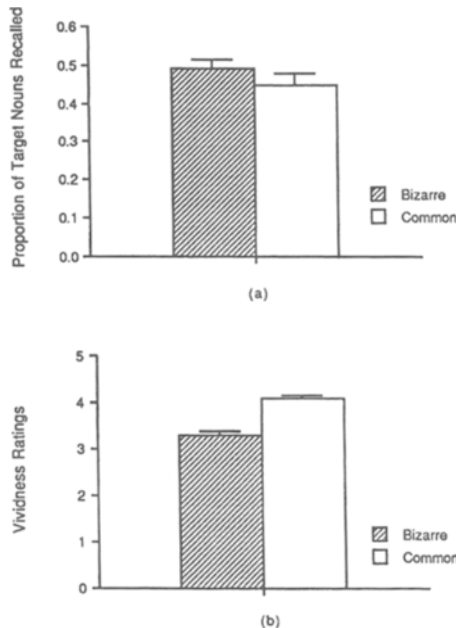


Figure 3. (a) The proportion of target nouns recalled and (b) the vividness ratings in Experiment 3, as a function of sentence type (bizarre or common).

ness effect in which bizarre sentences have been rated as less vivid than common images. This finding might indicate that bizarre sentences were at least as well imaged as common sentences in this experiment; nevertheless, a significant mnemonic advantage for bizarre sentences did not emerge. There was a nominal .04 advantage for bizarre sentences, a magnitude of difference that has been reported as statistically significant in a study utilizing complex sentences (with presentation times of 30 sec or more) but testing many subjects (Richman, personal communication). The important point here is that this slight difference, though perhaps reliable, is far less robust than that found with simple sentences, and the question why remains. The results of Experiments 1, 2, and 3 provided no evidence that manipulations presumed to enhance the imageability of the complex sentences affected the bizarreness pattern; accordingly, we next turned to another line of inquiry.

EXPERIMENT 4

Another possible explanation for the failure of complex sentences to exhibit a bizarreness effect is that the complexity of the sentence provides a mnemonic benefit. The additional adjectives used to describe the nouns in the complex sentences may produce mnemonically beneficial elaboration of the nouns. Craik and Tulving (1975) have shown that words embedded in a complex sentence frame are recalled better than the same words embedded in a simple sentence frame. They suggested that the complex sentence frame served to elaborate the target word and that this elaboration was beneficial for

recall. Similarly, the target nouns for both common and bizarre sentences were well elaborated in our complex sentences, so much so that the common sentences might have had sufficient retrieval cues (or routes) associated with them to be recalled well. To test this possibility, we examined the mnemonic effectiveness of complex sentences in relation to that of simple sentences. Although Craik and Tulving have demonstrated a mnemonic advantage for complex sentences as opposed to simple sentences, it is important to determine whether complexity functions similarly for both bizarre and common sentences in an imagery paradigm.

In addition, Experiment 4 allowed an evaluation of another possible explanation for the failure to find the bizarreness effect with the use of complex sentences. Richman, Dunn, Kahl, Sadler, and Simmons (1990) have suggested that increasing the number of words per sentence (i.e., in the present instance, the additional adjectives) encourages subjects to break the sentence down into phrases. When these phrases are logical (the noun modifiers make sense), they are considered by the subject to be common. Therefore, a bizarreness effect would not be expected because functionally all nouns are presumably encoded within a common context. If bizarre complex sentences are processed in such a way that they become functionally common, then a mixed list of bizarre complex and bizarre simple sentences should produce a recall advantage for bizarre simple sentences. This pattern of results would parallel the typical result found in a mixed-list design of common and bizarre items. Alternatively, if complexity confers a mnemonic benefit upon the sentences, then a recall advantage should be found for complex sentences, regardless of sentence type (bizarre or common).

Method

Subjects. Forty introductory psychology students at Purdue University participated to fulfill a course requirement.

