*REVISTA GALEGO-PORTUGUESA DE PSICOLOXÍA E EDUCACIÓN*

*Vol. 17, (1,2), Ano 13º-2009 ISSN: 1138-1663*

# *LIMITATIONS OF THE EFFICACY OF BIZARRE IMAGERY IN MIXED LISTS*

## *LIMITACIONES DE LA EFICACIA DE LAS IMÁGENES RARAS EN LISTAS MIXTAS*

***Rocío GÓMEZ-JUNCAL***
***María José PÉREZ-FABELLO***
***Alfredo CAMPOS****
*University of Vigo*
** University of Santiago de Compostela*

*Data de recepción: 16/07/2008*
*Data de aceptación: 05/03/2009*

***RESUMEN***

Desde hace siglos se utilizan las imágenes raras como ayuda a la memoria, sin embargo, se sigue estudiando en qué condiciones es efectiva. Se ha demostrado que es eficaz cuando se cumplen varias condiciones: recuerdo libre inmediato, listas mixtas, y aprendizaje incidental. En este trabajo deseábamos investigar la eficacia de las imágenes raras en comparación con las imágenes normales, en listas mixtas, con oraciones simples (Experimento1 y 2) y complejas (Experimento 3 y 4), con aprendizaje incidental (Experimto 1 y 3) y aprendizaje intencional (Experimento 2 y 4). Se midió el efecto de lo raro inmediatamente después del aprendizaje, al cabo de un día y de una semana. Como variables dependientes hemos utilizado el recuerdo, el acceso a la oración, el número de ítems recordados por oración, el número de oraciones totalmente recordadas, y el reconocimiento. No se ha encontrado un efecto claro de lo raro con ninguna de las variables estudiadas, ni independientes, ni dependientes. Los peores resultados se han conseguido con el aprendizaje intencional (Experimento 2 y 4). Se discuten los resultados y se proponen nuevas líneas de investigación.

***PALABRAS CLAVE*:** Imagen mixta, imagen mental, imagen rara, aprendizaje incidental.

Address correspondence to:
Alfredo Campos, University of Santiago de Compostela, Department of Basic Psychology, 15782 Santiago de Compostela (Spain), e-mail: pscampos@usc.es

*ABSTRACT*

For centuries bizarre imagery has been an aid to memory yet the precise conditions for its optimum effectiveness remain unknown. Bizarre imagery has been reported to be effective under certain conditions: free immediate recall, mixed lists, and incidental learning. The aim of this study was to assess the efficacy of bizarre imagery in comparison to normal imagery, in mixed lists with simple sentences (Experiment 1 and 2) and complex sentences (Experiment 3 and 4), with incidental learning (Experiment 1 and 3) and intentional learning (Experiment 2 and 4). In all the experimental conditions, bizarre imagery was assessed immediately after learning, at a 1-day, and at a 1-week interval. Dependent variables were as follows: recall, sentence access, number of items recalled per sentence, number of sentences fully recalled, and recognition. No clear effect was found between bizarre and any of the independent or dependent variables under study. The lowest efficacy was observed with intentional learning (Experiment 2 and 4). The results are discussed in the light of further lines of investigation.

***KEY WORDS*:** Mixed imagery, mental imagery, bizarre imagery, incidental learning.

## *LIMITATIONS OF THE EFFICACY OF BIZARRE IMAGERY IN MIXED LISTS*

Mental imagery plays a key role in several cognitive tasks such as memorizing, reasoning, problem-solving, etc. (Campos, González, & Amor, 2004a; Campos, Pérez-Fabello, & Calado, 2003; Denis, 1979; Higbee, 1993; Richardson, 1994). Mental imagery as a mnemotecnic technique (for further reference see Yates, 1966) has been employed for thousands of years, going back from ancient times to the present day (see Mercer, 1996; Worthen, 2006; Yates, 1966, for a full review). Bizarre imagery has been recommended instead of normal imagery though experimental studies have reported no clear differences in performance between normal imagery and bizarre imagery (Kroll & Tu, 1988; McDaniel & Einstein, 1986, 1989, 1991; Mercer, 1996).

Bizarre imagery can be either atypical (rarely occurring) or illogical (never occurring) (Mercer, 1996). Imai and Richman (1991) found that the bizarreness effect occurred only with atypical sentences when they were presented for a 7-second period; hence the recommendation for the use of atypical bizarre imagery over illogical bizarre imagery. Robinson-Riegler and McDaniel (1994) Experiment 1, found recall was greater with atypical bizarre imagery than with normal imagery though no differences were observed when compared to illogical bizarre imagery. A review of the literature on imagery concluded that bizarreness had an effect under the following conditions: a) when free recall is used as opposed to recognition, b) with immediate recall (5 minutes or less) in contrast to long-term recall, and c) and when mixed lists are used instead of pure lists (Mercer, 1996). Burns (1996) has added a fourth condition, d) with incidental versus non-intentional learning.

With mixed lists, simple sentences, and incidental learning, McDaniel and Einstein (1986, 1989) observed higher recall of bizarre image items than common images items. McDaniel and Einstein (1991) found that recall was greater with bizarre imagery than normal imagery when mixed lists and incidental learning was used regardless of the number (6 or 12) of sentences. Similar findings were reported by Worthen and Loveland (2000-2001) and Worthen, García-Rivas, Green and Vidos (2000). In contrast, Kroll and Tu (1988), Experiment 1, found no difference in recall between normal and bizarre imagery when mixed lists were used.

