Psychology
2011. Vol.2, No.3, 216-219
Copyright © 2011 SciRes.

Scientific Research
DOI:10.4236/psych.2011.23033

Memory Strength and Criterion Shift in the False Memory Paradigm: A Learning Case

Shahid Naved, Ameer Haider Ali, Khubaib Ahmed Qureshi
Hamdard University, Karachi, Pakistan.
Email: khubaib_ahmed@yahoo.com
Received February 28th, 2011; revised April 15th, 2011; accepted May 18th, 2011.

The attempt has been made to investigate the criterion shift hypothesis once again by re-evaluating the confidence measurement, which will possibly clarify the role that criterion shifts play in the false memory phenomenon (recollection of an event, or the details of an event, that did not occur). Literature review shows that this hypothesis still needs research upon the same topic. The study was experimental in which students of Hamdard University were selected as subjects—40 students from BBA and MBA programs. Both male/female and left/right handed subjects participated. All the subjects were not native English speakers. The experiment was conducted using a computer program to collect the data. The experiment had two parts, firstly a study/recall phase and secondly a test/recognition phase. The scale we introduced to allow participants to assess their own certainty about the classification of recognition items is more detailed than that used in the Roediger and McDermott study. Our hypothesis was that a shift in decision criterion would become evident by means of a lower certainty measure for lure words as compared to target words from the lists. This difference was found in our data. The mean certainty measure we found for the critical lures is significantly lower than the mean certainty for the targets.

Keywords: Memory, Recognition, Lures, Associate Memory

Introduction

Roediger and McDermott (1995), by applying methods first introduced by Deese (1959), have investigated the false memory phenomenon. In their study, subjects learned lists of words strongly related to one non-presented word. In recognition and recall tests, the non-presented strong association was recognized and recalled at a rate comparable to actually presented words. Specifically, in their first experiment, the non-presented associates were recalled 40% of the time. The authors have concluded that this procedure reliably creates false memories. They put a heavy emphasis on the illusionary character of the phenomenon, comparable to perceptual illusions. Additionally, the effect could also be observed in cases where the subjects were specifically informed that one would try to induce false memories in them. Roediger and McDermott provided some speculative explanatory approaches, one being that the false memory experience is the result of a reconstructive process that, in the case of auditory encoding, makes it possible for the subject to generate a representation of how the word would have sounded if presented in the speaker's voice. This clear representation would make it plausible that the subjects wrongly claim to remember the word's presentation.

Miller and Wolford (1999) have reliably reproduced the findings of Roediger and McDermott. In their opinion, however, the high rates of false recall and recognition are due to a change in the decision criterion. When asked to rate a presented item as new/old, the strength of that particular item is generally believed to be made up of two factors. Firstly, the item gains evidence of having been seen before through presentation. Additionally, however, the item also gains strength due to associative activation, meaning that an item semantically related to the list items will be considered more likely to be old than unrelated items. If the strength of activation for a particular item exceeds a certain threshold, it is considered to be old; otherwise it will be classified as new. Yet, according to Miller and Wolford (1999), the false memory phenomenon thus observed does not really represent false memory. In their opinion, a change in decision criteria, thus in the threshold for old/new classification, is responsible for the phenomenon. More specifically, for the critical lures Miller and Wolford claim that the decision criterion is lower than for unrelated lures, because subjects judge the critical lures to be more likely to be old due to the fact that they are strongly semantically related to the rest of the list.

The thesis of Miller and Wolford has been criticized, for instance by Wixted and Stretch (2000). Even though the claim that the false memory phenomenon is only due to criterion shift is quite controversial, even Wixted and Stretch admit that it is uncertain if and to what extent criterion shifts contribute to false memory. Further research is needed to answer this open question.

