A Study on the Implantation of False Memories

Sumeet Agarwal
Junior Undergraduate
Department of Computer Science and Engineering
Indian Institute of Technology
Kanpur, India - 208016

sagarwal@cse.iitk.ac.in

November 22, 2004

Abstract

In this paper, I present the results of an experiment done to measure the degree of implantation of 'false memories' in people, under varying conditions. Specifically, I wanted to test if 'clustering' of associated cues plays an important role in the formation of such strong, but inaccurate, memories. I also look at some of the existing literature in this area, and try to see what impact my results have on existing models and theories of human memory.

1 Introduction

It is a well known fact in psychology that people often remember things differently from the way they actually happened. In fact, memory recall is intrinsically a reconstructive process, involving various kinds of assumptions, inferences and mental biases. The phenomenon of formation of false memories, both under laboratory conditions (Roediger and McDermott [3]), as well as in real-life situations (Loftus [1]) has been well documented in the literature. For example, Loftus and Coan (cited in [1]) found that they were able to implant a real-life false memory in a 14-year old boy. They had the boy's parents and brother tell him a story about how he'd got lost in a mall at the age of 5. At first the boy denied remembering any such incident, but on hearing the story multiple times, he began to accept it as true, and sometime later, when asked about the event, he narrated it quite confidently, even adding on some details of his own which had not been part of the story told to him. What's more, the boy had been told this story along with several other actual incidents that had occurred in his childhood. When he was later informed that one of the stories had been false, and asked to identify it, he picked a real incident. This goes to show how confident we can be in our memories even when they are inaccurate.

In this study, I attempted to find out whether the formation of these so-called false memories is significantly influenced by whether the relevant cues are presented clustered together or in a distributed fashion. Roediger and McDermott [3] conducted an experiment where they presented subjects with lists of 15 words, all the words in a list being related to a single critical word, which was not presented. When the subjects were given words for recognition, it was found that they were almost as confident of the critical word being on the list as they were for the actual list words. I have altered this experiment to test whether similar results will be seen even when all the words presented are not related to a single critical word. For example, some of the lists I used had two critical words, while one had random unrelated distractors. My hypothesis is that formation of false memories will not be as strong under the altered conditions, since the mental connections that get activated would not be reinforced as much as they are when a long list of related cues is presented without any interruptions. The details of the experiment are given in the next section.

2 Experiment

The experiment consisted of two parts. One was an auditory test involving a recognition task, the other a visual test involving a recall task. For both tests the subjects were presented 5 lists of 12 words each. The words were spoken/shown at 2 second intervals. In the first experiment, after each list, the subjects were given another 12 words and asked to rate them based on whether they had been on the list or not. The rating was on a scale of 1-4, with 1 denoting definitely not present and 4 denoting definitely present. In the second experiment, the subjects just had to recall as many of the list words as they could, and rate their confidence in each word recalled on a scale of 1-4, 1 denoting high uncertainty and 4 denoting high certainty. The lists used for both the cases were the same, and were largely based on the lists used by Roediger and McDermott [3]. However, the 5 lists were all different. They were of the following kinds:

1. All the words were related to a single critical word.
2. There were two critical words, with words on the list being alternately related to one or the other.
3. There were two critical words, with the first three list words being related to one, the next three to the other and so on.
4. There was only one critical word, but 5 of the 12 list words were totally unrelated to it, and nowhere were more than 2 consecutive related words given.
5. Two critical words, with the first six words on the list being related to the first one, and the next six to the second.

In no case were any of the critical words presented on the list. In the recognition task, four kinds of words were given: words on the list, the critical word(s), words weakly related to the list words, and totally unrelated words.

3 Methodology

The recognition experiment was done on 17 subjects. The presentation modality was auditory. The subjects were given response sheets to write down their ratings. The list words were presented at two second intervals. Right after each list, the 12 words for recognition were given to the subjects, and they wrote down their rating (1-4) for each one. Then the next list would begin. All 5 lists were presented in this way.

List number	Cluster sizes	Critical words	Distractor words
1.	12	1	No
2.	1	2	No
3.	3	2	No
4.	1 or 2	1	Yes (5)
5.	6	2	No

Figure 1: Details of the five lists used for the experiment. The actual lists are given in Appendix A.

The recall experiment was done online. The presentation modality was visual. 20 subjects participated in the experiment. They were shown the list words at two second intervals by use of an applet. After each list, they had to fill up a form, writing down as many words as they could remember, and giving a confidence rating (1-4) for each word. On submitting this form, the display of the next list would begin. All 5 lists were presented in this way. The same lists (given in Appendix A) were used for both the recall and recognition experiments.