From the 80´s onwards memory measurements besides recall (recognition is not normally used with mixed lists for reasons that

will be explained in the methods section), such as sentences access (a sentence is scored as correct if at least one word per sentence is correctly recalled), and the number of items recalled per sentence. The efficacy of normal imagery versus bizarre imagery in terms of the number of full sentences recalled is not usually assessed using mixed lists).

With mixed lists, incidental learning and simple sentences, McDaniel and Einstein (1986) observed that more bizarre sentences were accessed than common sentences in the within-list design. Similar findings were reported by Worthen and Loveland (2000-2001), and Worthen, García-Rivas, Green and Vidos (2000). McDaniel and Einstein (1991) found greater sentence access with bizarre imagery than with normal imagery regardless of the list size (6 or 12 sentences).

The precise efficacy of bizarre and normal imagery on the number of items recalled per sentence remains unclear. Whereas McDaniel and Einstein (1986), and Worthen and Loveland (2000-2001) found no differences between normal and bizarre imagery in terms of the number of items recalled per sentence, Worthen et al. (2000) found the highest number of items recalled per sentence using normal imagery as compared to bizarre imagery. Moreover, McDaniel and Einstein (1991) found no difference between normal and bizarre imagery in the number of items per sentence recalled, regardless of whether the list had 6 or 12 sentences.

The results for complex sentences contrast with those observed with are very different of simple sentences. McDaniel and Einstein (1989) found bizarreness to be more effective when simple sentence frames were used whereas recall was greater for normal imagery when complex sentence frames were used. According to these authors, sentence complexity eliminates the mnemonic advantage of bizarre imagery when bizarreness is manipulated within-subjects. In contrast, McDaniel, Einstein, DeLosh, May, and Brady (1995), and Robinson-Riegler, and McDaniel (1994) found no differences in recall between normal and bizarre imagery in complex lists.

Mental imagery is characterised by the rapid loss of image memory recall. Campos, Amor, and González (2004a), in line with the work of Ashcraft (1998), interpret the quick memory loss of imagery strategies in terms of Tulving´s (1972, 1989, 1993) theory of episodic memory, which is characterized by good immediate recall but rapid decline over time. Visually coded materials are particularly susceptible to interference through time thus they are quickly forgotten. In order to avoid interference and mental imagery memory loss, Campos, Gómez-Juncal, and Pérez-Fabello (2007, in press-a, b, d) have attempted to consolidate recall to delay memory loss by determining the efficacy of recall at different time intervals. The authors observed the effects of bizarre on: imagery recall, recognition, sentence access, and the number of full sentences recalled after a one-week interval but no effect was found in the number of items recalled per sentence.

In this study the efficacy of bizarre imagery versus normal imagery on memory was assessed using: mixed lists with simple sentences (Experiment 1 and 2) and complex sentences (Experiment 3 and 4), with incidental learning (Experiment 1 and Experiment 3), and intentional learning (Experiment 2 and Experiment 4). The scores for recall, sentence access, number of items recalled per sentence, number of fully recalled sentences, and the recognition task were used as measurements of memory. The aim was to determine the efficacy of bizarre imagery on memory immediately following learning, and at one-day and one-week intervals.

## *EXPERIMENT 1*

Experiment 1 was designed to assess the difference in recall between normal imagery

versus bizarre imagery using mixed lists (intralist), with the incidental learning of simple sentences.

*METHOD*

*Participants*. The sample consisted of 41 Spanish compulsory secondary education students (21 boys and 20 girls with a mean age of 14.5 years (*SD* = 1.13), and an age range of 12 to 17 years.

*Materials*. A list of 96 words (all nouns) taken from the official Spanish dictionary (Diccionario de la Lengua Española) was used. Fifty percent of a list of 48 words were selected to make a list of 16 simple sentences describing normal situations with three words per sentence. The same words were used in the same sentences to describe the list of 16 atypical bizarre imagery simple sentences. In order to ensure that both lists were equal, the same verbs, defining and non-defining articles, number (singular or plural), gender etc., were used in both lists. The mean words per sentence for the normal and atypical bizarre simple sentences were 7.56 and 7.56 respectively. Thus, no significant differences were observed between the normal and the atypical bizarre imagery lists in terms of the number of words per sentence ($t = .001$, $p > .05$). The correlation between number of words per sentence in the normal and atypical bizarre imagery lists was .77, $p < .001$.

*Procedure*. Two lists, one with 16 normal sentences and the other 16 atypical bizarre imagery sentences, were randomly mixed before being equally divided to form two mixed lists containing 8 normal sentences and 8 standard bizarre imagery sentences. Subjects were given one of the two lists thus word difficulty was controlled as the same words were used in both the normal a bizarre sentences. The three words in each sentence, taken from the original lists, were written in upper case.

As the experiment involved incidental learning, the task required subjects to score imagery vividness for each sentence using a five-point scale where 1 indicated "no image at all", and 5 "image as clear and vivid as normal vision" (McDaniel et al., 1995; Campos et al., in press-a, b; Campos, Gómez-Juncal, & Pérez-Fabello, in press-c). The instructions were written on one sheet of paper and the list of sentences on another. Subjects were given a 5-sentence training session prior to being instructed they had 4 minutes to complete the task i.e., 15 seconds per sentence. The vividness scores were not used for later analysis, and were only used as an incidental learning strategy.