The objective of this experiment is to test the criterion shift hypothesis once again. We think that the introduction of a confidence measurement could possibly clarify the role that criterion shifts play in the false memory phenomenon. Thus, when asked to judge an item, subjects will not only be asked to classify that item as being either new or old, but they will also be asked to rate their own confidence about their decision on a scale. If there is a change in the decision criterion during recognition of items and lures from a given list, then the rate of false alarm for critical lures should be the same as in the Roediger/McDermott experiment. However, the confidence ratings for those items should generally be lower than the con-

fidence/memory strength ratings for presented items. This is, after all, what is meant by a change in decision criterion: items are judged to be old even though they have a lower activation because the threshold for a positive decision is lower. Similar measurements have been taken in the experiments by Roediger/McDermott and Miller/Wolford, but with a scaling that is too rough to properly examine the distribution of memory strength values.

Thus this experimental hypothesis is that criterion shifts should become evident in our setting by means of lower memory strength/confidence measures for critical lures.

Methodology

Subjects

Due to the nature of the experiment, many more students wanted to participate, however 40 subjects were randomly selected from BBA and MBA programs. The ages ranged between 18 - 25 years, included both male/female and left/right handed subjects participated. All subjects were not native English speakers.

Materials

Six lists of items by consulting the data from "Norms of word association" by Russell (1954) were created. Every list contained the 12 associates with the highest frequency from one of Russell's lists. We created seven recognition items per list, one of which was the main associate of the original list. Among these recognition items there were 2 unrelated items. These were picked arbitrarily from the unused lists of Russell (1954).

Also, two weak associates were part of the recognition items for each particular list. Those weak associates were taken from the positions 13 and lower of the association norms. Finally, two of the recognition items were studied words from the list. One of these was the first item in the association norms list, the other occurred somewhere in the first six positions. The recognition items were presented in blocks of seven items, each block corresponding to one studied list. One block always started with a studied word and ended with the critical lure. The order of the blocks corresponded to the order in which the lists had been studied.

Procedure

The experiment was conducted using a computer program to collect the data. Generally, the experiment had two parts, firstly a study/recall phase and secondly a test/recognition phase. During study phase, the lists were presented to the Subjects on a computer screen one word at a time, each word presentation lasting about 2 seconds. After the presentation of each list there was a recall test, in which the Subjects were instructed to write down as many words as they could remember from the previous presentation. This process was repeated six times, once per list. After the completion of the study/recall phase, there was a short break, after which the recognition test took place. The instructions for this recognition test were displayed on each Subject's computer screen. Then the recognition items were presented one at a time in the order described in the materials section. For each item, Subjects first had to decide whether it was a studied item or an unstudied one. In both cases, they afterwards had to rate their confidence on a scale of 1 (absolutely sure) ... 7 (not sure at all).

At the end of the experiment, Subjects were given handouts explaining the purpose of the experiment in which they had just participated.

Results

We collected data for the old/new classification and for the certainty measure. As expected, a high probability for "old" classification was found for both lures and targets.

The certainty and old/new data was then compiled onto one scale by multiplying the certainty value by –1 for words classified as "new" and by 1 for words classified as "old". Thus, the resulting scale ranged from –7 (absolutely sure the item is new) to +7 (absolutely sure the item is old).

From Table 2 it can be seen that a difference in the mean certainty for lure words and for targets was found.

The frequency distributions of the certainty values were also computed, separately for each recognition item type:

Subsequently the standard error for both the mean certainty values and the mean probability for an "old" response were computed (Figure 5, Figure 6). Here the standard error of the mean is indicated by the vertical lines through the data points. A significant difference in mean certainty and mean probability of old classification was found between targets and lure words.

Discussion

First of all, we reproduced the findings of Roediger/

Table 1.
Mean probability of "old" classification.

Item Type	Mean P (Old)
Lures	0.7
Target	0.88
Weakly related	0.09
Unrelated	0.04

Table 2.
Mean certainty measure after rescaling (Scale from −7 to 7).