Design and Materials. The design was a two-factor mixed factorial, with sentence type (common vs. bizarre) as the between-subjects factor and sentence complexity (complex vs. simple) as the within-subjects factor. Twenty subjects were randomly assigned to each of the between-subjects conditions.

Two lists consisting of 12 sentences each (6 complex and 6 simple sentences per list) were constructed for each condition. The complex sentences were those utilized in Experiments 1 and 2. The simple sentences were constructed by eliminating the adjectives from these complex sentences. One list contained the common forms of the sentences, and the second list contained the bizarre forms. Two versions of each list were constructed so that each target noun pair would appear once in a complex context and once in a simple context.

Procedure. The subjects were treated identically to those in the no-guidance instructions condition in Experiment 2. The only difference was that they were not required to finish the mental rotation test after completing the free recall task.

Results

The rejection level was again set at .05. Free recall of the target nouns was evaluated by assessing the proportion of target nouns recalled (see Figure 4). These data

were subjected to a 2×2 mixed factor ANOVA. This analysis showed a significant effect of complexity. Target nouns embedded within complex sentences were recalled significantly more often than those embedded within simple sentences [$F(1,38) = 4.12, MS_e = .03$].

In order to determine the locus of this effect, two subsequent measures of recall performance were evaluated—the proportion of sentences accessed (a sentence

was scored as accessed if at least one word from that sentence was recalled), and the number of nouns per sentence recalled. Each of these measures (see Figure 4) was submitted to a separate 2×2 mixed ANOVA. The sentence access measure showed a significant effect, with complex sentences being accessed more often than simple sentences [$F(1,38) = 6.94, MS_e = .04$], whereas the number of nouns recalled per sentence was not sig-