Alter learning the sentences, subjects were given a 2-minute distracting task consisting of spotting the differences between two similar pictures. Thereafter, subjects were given a blank sheet of paper and asked to write for a 5-minute period as many sentences as they could recall or as much of each sentence they could recall. Following the recall task, subjects were administered a recognition task consisting of identifying as many of the 48 words (initially written in upper case in the sentences) from the original list of 96 words.

Following a one-day interval, and without the subjects' prior knowledge, they were asked to repeat the same recall and recognition task under the same experimental conditions. Thereafter subject were administered the same 4-minute imagery questionnaire as the day before as part of incidental relearning. 1 week later, without the subjects' prior knowledge, they underwent the same 5-minute recall and recognition tasks.

The data obtained was assessed in terms of 5 measurements of memory: a) Recall. Number of correctly recalled words. Plural/singular, and masculine/feminine variants for each word were accepted as good but synonyms were rejected. b)

Sentence access. A point was awarded for each correctly recalled sentence if the subject recalled at least one of the three words written in upper case in each sentence. c) Items per sentence recalled. Scoring entailed dividing the total of number words recalled between the total number of sentences recalled if at least one of the three words written in upper case was recalled (i.e., sentence access). d) Number of correct full sentences recalled. One point was awarded for each sentence if the three words written in upper case were correctly recalled. e) Recognition. Number of recognized words. This procedure was employed as no other method was possible (see for example, Einstein, McDaniel, & Lackey, 1989; Zechmeister & Nyberg, 1982) i.e., as the method employed intralists, false recognitions cannot be discriminated.

*RESULTS*

In order to assess the differences between normal and bizarre imagery word recall during the three time intervals under study (i.e., immediate recall, recall following a 1-day and 1-week interval), a repeated measurements MANOVA was performed. The mean and standard deviations for each group are shown in Table 1. The Roy's largest root indicated significant differences in recall according to the type of imagery used, Roy's largest root = .23, $F(3, 38) = 2.948$, $p <. 05$, power = .65. No significant differences were observed between normal and bizarre imagery after immediate recall, $F(1, 40) = 2.712$, $p >. 05$, power = .36, nor after a 1-day lapse, $F(1, 40) = 3.867$, $p >. 05$, power = .48. In contrast, significant differences were found between normal and bizarre imagery recall after a 1-week interval, $F(1, 40) = 8.542$, $p < . 01$, power = .81. Bizarre imagery accounted for the highest recall levels as compared to normal imagery.

**TABLE 1.** Mean and Standard Deviations (in Brackets) of the Different Measurements of Memory i. e., Immediately after learning, at 1-Day, and at the 1-Week Interval According to the Learning Strategy (Simple Sentences and Incidental Learning)

| | Memory Measures | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Recall | | Sentence Access | | Items per Sentence | | Full Sentences | | Recognition | |
| | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre |
| Immediate | 9.32 | 10.56 | 3.59 | 3.83 | 2.58 | 2.70 | 2.44 | 2.93 | 18.68 | 19.32 |
| | (3.65) | (3.62) | (1.30) | (1.30) | (.44) | (.48) | (1.32) | (1.23) | (3.02) | (3.54) |
| At 1-Day | 12.27 | 13.85 | 4.73 | 5.10 | 2.60 | 2.73 | 3.07 | 3.76 | 19.63 | 20.59 |
| | (4.48) | (4.86) | (1.66) | (1.79) | (.32) | (.24) | (1.46) | (1.48) | (3.23) | (2.78) |
| At 1-Week | 14.88 | 16.88 | 5.71 | 5.90 | 2.59 | 2.84 | 3.80 | 5.10 | 21.49 | 21.90 |
| | (5.52) | (4.73) | (1.97) | (1.59) | (.36) | (.17) | (1.95) | (1.72) | (2.49) | (2.80) |

To analyze difference between normal and bizarre imagery word recall during the three time intervals, a repeated measurements MANOVA was undertaken. The mean and standard deviations for each group are shown in Table 1. Roy's largest root revealed no significant differences between sentence access and image type, Roy's largest root = .05, $F(3, 38) = .573$, $p >. 01$, power = .16.

To examine the difference between normal and bizarre imagery in relations to the number of items recalled per sentence immediately, at 1-day, and at a 1-week interval, a repeated measurements MANOVA was carried out. The mean and standard deviations for each group are shown in Table 1. The Roy's largest root revealed significant differences in the number of items recalled per sentence according to the image type used, Roy's largest root = .69, $F(3, 38) = 8.747$, $p <. 01$, power = .99. No significant differences were found between normal and bizarre imagery and the number of items recalled per sen-

tence immediately after learning, $F(1, 40) = 1.555$, $p >. 05$, power = .23, or at the 1-day interval, $F(1, 40) = 3.964$, $p >. 05$, power = .49; however, significant differences were observed between normal and bizarre imagery at one-week, $F(1, 40) = 27.447$, $p < . 01$, power = 1. Sentence access was greater for bizarre imagery than for normal imagery.