Item Type	Mean Certainty
Lures	2.88
Target	5.05
Weakly related	−4.34
Unrelated	−5.23

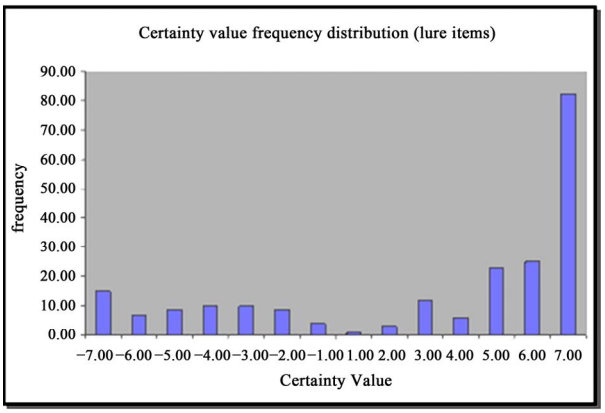

Figure 1.
Histogram lure items certainty.

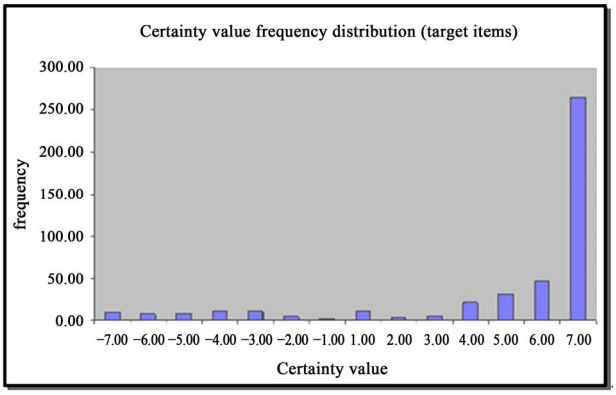

Figure 2.
Histogram target item certainty.

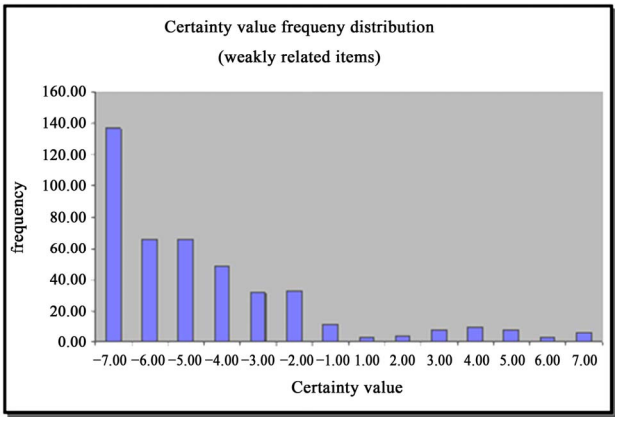

Figure 3.
Histogram weekly related items certainty.

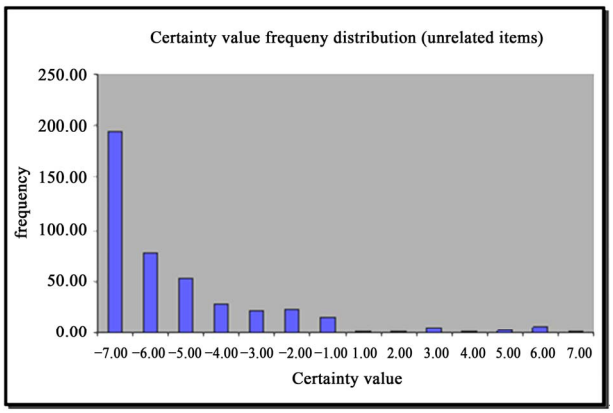

Figure 4.
Histogram unrelated item certainty text with figure.

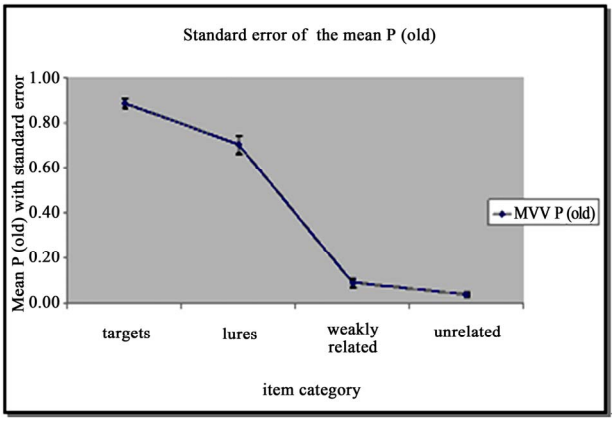

Figure 5.
Standard error of mean P (*old*) *per item type*.