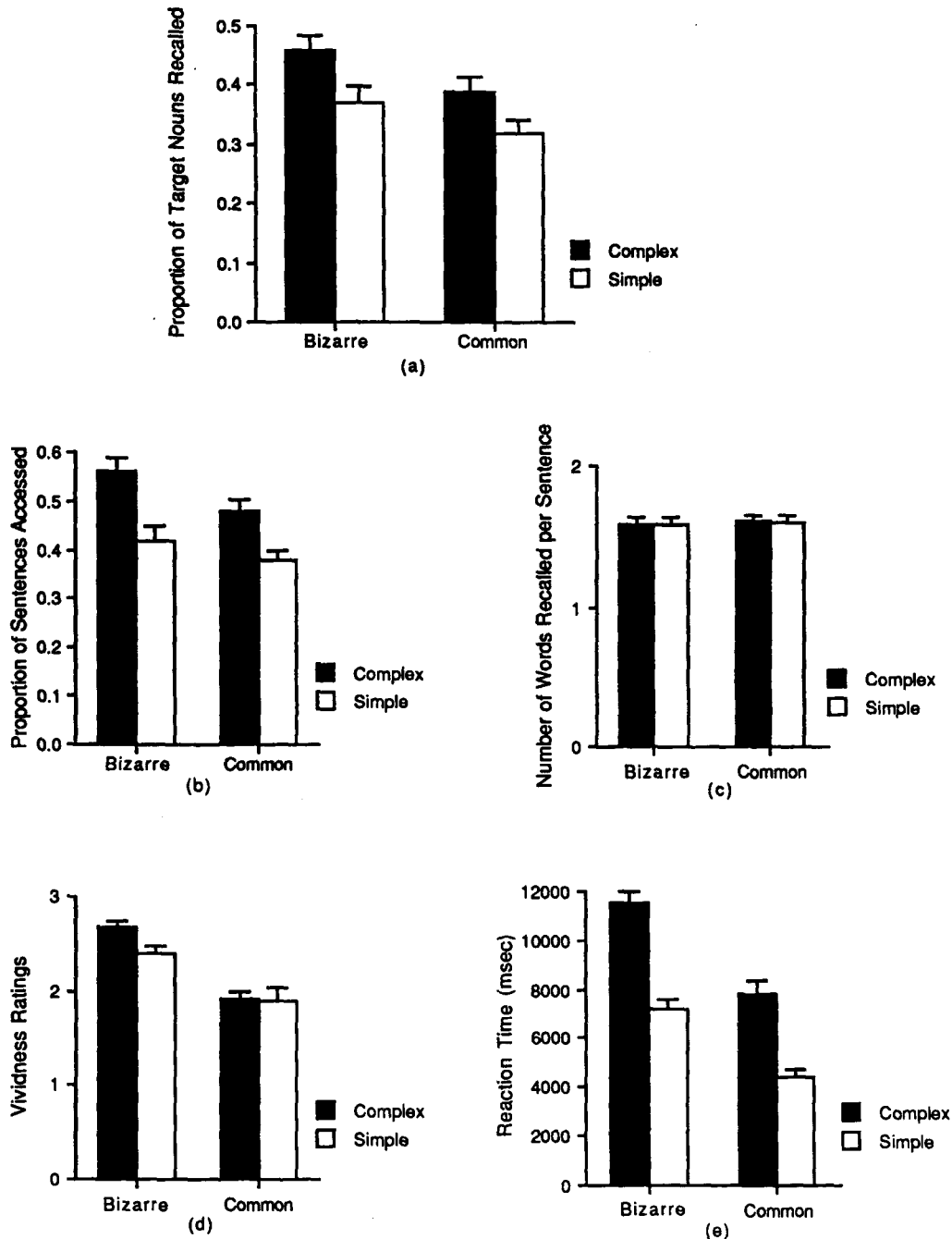


Figure 4. (a) The proportion of target nouns recalled, (b) the proportion of sentences accessed, (c) the number of words recalled per sentence, (d) the vividness ratings, and (e) reaction time in Experiment 4, as a function of complexity (complex or simple) and sentence type (bizarre or common).

nificantly different. The two-way interactions with either of these measures and sentence type were not significant, indicating that the advantage of complexity was the same for both common and bizarre sentences. In order to ensure that this effect was due to sentence access and not a strategy in which subjects recalled nouns from complex sentences first, the first half of each subject's free recall protocol was analyzed to determine the proportion of complex and simple sentences recalled. The proportion of target nouns recalled was .46 and .54 for simple and complex versions, respectively. This difference was not statistically significant [$F(1,38) = 0.92$, $MS_e = .12$]. The two-way interaction with sentence type was not significant [$F(1,38) = 0.92$, $MS_e = .12$].

These results demonstrate that complex images have their advantage in access (sentence access measure) and not in accessibility to the component parts of the image (number of nouns per sentence and the number of target nouns recalled) or due to a particular recall strategy in which nouns from complex sentences are recalled first. In addition, bizarre sentences did not exhibit a memory advantage over common sentences on any measure of free recall. This is consistent with the finding that a bizarreness effect is typically not found when unmixed lists are employed (see Einstein & McDaniel, 1987).

The average vividness ratings as well as the average image formation times (see Figure 4) were each subjected to a separate 2×2 mixed factor ANOVA. The analysis of the vividness ratings showed that complex and simple sentences were not rated as significantly different. The two-way interaction with sentence type was not significant. In addition, bizarre sentences were rated as significantly less vivid than common sentences [$F(1,38) = 8.04$, $MS_e = .98$].

The analysis of the reaction time data showed that complex images took significantly longer to form than the simple images [$F(1,38) = 90.72$, $MS_e = 3,452,152.27$]. The two-way interaction with sentence type was not significant. It is important to keep in mind, however, that the reaction time data are not indicative of pure image formation but also include the time needed to read the sentence. Given that complex sentences were longer in length because of the additional adjectives, these sentences would require a longer time to read than the simple sentences.