To determine the differences between normal and bizarre imagery on the number of fully recalled sentences immediately after learning, at 1-day, and at a 1-week time interval, a repeated measurements MANOVA was carried out. The mean and standard deviation for each group is shown in Table 1. Roy's largest root showed significant differences in the number of fully recalled sentences in relation to the imagery used, Roy's largest root = .69, $F(3, 38) = 7.543$, $p <. 01$, power = .98. Moreover, univariate analysis found no significant difference between normal and bizarre imagery in the number of sentences fully recalled immediately, $F(1, 40) = 3.300$, $p >. 05$, power = .43. However, significant differences were observed between normal and bizarre imagery in the number of sentences fully recalled at the 1-day interval, $F(1, 40) = 7.157$, $p <. 05$, power = .72. Bizarre imagery had greater impact on sentence access at 1-day than normal imagery. Whilst at 1- week, significant differences were found between normal and bizarre imagery, $F(1, 40) = 23.526$, $p < . 01$, power = 1. Greater sentence access was observed for bizarre imagery than with normal imagery.

To analyse the difference between normal and bizarre imagery and word recognition immediately after learning, at 1- day, and at 1 week, a repeated measurements MANOVA was performed. The mean and standard deviations for each group are shown in Table 1. Roy's largest root found no significant differences in terms of recognition and the image type used, Roy's largest root = .10, $F(3, 35) = 1.293$, $p > .05$, power = .32.

## *EXPERIMENT 2*

Experiment 2 was designed to examine the differences among the different measurements of recall and normal or bizarre imagery using intralist, simple sentences, and intentional learning.

### *METHOD*

*Participants*. A total of 42 Spanish compulsory secondary education students (20 boys and 22 girls), with a mean age of 14.5 years ($SD$ = 1.18) and age range of 12 to 17 years were included for Experiment 2.

*Materials and Procedure*. The materials and procedure were the same as those in Experiment 1, with the only exception being that Experiment 1 involved incidental learning whereas Experiment 2 assessed intentional learning.

### *RESULTS*

To determine the difference between normal and bizarre imagery word recall immediately after learning, at 1-day, and at 1-week, a repeated measurements MANOVA was performed. The mean and standard deviations for each group are shown in Table 2. Roy's largest root showed no significant differences in recall in terms of the image type used, Roy's largest root = .02, $F(3, 35) = .226$, $p > . 05$, power = .09. A MANOVA was also undertaken to examine the differences between normal and bizarre imagery and sentence access after immediate learning, at 1-day, and 1-week. The mean and standard deviations for each group are shown in Table 2. Roy's largest root revealed no significant differences in sentence access in relation to the image type used, Roy's largest root = .06, $F(3, 35) = .721$, $p >. 05$, power = .19.

The MANOVA to examine difference between normal and bizarre imagery and the number of items recalled per sentence, imme-

diately, at 1-day or at the 1-week interval, (see mean in Table 2), indicated no significant differences in the number of items per sentence recalled according to the image type used, Roy's largest root = .15, $F(3, 35) = 1.788$, $p >$ . 05, power = .43.

**TABLE 2.** Mean and Standard Deviations (in Brackets) of the Different Measurements of Memory i. e., Immediately after learning, at 1-Day, and at the 1-Week Interval According to the Learning Strategy (Simple Sentences and Intentional Learning)

| | Memory Measures | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Recall | | Sentence Access | | Items per Sentence | | Full Sentences | | Recognition | |
| | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre |
| Immediate | 10.95 | 10.47 | 4.32 | 3.87 | 2.55 | 2.67 | 2.71 | 3.00 | 19.58 | 18.53 |
| | (4.21) | (4.58) | (1.49) | (1.47) | (.41) | (.40) | (1.66) | (1.72) | (3.35) | (4.17) |
| At 1-Day | 13.45 | 13.08 | 5.18 | 4.82 | 2.59 | 2.70 | 3.47 | 3.66 | 19.84 | 19.58 |
| | (4.56) | (5.67) | (1.63) | (1.89) | (.38) | (.42) | (1.87) | (2.03) | (3.56) | (3.47) |
| At 1-Week | 14.13 | 14.16 | 5.34 | 5.08 | 2.62 | 2.69 | 3.68 | 4.16 | 21.11 | 20.53 |
| | (4.91) | (6.43) | (1.65) | (2.16) | (.33) | (.55) | (1.95) | (2.22) | (3.14) | (2.93) |

Differences between normal and bizarre imagery and the number of sentences fully recalled were also analysed immediately after learning, at1-day, and at 1-week interval. The mean and standard deviations for each group are shown in Table 2. A repeated measurements MANOVA found a Roy's largest root = .11, $F(3, 35) = 1.259$, $p >$ . 05, power = .31, indicating no significant differences in the number of sentences fully recalled in relation to the image type used.

To assess the differences between normal and bizarre imagery in relation to word recognition immediately after learning, at 1-day or at the 1-week interval, a repeated measurements MANOVA was carried out. The mean and standard deviations for each group are shown in Table 2. Roy's largest root showed significant differences in recognition according to the type of imagery, Roy's largest root = .25, $F(3, 35) = 2.926$, $p < .05$, power = .65. The univariate analysis showed that the image strategy (normal or bizarre in mixed groups) influenced immediate recognition, $F(1, 37) = 5.688$, $p < .05$, power = .64. Immediate recognition was greater for normal imagery than for bizarre imagery. In contrast, no significant differences were observed between either strategy type and recognition at 1-day, $F(1, 37) = .390$, $p > .05$, power = .09, or at the 1-week interval, $F(1, 37) = 3.951$, $p > .05$, power = .49.