The methodology adopted had some limitations, due to which the results obtained cannot be regarded as necessarily very accurate or representative. These limitations were:

1. The number of subjects was not very large, implying that statistically, the data obtained is not very robust.
2. The recognition experiment was done online, and the conditions under which it was taken by the subjects could not be verified.
3. The recall experiment too was carried out in hostel rooms, and as such was not entirely free from disturbance.
4. All the subjects were from the same demographic profile, male college students between the ages

	Cluster Size	Studied	Unrelated	Weakly related	Critical word
List 1	12	3.89	1.00	1.09	2.41
List 2	1	3.69	1.00	1.29	2.15
List 3	3	3.72	1.00	1.03	2.26
List 4	1 or 2	3.65	1.00	1.15	2.24
List 5	6	3.43	1.09	1.09	2.41

Figure 2: The results of the recognition experiment. The ratings are the averages over all the subjects, and are on a scale of 1 to 4, 1 denoting definitely not present and 4 denoting definitely present. Detailed results are in Appendix B.

	Studied words		Critical words	
	Percent Recall	Mean Confidence	Percent Recall	Mean Confidence
List 1	69.9	3.91	38.9	3.29
List 2	60.2	3.90	19.4	2.57
List 3	65.8	3.83	17.5	2.71
List 4	65.2	3.87	35.3	3.67
List 5	70.8	3.93	9.4	2.33

Figure 3: The results of the recall experiment. The figures are the averages over all the subjects. The confidence ratings are on a scale of 1 to 4, 1 denoting low confidence and 4 denoting high confidence.

of 18-22. Although there is no reason to expect the trends to have been different for any other group of subjects, the actual numbers may well have varied.

4 Results

The results of the recognition experiment are summarized in Figure 2, and presented in more detail in Appendix B. It was found that there is a difference in the extent to which the critical words are rated as having been on the list, based on the nature of the list. It can be seen that the lists where the critical words received the highest mean rating (i.e. the lists where the confidence in those words having been there was highest) were lists 1 and 5. These were also the lists with the largest sized clusters of related cues, 12 and 6 words respectively. On the other hand, the lowest mean rating for critical words was found to be in list 2, which had no clusters even of size 2.

The results of the recall experiment are summarized in Figure 3. The first noticeable thing is that the recall of the critical words is significantly lower than their recognition was. Also, the amount of recall does not follow the same pattern as for the recognition test; here, recall of the critical words is least for list 5 (with clusters of 6). In fact, the maximum critical word recall is found to occur for lists 1 and 4, the two lists with one critical word each (all the others have two each). The average confidence in the critical words is also maximum for these two lists, being only a bit less than the average confidence for the actual list words.

These results are analysed in the next section, where I also look at their implications on my hypothesis.

5 Analysis

The recognition experiment results seem to be in accordance with the stated hypothesis. Although it is not very large, there is a difference between the extent of false memory formation under the different conditions. The maximum effect is seen when the clusters are largest, although there seems to be no difference between the cases with a cluster of 12 (i.e. all the words on the list related to the critical word) and clusters of 6 (i.e. the first half of the list related to one critical word and the next half to the other). This may show that the activation of mental connections saturates after a point, and further cues after that do not matter much. Also, the effect decreases slowly but steadily with decrease in cluster size below 6, as expected. Models like the spreading activation and compound cue models (Neath [2], chapter 10) can account for these results. One way of looking at it might be that activations fade out slowly, but can become stronger if quickly reinforced. When several related cues are received in succession, the same mental image gets activated quickly from several directions, leading to a stronger impression.

One thing is clear: false memories are formed, and subjects can be quite confident of them despite

their total inaccuracy. This is seen across all the lists, so it cannot really be said that clustering of cues is a pre-requisite for false memory formation: they can be formed even when the cues come in a staggered fashion, even if they are slightly weaker in that case. These results clearly reinforce the findings of Roediger and McDermott [3], though false memory formation is found to be not as strong here. This may be due to differences in the nature of the subjects.