Discussion

The results of this experiment demonstrate that complex sentences exhibit a recall advantage over simple sentences, and this benefit is manifested through the measure of sentence access. Furthermore, the benefit for complex sentences was not significantly different in the bizarre and common conditions, indicating that embellishing a sentence with adjectives increases the memorability for bizarre and common sentences equally.

In addition, the results of this experiment appear to counter the idea, suggested by Richman et al. (1990), that the failure to find a bizarreness effect with the use of complex sentences is due to the bizarre sentences'

being functionally common. If bizarre complex sentences were functionally common, it would seem that the bizarre simple sentences should have been recalled better than the bizarre complex sentences. This pattern would have paralleled those found when bizarre and common sentences were encoded in a mixed-list design. Therefore, it seems reasonable to conclude that bizarre complex sentences are indeed considered bizarre and not common.

GENERAL DISCUSSION

"People—including mnemonists, experimental psychologists, and experimental subjects—believe that bizarre imagery helps their memories, but . . . it does not." Kroll et al. (1986, p. 50) based this statement on the failure to find a bizarreness effect for their complex sentences. Recent failures to consistently find the effect with the use of complex sentences (Experiments 1, 2, and 3 of the study; McDaniel & Einstein, 1989; Richman et al., 1990) strengthens this contention. Robust bizarreness effects have been found for simple sentences (Hirshman et al., 1989; McDaniel & Einstein, 1986), however, and no current theoretical account would expect the effect to be different for complex sentences. In the present investigation, we have examined this apparent disparity.

The results of the first three experiments demonstrated that subjects did not utilize the bizarreness of the complex sentences to assist them during retrieval. In light of the results of Experiment 4, showing a mnemonic benefit of complexity, we propose that the cues associated with the complex sentences are more salient or precise (Stein & Bransford, 1979) than whatever retrieval cues are associated with the bizarre sentences (cf. Hirshman et al., 1989; McDaniel & Einstein, 1986; Wollen & Magres, 1987). Therefore, the cues associated with complexity may dominate during retrieval. The idea here is similar to the "outshining" explanation of why environmental reinstatement effects are not always observed (cf. Smith, 1988). According to this explanation, other, stronger, cues present in the to-be-remembered material preempt the use of environmental context cues. Perhaps in much the same way, subjects in this study may have utilized the cues associated with complexity during retrieval, thereby "outshining" the cues tied to sentence bizarreness.

This view would suggest that if the mnemonic advantage of complexity could be diminished, a bizarreness effect should be obtained. McDaniel and Einstein (1991) reported a preliminary experiment (conducted by B.R.-R.) attempting to do just this. We attempted to reduce the presumed mnemonic benefit of the adjectives in the complex sentences by producing cue overload (see Watkins & Watkins, 1975). Each noun in the complex sentence was described by adjectives chosen from a small set, resulting in each adjective's being used five or six times across the sentence set. The initial experiment obtained a significant bizarreness effect, but subsequent

experiments (unpublished) showed the effect not to be reliable. It is possible, though, that if a more effective technique were found to neutralize the retrieval advantage associated with sentence complexity, a reliable bizarreness effect would emerge.

In sum, it appears that the mnemonic effectiveness of bizarre imagery is further constrained to an impoverished encoding context (i.e., simple, unelaborated sentences). When the encoding context is relatively impoverished (e.g., simple sentences), subjects appear to utilize the bizarreness of the sentences as a cue for sentence retrieval (as is evidenced by the fact that the locus of the bizarreness effect is in sentence access and not in the number of words recalled per sentence), and a bizarreness effect is observed. Conversely, when the encoding context is relatively rich (e.g., complex sentences), more specific retrieval cues may be available to aid in recovering the target nouns at test. Hence, a bizarreness effect is not found. Whether the cue associated with bizarreness is the surprise response made in reaction to the bizarre sentences (Hirshman et al., 1989), the distinctiveness of the bizarre sentences (McDaniel & Einstein, 1986), or cues associated with differential processing in the formation of bizarre versus common images (Wollen & Magres, 1987), all theories must account for the fact that the effect appears to be limited by the availability of other more effective cues. If such cues are available, subjects will apparently utilize them, thereby diminishing the mnemonic advantage of bizarreness.