## *EXPERIMENT 3*

Experiment 3 aimed to assess the differences among the different measurements of recall in relation to normal and bizarre imagery using mixed lists (intralist), complex sentences, and incidental learning.

### *METHOD*

*Participants*. A total 59 Spanish compulsory secondary education students (40 boys and 19 girls with a mean age of 14.29 years ($SD = 1.12$) and an age range of 12 to 17 years participated in this experiment.

*Material and Procedure*. The materials and procedure were the same as those used in Experiment 1, with the only exception being that in this experiment complex sentences were used with a mean of 13 words for each normal and bizarre sentence. Significant difference between simple normal sentences (Experiment 1), and normal complex sentences (Experiment 3) were observed in terms of the number of words per list ($t = 23.38$, $p <$ .001). Significant differences were also found

between bizarre imagery with simple sentences and bizarre imagery with complex sentences in terms of the number of words per list ($t$ = 22.56, $p < .001$).

The procedures were the same as those used in Experiment 1, with the exception that in Experiment 3, subjects were allowed 30 seconds to generate their normal or bizarre imagery and score their imagery vividness whereas in Experiment 1 subjects were allowed 15 seconds per sentence. The difference in time period was due to the longer time period required for image formation using complex sentences as opposed to simple sentences.

*Results.* To analyze the difference between normal and bizarre imagery in relation to word recall, immediately after learning, at 1-day, and at 1-week, a repeated measurements MANOVA was undertaken. The mean and standard deviations for each group are shown in Table 3. Roy's largest root showed significant differences in recall according to the type of image used, Roy's largest root = .29, $F(3, 52) = 5.035$, $p <. 01$, power = .90. Significant difference were observed between normal and bizarre imagery and immediate recall, $F(1, 54) = 7.738$, $p < . 01$, power = .78. Immediate recall was greater for normal imagery than for bizarre imagery. Significant difference were also found between normal and bizarre imagery at 1-day, $F(1, 54) = 7.692$, $p < . 01$, power = .78. Recall was greater with normal imagery than with bizarre imagery at 1-day. However, no differences between image type and recall were observed at 1-week, $F(1, 54) = 2.345$, $p > . 05$, power = .33.

**TABLE 3.** Mean and Standard Deviations (in Brackets) of the Different Measurements of Memory i. e., Immediately after learning, at 1-Day, and at the 1-Week Interval According to the Learning Strategy (Complex Sentences and Incidental Learning)

| | Memory Measures | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Recall | | Sentence Access | | Items per Sentence | | Full Sentences | | Recognition | |
| | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre |
| Immediate | 8.53 | 6.60 | 3.49 | 2.45 | 2.51 | 2.39 | 1.78 | 1.65 | 18.67 | 16.98 |
| | (3.37) | (4.36) | (1.53) | (1.55) | (.36) | (.89) | (.99) | (1.28) | (3.23) | (4.25) |
| At 1-Day | 10.00 | 8.13 | 4.00 | 2.98 | 2.46 | 2.56 | 2.00 | 2.05 | 18.71 | 17.44 |
| | (3.51) | (4.69) | (1.36) | (1.67) | (.50) | (.72) | (1.22) | (1.46) | (2.97) | (3.35) |
| At 1-Week | 10.24 | 11.13 | 4.16 | 4.16 | 2.37 | 2.59 | 2.04 | 2.73 | 19.98 | 19.38 |
| | (4.67) | (4.87) | (1.90) | (1.66) | (.54) | (.50) | (1.49) | (1.51) | (3.29) | (3.37) |

To evaluate the difference between normal and bizarre imagery in terms of sentence access immediately after learning, at 1-day, and at a 1-week interval, a repeated measurements MANOVA was performed. The mean and standard deviations for each group are shown in Table 3. Roy's largest root revealed significant differences in sentence access according to the type of image used, Roy's largest root = .37, $F(3, 52) = 6.483$, $p <. 001$, power = .96. Significant differences were found between normal and bizarre imagery in sentence access immediately after learning $F(1, 54) = 13.636$, $p < . 001$, power = .95. Sentence access was greater with normal imagery than with bizarre imagery after immediate learning. Significant difference between normal and bizarre imagery were also observed at 1-day, $F(1, 54) = 15.320$, $p < . 001$, power = .97. Sentence access was greater with normal imagery than with bizarre imagery at 1-day. However, 1-week interval no significant difference were observed between both image types in sentence access at 1-week, $F(1, 54) = .001$, $p > . 05$, power = .05.

A repeated measurements MANOVA was performed to examine the differences between normal and bizarre imagery according to the number of items recalled per sentence immediately after learning, at 1-day, and at 1-week. The mean and standard deviations for each

group are shown in Table 3. Roy's largest root showed significant differences in the number of items recalled per sentence according to the image type used, Roy's largest root = .19, $F(3, 52) = 3.270$, $p <$. 05, power = .72. No significant difference between normal and bizarre imagery in relation to the number of items recalled per sentence were observed immediately after learning, $F(1, 54) = .913$, $p >$. 05, power = .16, or at 1-day, $F(1, 54) = 1.063$, $p >$. 05, power = .17. In contrast, significant difference between normal and bizarre imagery were found in terms of the number of items recalled per sentence at 1-week, $F(1, 54) = 8.811$, $p <$ . 01, power = 83. The greatest number of items recalled per sentence was observed with bizarre imagery than with normal imagery.