The results of the recall experiment are more puzzling. Here the pattern is quite different. In fact, there seems to be no clear-cut pattern. One thing of interest, however, is that recall of the critical word(s) is highest for the two lists which had only one such word each. For both these lists, over 35 percent of the subjects recalled the critical word as having been on the list, while in none of the other lists was this figure over 20 percent. This may suggest that during a recall task, there is interference between cues relating to multiple critical words, while more coherent false memories are formed if only a single critical word is present. During recognition, the critical words are explicitly presented as cues, which can serve to activate the associated memory trace, even if it is weak, but this is not possible during recall. This of course also explains the fact that the false memories show up more strongly during recognition.

Another point to note is that the memory implantation effect does show up fairly significantly in the recall experiment, which involved visual presentation of data. This suggests that words presented visually are encoded semantically, much as words presented via the auditory modality are.

6 Conclusions

The results of the recognition experiment support the hypothesis that the extent of false memory formation is influenced by the distribution of the associated cues, being stronger if the cues appear clustered together. However, this influence does not appear to be very strong, and may well be within the margins of experimental error. It is clear that false memories do get formed under several kinds of conditions, and multiple ones can even be implanted in parallel. However, the presence of multiple critical words appears to affect the recall of such memories; during recall, the implantation effect shows up most strongly for lists with just one critical word. The one thing that does come up most clearly, yet again, is that memory is largely reconstructive, and as a consequence, we can remember, and be quite confident about, things that never actually happened.

7 Acknowledgements

I would like to thank Dr. Harish Karnick for his guidance and motivation. I would also like to thank all the people who consented to be subjects for my experiment, most of them being undergraduate students at IIT Kanpur.

References

[1] Loftus, E.F. (1993). The reality of repressed memories. *American Psychologist*, 48, 518-537.

[2] Neath, I. and Surprenant, A.M. (2003). *Human Memory: An Introduction to Research, Data and Theory, second edition.* Belmont, CA: Wadsworth.

[3] Roediger, H.L. and McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 803-814.

Appendix A

Word Lists

	I	*II*	*III*	*IV*	*V*
1	table	thread	hill	bed	sing
2	legs	bumpy	valley	rest	music
3	seat	pin	summit	computer	tune
4	desk	tough	sour	awake	melody
5	sofa	sewing	candy	dream	album
6	wood	smooth	sugar	library	sweet
7	cushion	sharp	top	gold	blue
8	stool	road	peak	yawn	space
9	sitting	point	glacier	sing	cloudy
10	bench	coarse	bitter	nap	stars
11	couch	prick	taste	snore	sunny
12	folding	uneven	honey	waterfall	clear

Recognition Task Lists

	I	*II*	*III*	*IV*	*V*
1	legs	cloth	range	**sleep**	air
2	**chair**	smooth	summit	rest	sweet
3	arms	thread	**sweet**	coma	**sky**
4	apple	**needle**	forest	yawn	music
5	plastic	banana	book	gold	track
6	cushion	**rough**	sugar	night	orange
7	sitting	eye	**mountain**	snore	sunny
8	couch	pin	bitter	shirt	black
9	shoe	rude	cute	machine	**song**
10	pillow	prick	honey	baby	space
11	sofa	fair	dessert	dream	flute
12	soft	sugar	star	alarm	paper

Note: Critical words are in boldface.

Appendix B
Detailed Results of the Recognition Experiment

Rating given	**4**	**3**	**2**	**1**	**Mean**
List 1					
Studied	96.5	0	0	3.5	3.89
Unrelated	0	0	0	100	1.00
Weakly related	1.5	0	4.4	94.1	1.09
Critical	41.2	5.9	5.9	47.1	2.41
List 2					
Studied	88.2	0	4.4	7.4	3.69
Unrelated	0	0	0	100	1.00
Weakly related	5.9	0	11.8	82.4	1.29
Critical	32.4	5.9	5.9	55.9	2.15
List 3					
Studied	83.8	8.8	2.9	4.4	3.72
Unrelated	0	0	0	100	1.00
Weakly related	0	0	2.9	97.1	1.03
Critical	32.4	5.9	17.6	44.1	2.26
List 4					
Studied	85.9	2.4	2.4	9.4	3.65
Unrelated	0	0	0	100	1.00
Weakly related	0	2.9	8.8	88.2	1.15
Critical	23.5	17.6	17.6	41.2	2.24
List 5					
Studied	75	5.9	5.9	13.2	3.43
Unrelated	0	0	8.8	91.2	1.09
Weakly related	0	2.9	2.9	94.1	1.09
Critical	38.2	8.8	8.8	44.1	2.41

Figures in columns 2-5 are percentages.
Note: Percentages may not always add up to 100 due to rounding.