REFERENCES

- BEECH, J. R., & ALLPORT, D. A. (1978). Visualization of compound scenes. *Perception*, *7*, 129-138.
- BOWER, G. H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning & Verbal Behavior*, *9*, 529-533.
- 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.
- COX, S. D., & WOLLEN, K. A. (1981). Bizarreness and recall. *Bulletin of the Psychonomic Society*, *18*, 244-245.
- CRAIK, F. I. M., & TULVING, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *55*, 467-517.
- 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., & LACKEY, S. (1989). Bizarre imagery, interference, and distinctiveness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 137-146.
- HIRSHMAN, E., WHELLEY, M. M., & PALIJ, M. (1989). An investigation of paradoxical memory effects. *Journal of Memory & Language*, *28*, 594-609.
- HUBBARD, T. L., & STOECKIG, K. (1988). Musical imagery: Generation of tones and chords. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 656-667.
- KLINE, S., & GRONINGER, L. D. (1991). The imagery bizarreness effect as a function of sentence complexity and presentation time. *Bulletin of the Psychonomic Society*, *29*, 25-27.
- KOSSLYN, S. M., CAVE, C. B., PROVOST, D. A., & VON GIERKE, S. M. (1988). Sequential processes in image generation. *Cognitive Psychology*, *20*, 319-343.
- KOSSLYN, S. M., REISER, B. J., FARAH, M. A., & FLIEGEL, S. L. (1983). Generating visual images: Units and relations. *Journal of Experimental Psychology: General*, *112*, 278-303.
- KROLL, N., SCHEPELER, E. M., & ANGIN, K. T. (1986). Bizarre imagery: The misremembered mnemonic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *12*, 42-54.
- KROLL, N., & TU, S. F. (1988). The bizarre mnemonic. *Psychological Research*, *50*, 28-37.
- LORAYNE, H., & LUCAS, J. (1974). *The memory book*. New York: Stein & Day.
- 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., ANDERSON, D. C., EINSTEIN, G. O., & O'HALLORAN, C. M. (1989). Modulation of environmental reinstatement effects through encoding strategies. *American Journal of Psychology*, *102*, 523-548.
- MCDANIEL, M. A., & EINSTEIN, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *12*, 54-65.
- MCDANIEL, M. A., & EINSTEIN, G. O. (1989). Sentence complexity eliminates the mnemonic advantage of bizarre imagery. *Bulletin of the Psychonomic Society*, *27*, 117-120.
- MCDANIEL, M. A., & EINSTEIN, G. O. (1991). Bizarre imagery: Mnemonic benefits and theoretical implications. In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 183-192). Amsterdam: Elsevier.
- O'BRIEN, E. J., & WOLFORD, C. R. (1982). Effect of delay in testing on retention of plausible versus bizarre mental images. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*, 148-152.
- PAIVIO, A., & HARSHMAN, R. (1983). Factor analysis of a questionnaire on imagery and verbal habits and skills. *Canadian Journal of Psychology*, *37*, 461-483.
- 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.
- RICHMAN, C. L., DUNN, J., KAHL, G., SADLER, L., & SIMMONS, K. (1990). The bizarre sentence effect as a function of list length and complexity. *Bulletin of the Psychonomic Society*, *28*, 185-187.
- SMITH, S. (1988). Environmental context-dependent memory. In G. M. Davies & D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 13-34). Wiley.
- STEIN, B. S., & BRANSFORD, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning & Verbal Behavior*, *18*, 769-777.
- WATKINS, O. C., & WATKINS, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning & Memory*, *104*, 442-452.
- WOLLEN, K. A., & COX, S. D. (1981a). The bizarreness effect in a multi-trial intentional learning task. *Bulletin of the Psychonomic Society*, *18*, 296-298.
- WOLLEN, K. A., & COX, S. D. (1981b). Sentence cuing and the effectiveness of bizarre imagery. *Journal of Experimental Psychology: Human Learning & Memory*, *104*, 442-452.
- WOLLEN, K. A., & MAGRES, 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.
- YATES, F. A. (1966). *The art of memory*. London: Routledge & Kegan Paul.