To examine the differences between normal and bizarre imagery in relation to the number of sentences fully recalled immediately, at 1-day and at the 1-week interval, a repeated measurements MANOVA was carried out. The mean and standard deviations for each group are shown in Table 3. Roy's largest root showed significant differences in the number of sentences fully recalled according to the type of image used, Roy's largest root = .27, $F(3, 52) = 4.714$, $p <$. 01, power = .87. Later univariate analysis revealed no significant difference between normal and bizarre imagery in terms of the number of sentences fully recalled immediately, $F(1, 54) = .429$, $p >$. 05, power = .10. No significant difference were found between normal and bizarre imagery in relation to the number of sentences fully recalled at 1-day, $F(1, 54) = .07$, $p >$ . 05, power = .06. However, significant difference were observed between normal and bizarre imagery at 1-week, $F(1, 54) = 13.934$, $p <$ . 001, power = 96. The greatest number of fully recalled sentences was observed with bizarre imagery than with normal imagery.

To evaluate the difference between normal and bizarre imagery word recognition using intralists immediately after learning, at 1-day, and at 1-week, a repeated measurements MANOVA was performed. The mean and standard deviations are shown in Table 3. Roy's largest root indicated significant differences in word recognition in accordance with image type used, Roy's largest root = .26, $F(3, 52) = 4.480$, $p < .01$, power = .86. Univariate analysis showed that the type of image strategy (normal or bizarre in mixed groups) influenced immediate recognition, $F(1, 54) = 12.794$, $p < .001$, power = .94. Immediate word recognition was greater with normal imagery than with bizarre imagery. Significant differences were also observed between the type of image strategy and word recognition at 1-day, $F(1, 54) = 9.307$, $p < .01$, power = .85. Word recognition was greater with normal than with bizarre imagery at 1-day. No significant difference were observed between image strategy type and word recognition at 1-week, $F(1, 54) = 2.912$, $p > .05$, power = .39.

## *EXPERIMENT 4*

Experiment 4 was designed to examine the differences among the different measurements of recall and normal imagery and bizarre imagery using intralist, complex sentences, and intentional learning.

### *METHOD*

*Participants*. A total 55 Spanish compulsory secondary education students (25 boys and 28 girls with a mean age of 14.46 years ($SD$ = .96) and an age range of 13 to 17 years participated in this experiment.

*Material and Procedure*. The materials and procedure were the same as those in Experiment 3, with the exception that in this experiment intentional learning was used whereas Experiment 3 involved incidental learning.

### *RESULTS*

To determine the differences between normal and bizarre imagery in relation to word

recall immediately after learning, at 1-day, and at 1-week, a repeated measurements MANOVA was performed. The mean and standard deviations for each group are shown in Table 4. Roy's largest root revealed no significant differences in word recall between the image type used, Roy's largest root = .08, $F(3, 44) = 1.227$, $p > .05$, power = .31.

**TABLE 4.** Mean and Standard Deviations (in Brackets) of the Different Measurements of Memory i. e., Immediately after learning, at 1-Day, and at the 1-Week Interval According to the Learning Strategy (Complex Sentences and Intentional Learning)

| | Memory Measures | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Recall | | Sentence Access | | Items per Sentence | | Full Sentences | | Recognition | |
| | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre | Normal | Bizarre |
| Immediate | 8.47 (3.86) | 7.49 (5.09) | 3.40 (1.50) | 2.81 (1.84) | 2.45 (.51) | 2.36 (.86) | 1.81 (1.19) | 1.94 (1.52) | 19.96 (3.16) | 18.89 (3.87) |
| At 1-Day | 11.34 (5.08) | 10.96 (5.66) | 4.60 (1.18) | 4.04 (2.01) | 2.44 (.42) | 2.57 (.63) | 2.45 (1.63) | 2.83 (1.74) | 20.22 (3.02) | 18.72 (4.22) |
| At 1-Week | 13.02 (5.58) | 13.79 (5.43) | 5.13 (2.03) | 5.06 (1.77) | 2.50 (.41) | 2.57 (.53) | 2.96 (1.88) | 3.72 (1.86) | 20.87 (2.88) | 20.96 (2.84) |

To examine the difference between normal and bizarre imagery in sentence access immediately after learning, at 1-day, and at 1-week, a repeated measurements MANOVA was undertaken. The mean and standard deviations for each group are shown in Table 4. Roy's largest root showed significant differences between image type in sentence access, Roy's largest root = .24, $F(3, 44) = 3.451$, $p < .05$, power = .74. Univariate analysis revealed significant difference between normal and bizarre imagery in relation to sentence access immediately after learning, $F(1, 46) = 4.377$, $p < .05$, power = .54. Sentence access was greater with normal than with bizarre imagery. Significant differences were also found between normal and bizarre imagery at 1-day, $F(1, 46) = 6.511$, $p < .05$, power = .71. Sentence access was greater with normal than with bizarre imagery at 1-day. However, no significant differences were observed between the mean sentence access of both image types, $F(1, 46) = .063$, $p > .05$, power = .06.