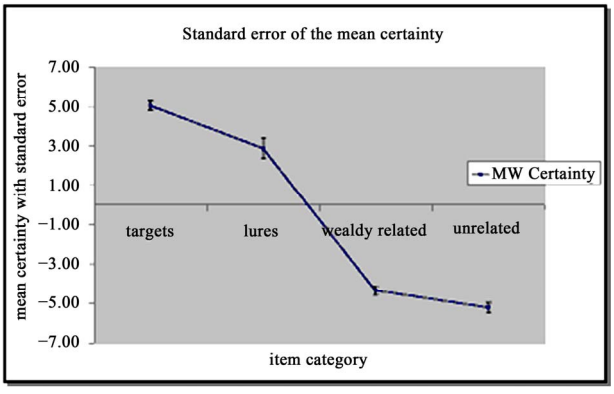

Figure 6.
Standard error of mean certainty per item type.

McDermott (1995). The false alarm rates found for critical non-studied lure words were almost as high as in the Roediger/McDermott study. The scale we introduced to allow participants to assess their own certainty about the classification of recognition items is more detailed than that used in the Roediger/McDermott study (1995).

Our hypothesis was that a shift in decision criterion would become evident by means of a lower certainty measure for lure words as compared to target words from the lists. This difference was found in our data. The mean certainty measure we found for the critical lures is significantly lower than the mean certainty for the targets. The reason for this is most probably

the range of our certainty scale, since similar measurements with a rougher scaling were taken for instance by Roediger/ McDermott (1995) and no such effect was found.

The question remains what this means for the ongoing discussion about the role that criterion shift plays in the false memory phenomenon.

Conclusion

The fact that lower certainty measures could actually be collected for critical lures indicates that criterion shift cannot be ignored when discussing the false memory phenomenon. This implies that it is very unlikely that associative models alone can account for the false memory phenomenon.

At the same time, our results provide no further information concerning the role of associative processes in the phenomenon. Further research examining the exact mechanisms at work here is needed to clarify the role that these processes play in the phenomenon. Thus, it can by no means be ruled out that in fact an interaction of associative mechanisms and mechanisms of shifts in decision criterion result in the false memory effect.

Possibly the effect we found is weaker than it actually could have been. This is because a number of participants reported that they found the scaling range from 1 (not sure) to 7 (absolutely sure) confusing. Thus it can be assumed that a number of them made minor rating mistakes while assessing their certainty.

References

Deese, J. (1995). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports, 5,* 305-312. doi:10.2466/PR0.5.3.305-312

Miller, M. B., & Wolford, G. L. (1999). Theoretical commentary: The role of criterion shift in false memory. *Psychological Review, 106,* 398-405. doi:10.1037/0033-295X.106.2.398

McDermott, K. B., & Roediger H. L. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language, 39,* 508-520. doi:10.1006/jmla.1998.2582

Roediger, H. L., & McDermott K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 803-814. doi:10.1037/0278-7393.21.4.803

Roediger, H. L., & McDermott K. B. (1999). False alarms and false memories. *Psychological Review, 106,* 406-410. doi:10.1037/0033-295X.106.2.406

Russell (1954). *Norms of word association.* L. Postman (Ed.), 1970, New York: Academic Press.

Smith, R. E., & Hunt R. R. (1999). Presentation modality affects false memory. *Psychonomic Bulletin and Review, 5,* 710-715. doi:10.3758/BF03208850

Rhodes, M. G., & Anastasi J. S. (2000). The effects of levels of processing manipulation on false recall. *Psychonomic Bulletin and Review, 7,* 158-162. doi:10.3758/BF03210735

Wixted, J. T., & Stretch V. (2000). The case against a criterion shift account of false memory. *Psychological Review, 107,* 368-376. doi:10.1037/0033-295X.107.2.368