NOTE

1. This study was cited in a previous review of the present article.

APPENDIX

Kroll et al.'s (1986) Sentences

The new sleek TRAIN is derailed by the fresh juicy STRAWBERRIES
 The new sleek TRAIN passes a field of fresh juicy STRAWBERRIES

The black leather COAT crawls from under the king-size BED
 The black leather COAT is on the king-size BED

The large cracked VASE sings loudly as it plays a dusty PIANO
 The large cracked VASE is on top of a dusty PIANO

The slender rookie POLICEWOMAN hungrily eats the tall stone GATE
 The slender rookie POLICEWOMAN quickly climbs the tall stone GATE

The large reddish COCKROACH, waving its feelers, carries off the dirty STOVE
 The large reddish COCKROACH, waving its feelers, climbs up the dirty STOVE

The old bearded doctor operates urgently on the pale green TELEPHONE
 The old bearded doctor speaks urgently into the pale green TELEPHONE

A shaggy brown DOG balances on top of the tossed FRISBEE
 A shaggy brown DOG jumps to catch the tossed FRISBEE

A white REFRIGERATOR dances merrily with a happy pine TABLE
 A white REFRIGERATOR is in the corner by a thick pine TABLE

The new MAGAZINE is read by a brown dried-out APPLE
 The new MAGAZINE partially covers a brown dried-out APPLE

The brown TROUT paddles his own birch-bark CANOE
 The brown TROUT is dropped into the birch-bark CANOE

A sturdy wooden CRATE relaxes in an old uncomfortable CHAIR
 A sturdy wooden CRATE is next to an old uncomfortable CHAIR

The lazy arrogant WAITER serves a plate containing a narrow white NECKTIE
 The lazy arrogant WAITER ties a narrow white NECKTIE

Revised Sentences

The smoking freight TRAIN punches a field of jumbo red STRAWBERRIES
 The smoking freight TRAIN passes a field of jumbo red STRAWBERRIES

The long fur COAT skates on the canopy BED
 The long fur COAT is on the canopy BED

A cracked crystal VASE does cartwheels across a black player PIANO
 A cracked crystal VASE is on top of a black player PIANO

The slender blonde POLICEWOMAN eats the arched wrought-iron GATE
 The slender blonde POLICEWOMAN climbs the arched wrought-iron GATE

The six-legged COCKROACH lassos the grimy pot-belly STOVE
 The six-legged COCKROACH crawls across the grimy pot-belly STOVE

The bearded DOCTOR, wearing glasses, surfs on the cordless TELEPHONE
 The bearded DOCTOR, wearing glasses, speaks into the cordless TELEPHONE

A short-haired spotted DOG cooks the teeth-marked FRISBEE
 A short-haired spotted DOG jumps to catch the teeth-marked FRISBEE

A side-by-side-REFRIGERATOR-freezer dances with a round glass TABLE
A side-by-side-REFRIGERATOR-freezer is by a round glass TABLE

The opened tattered MAGAZINE bowls with a half-eaten APPLE
The opened tattered MAGAZINE is under a half-eaten APPLE

The squirming scaly TROUT kisses a muddy aluminum CANOE
The squirming scaly TROUT is dropped into a muddy aluminum CANOE

An empty wooden CRATE tickles a wicker rocking CHAIR
An empty wooden CRATE is next to a wicker rocking CHAIR

The short bald WAITER drives a wide paisley NECKTIE
The short bald WAITER ties a wide paisley NECKTIE

(Manuscript received June 1, 1993;
revision accepted for publication March 4, 1994.)