To examine the difference between normal and bizarre imagery in relation to the number of items recalled per sentence immediately after learning, at 1-day, and at 1-week, a repeated measurements MANOVA was performed. The mean and standard deviations for each group are shown in Table 4. Roy's largest root found no significant differences in the number of items recalled per sentence in relation to the image type used, Roy's largest root = .11, $F(3, 44) = 1.624$, $p > .05$, power = .40.

To determine the differences between normal and bizarre imagery in relation to the number of sentences fully recalled immediately, at 1-day, and at the 1-week interval, a repeated measurements MANOVA was undertaken. The mean and standard deviations for each group are shown in Table 4. Roy's largest root revealed significant differences in the number of sentences fully recalled in accordance to the type of image used, Roy's largest root = .23, $F(3, 44) = 3.367$, $p < .05$, power = .72. Later univariate analysis found no significant difference between normal and bizarre imagery in the number of sentences fully recalled immediately after learning, $F(1, 46) = .341$, $p > .05$, power = .09 or at 1-day, $F(1, 46) = 4.009$, $p < .05$, power = .50. In contrast, significant differences were observed between normal and bizarre imagery, $F(1, 46) = 7.356$, $p < .01$, power = 76 at 1-week. The greatest number of full sentences recalled was obtained with bizarre imagery than with normal imagery.

To evaluate the differences in word recognition with normal or bizarre imagery using a mixed lists immediately after learning, at 1-day and at 1-week, a repeated measurements MANOVA was carried out. The mean and standard deviations are shown in Table 4. Roy's largest root showed significant differences in word recognition in relation to the type of image used (normal or bizarre), Roy's largest root = .36, $F(3, 43) = 5.150$, $p < .01$, power = .90. The univariate analysis found that image strategy (normal or bizarre in mixed groups) influenced immediate word recognition, $F(1, 45) = 6.510$, $p < .05$, power = .70. Immediate recognition was greater with normal imagery than with bizarre imagery. Significant difference between normal and bizarre image strategies were observed in terms of recognition at 1-day, $F(1, 45) = 11.429$, $p < .01$, power = .91. Word recognition was greatest with normal than with bizarre imagery at 1-day. No significant differences in word recognition were observed between the image strategies at 1-week, $F(1, 45) = .97$, $p > .05$, power = .06.

## *DISCUSSION*

With mixed lists, simple sentences and incidental recall (Experiment 1), the bizarre effect on recall was observed at 1-week, which is in agreement with the findings of Campos et al. (in press-a, b), who using pure lists and intermediate relearning report the efficacy of bizarre imagery at 1-week. In contrast, bizarre imagery was not effective immediately after learning as reported by McDaniel and Einstein (1986, 1989, 1991), Worthen and Loveland (2000-2001), and Worthen et al. (2000). The discrepancies with these authors on immediate recall may arise from the different ways in which the sentences were presented i.e., sentences were presented individually, normally on computer screens during a fixed time period whereas in the present study sentences were presented sentences on a sheet of paper and had to distribute the time period as they wished. The aim was to ensure the learning procedure resembled real-life learning as closely as possible. Several studies (see Campos, Amor, & González, 2004a,b; Campos, González, & Amor, 2004b; Campos, González et al., 2004a; King-Sears, Mercer, & Sindelar, 1992;Thomas & Wang, 1996, for a review) have suggested that the procedure for presenting learning material influences learning through the keyword-generation method in keyword mnemonics.

As for sentence access and recognition no difference between normal and bizarre imagery were found. To our knowledge, no previous studies have been undertaken on recognition using mixed lists. The bizarre effect on the number of fully recalled sentences was observed at the 1-day and at 1-week interval. Likewise, no previous studies have been carried out on the measurement of fully recalled sentences. The bizarre effect on the number of items recalled per sentence was only observed at the1-week interval. Previous studies (McDaniel & Einstein, 1986, 1991; Worthen & Loveland, 2000-2001) have not found differences between normal and bizarre imagery in terms of the number of items recalled per sentence immediately after learning.

With lists mixed lists, simple sentences and intentional recall (Experiment 2) no significant differences were found between normal and bizarre imagery in recall, sentence access, number of items recalled per sentence, and the number of fully recalled sentences. This is in agreement with previous studies (see Burns, 1996 for a review), that have reported that bizarre imagery is not effective with intentional learning. A significant relationship between normal and with bizarre imagery was only found in terms of recognition. Recognition was greater with normal imagery than bizarre imagery. To date, data are not available to draw a comparison with our findings.

With mixed lists, complex sentences incidental recall (Experiment 3) normal imagery

was more effective than bizarre imagery in recall (immediate at 1-day), sentence access (immediately after learning and at 1-day), and recognition (immediately alter learning and at 1-day). Previous studies have reported no bizarre effect using complex sentences (McDaniel and Einstein, 1989; McDaniel et al.,1995; Robinson-Riegler, & McDaniel, 1994). McDaniel and Einstein (1989) found recall was greater with normal imagery when complex sentence frames were used which coincides with our results. The bizarre effect was only observed in the number of items recalled per sentence at 1-week, and the number of fully recalled sentences. Similar bizarre effects were observed at 1-week using pure lists by Campos et al. (in press-a, b).

With mixed lists, complex sentences and intentional learning (Experiment 4) no significant differences between normal and bizarre imagery were observed in recall and in the number of items recalled per sentence. This learning strategy had the inconvenience of using complex sentences, which is not effective as was the case with bizarre imagery and complex sentences (McDaniel and Einstein, 1989; McDaniel et al.,1995; Robinson-Riegler, & McDaniel, 1994). As stated in the introduction, a further inconvenience of using intentional learning (for a review see Burns, 1996), is that bizarre imagery is not effective with intentional learning. The bizarre effect was observed on the number of fully recalled sentences at 1-week, similar results were obtained in Experiment 1 and 3. Sentence access at 1-day and immediate recognition at 1- day were greater with normal imagery than with bizarre imagery. Unfortunately, no previous studies with these variables are available to compare our results.

In conclusion our findings reveal that under our experimental conditions, on the whole learning was more effective using normal imagery than bizarre imagery. Bizarre imagery was only more effective with simple lists and incidental learning. Normal imagery was more effective with complex lists particularly in terms of recognition. Nevertheless, most of the analysis undertaken in this study show no significant differences between normal and bizarre imagery. Further studies are required to ascertain the efficacy of different learning material presentation techniques on memory.

## *REFERENCES*


Ashcraft, M. H. (1998). *Fundamentals of cognition*. New York: Longman.

Burns, D. J. (1996). The bizarre imagery effect and intention to learn. *Psychonomic Bulletin & Review, 3*, 254-257.

Campos, A., Amor, A., & González, M. A. (2004a). The importance of the keywor-generation method in keyword mnemonics. *Experimental Psychology, 51*, 125-131.

Campos, A., Amor, A., & González, M. A. (2004b). Drawing-assited strategies in keyword mnemonics. *Studia Psychologica, 46*, 211-218.

Campos, A., González, M. A., y Amor, A. (2004a). Different strategies for keyword generation. *Journal of Mental Imagery, 28*, 51-58.

Campos, A., González, M. A., y Amor, A. (2004b). Limitations of the mnemonic-keyword method. *Journal of General Psychology, 130,* 399-413.

Campos, A., Gómez-Juncal, R., & Pérez-Fabello, M. J. (2007). Imágenes mentales, recuerdo y reconocimiento en sujetos altos en viveza de imagen. *Revista Galego-Portuguesa de Psicoloxía e Educación, 14*, 121-130.

Campos, A., Gómez-Juncal, R., & Pérez-Fabello, M. J. (in press-a). Experiência na mnemotécnica e aprendizagem incidental

com imagens normais e rara. *Estudos de Psicologia*.

Campos, A., Gómez-Juncal, R., & Pérez-Fabello, M. J. (in press-b). Mnemotecnia mediante imágenes y aprendizaje incidental de sujetos altos en viveza de imagen. *Adaxe*.

Campos, A., Gómez-Juncal, R., & Pérez-Fabello, M. J. (in press-c). Experience in imagery and imagery vividness. *Imagination, Cognition and Personality*.

Campos, A., Gómez-Juncal, R., & Pérez-Fabello, M. J. (in press-d). Efecto de las imágenes raras y del reaprendizaje sobre la memoria. *Revista Galego-Portuguesa de Psicoloxía e Educación.*

Campos, A., Pérez-Fabello, M. J., & Calado, M. (2003). Imagen normal-rara y su efecto en el recuerdo. *Revista Galego-Portuguesa de Psicoloxía e Educación, 9,* 159-165.

Denis, M. (1979). *Les images mentales*. Paris: PUF.

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

Higbee, K. L. (1993). *Your memory*. New York: Paragon House.

Imai, S., & Richman, C. L. (1991). Is the bizarreness effect a special case of sentence reorganization? *Bulletin of the Psychonomic Society, 29*, 429-432.

King-Sears, M. E., Mercer, C. D., & Sindelar, P. T. (1992). Toward independence with keyword mnemonics: A strategy for science vocabulary instruction. *Remedial and Special Education, 13*, 22-33.

Kroll, N. E. A., & 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 and 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). North-Holland: Elsevier Science Publishers.

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

Mercer, C. (1996). The bizarre imagery effect on memory. *Journal of Mental Imagery*, *20*, 141-152.

Richardson, A. (1994). *Individual differences in imaging*. Amityville, NY: Baywood.

Robinson-Riegler, R., & McDaniel, M. A. (1994). Further constraints on the bizarreness effect: Elaboration at encoding. *Memory & Cognition, 22*, 702-712.

Thomas, M. H., & Wang, A. Y.(1996). Learning by the keyword mnemonic: Looking for long-term benefits. *Journal of Experimental Psychology: Applied, 2*, 330-343.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donalson (Eds.). *Organization of memory* (pp. 381-403). New York: Academic Press.

Tulving, E. (1989). Remembering and knowing the past. *American Scientist, 77*, 361-367.

Tulving, E. (1993). What is episodic memory? *Current Direction in Psychological Science, 2*, 67-70.

Worthen, J. B. (2006). Resolution of discrepant memory strengths: An explanation of the effects of bizarreness on memory. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp.133-156). New York: Oxford Press.

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

Worthen, J. B., & Loveland, J. M. (2000-2001). Imagery nonvividness and the mnemonic advantage of bizarreness. *Imagination, Cognition and Personality, 20*, 373-381.

Yates, F. A. (1966). *The art of memory*. London: Routledge & Kegan.

Zechmeister, E. B., & Nyberg, S. E. (1982). *Human memory: An introduction to research and theory.* Monterey, CA: Brooks/Col