

# Are Experts Unbiased? Effects of Knowledge and Attitude on Memory for Text

Jennifer Wiley

Department of Psychology and  
Learning Research and Development Center  
University of Pittsburgh  
Pittsburgh, PA 15260  
wiley@pitt.edu

## Abstract

Subjects with varying amounts of domain knowledge read texts on two controversial issues: whether the US should participate in the Persian Gulf War and whether abortion should be legal. Each text contained ten arguments for each side of the issue. Subjects with the most knowledge about the topics recalled roughly equal numbers of arguments from either side of the issue, while subjects with less knowledge recalled more arguments for the side they agreed with. The results were replicated with a third topic, the OJ Simpson case. The results of both experiments suggest that recall bias due to attitude may be eliminated by the possession of domain knowledge. Implications for instructional programs using expert models are discussed.

## Experiment 1

Recent conceptions of expertise have stressed that it is not just the sheer quantity of domain-relevant knowledge possessed that separates experts from novices, but also the organization, integration and accessibility of that knowledge (e.g. Bedard & Chi, 1992). Ericsson and colleagues (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989) have advanced a "skilled memory" theory which suggests that as people acquire content knowledge, they also develop mechanisms, like retrieval structures, that enable them to use their knowledge. It is argued that the presence of such a mechanism enables experts to rapidly encode, store and retrieve information related to their domain. One possible consequence of well-developed retrieval structures among experts is that expert recall may be qualitatively different than that of novices. In particular, the present study investigates whether expertise can eliminate (or reduce) the influence of attitude on recall.

Since the early days of psychology, it has been observed that information which is consistent or congruent with our attitudes, opinions or beliefs is better remembered than attitude-opposing information. Bartlett (1932) noted that "When a subject is being asked to remember, very often the first thing that emerges is something of the nature of attitude. The recall is then a construction made largely on the basis of attitude, and its general effect is that of a justification of the attitude." (pp. 206-7) However, over fifty years and fifty experiments later, empirical findings on "consistency bias" provide no clear support for these ideas, with some studies finding better recall for attitude-supporting information, and other studies finding no differences at all (See Pratkanis, 1989; Roberts, 1985 for further discussion). One possible reason for the failure to find a consistency bias effect overall

may be that consistency bias is moderated by the possession of domain knowledge.

The present study was designed to examine whether expertise (the possession of a large amount of domain knowledge) can reduce or eliminate consistency bias related to attitude in the recall of text. The position taken is that experts, because they have the advantage of a well-developed and accessible store of domain knowledge, would not need to use their attitude as a cue for retrieval. Thus, experts would not be susceptible to consistency bias. Nonexperts, however, lack a well-developed, accessible store of domain knowledge, and may default to attitude as a basis for encoding or retrieval. Thus, nonexperts would be susceptible to consistency bias.

Three specific hypotheses are made with respect to this general prediction that consistency bias should be moderated by the possession of domain knowledge. First, nonexperts should recall more information consistent with their attitudes than inconsistent while experts should recall at least as many attitude-inconsistent items as consistent.

Second, less-educated nonexperts may show more bias than more-educated nonexperts. In order to separate the effects of possession of domain knowledge from general ability or motivation differences, experts in a specific domain read two kinds of texts. Some experts served as subjects with a large amount of domain knowledge on a text within their domain, while some read a text outside their domain, serving as more-educated nonexperts.

Undergraduates in an Introductory Psychology course served as less-educated nonexperts for both texts. Even though more-educated nonexperts may not possess more domain knowledge than the less-educated nonexperts, they may better appreciate opposing arguments. Thus, the more-educated nonexperts may recall more opposing arguments, making them less biased than their less-educated counterparts.

Third, recognition performance should not be affected by the reader's attitude or amount of domain knowledge. All information should be encoded to some degree by all subjects. As a result, roughly equal numbers of consistent and inconsistent items should be recognized by all subjects.

## Method

### Subjects

Twenty undergraduates at the University of Pittsburgh enrolled in an Introductory Psychology course, 20 graduate students in Political Science, and 20 Law students served as subjects. Subjects were paid for their participation.

## Materials

Two texts were written of approximately equal length. The first, on abortion, was 533 words and titled "Should abortion be legal?"; the second on the Persian Gulf War was 548 words and titled "Should the United States use force against Iraq?" The topics were chosen as arguments about them would be familiar and of interest even to nonexperts. Also, presentation in the form of a prose passage was chosen over a list of statements in the interest of keeping the task more incidental, and hence, naturalistic.

Each text contained 20 arguments, 10 supporting each side of the issue. Arguments for either side of the issue were intermingled throughout the text. Because coherence of the text was a primary consideration, the exact pro/con ordering was not identical for the two texts. Most importantly, arguments for each side of the text were matched for strength and familiarity through piloting. For each text, recognition tests were developed consisting of the 20 arguments from the text and 10 distractor items (5 for each side). Recognition items were not taken verbatim from the texts, rather they were rephrased to prevent surface recognition. In addition, a 10-item knowledge test was developed for each topic.

## Procedure

The experiment took place in 1991 after the war had ended. Subjects were told they were piloting materials and rated the familiarity of each sentence of the text on a 10-point scale while reading. Familiarity judgement was chosen as a task because each sentence would be read, but as opposed to an agreement judgement task, would not necessarily invoke attitude-based processing.

Subjects first read and rated the experimental text. After a 30 minute filler task, subjects were asked to write down as much as they could of the article they read at the beginning of the session, referred to by title, and prompted to remember as many points from the article as possible. When finished recalling, subjects were given a 30-item recognition test and a 10-item domain knowledge test on the same topic as the text they had read. Position on issues was assessed using a 10-point scale before and after the task. Before the task subjects were asked to rate their opinion on many issues, and the task-relevant issue was embedded in this list.

## Design

Ten undergraduates (less-educated nonexperts), 10 Political Science graduate students (more-educated nonexperts), and 10 Law students (experts) read the text on abortion; while 10 undergraduates (less-educated nonexperts), 10 Law students (more-educated nonexperts) and 10 Political Science graduate students (experts) read the text on the Persian Gulf War. Thus, three levels of expertise were split across two texts to yield six independent cells.

Each text contained 20 arguments in total, 10 arguments for each side of the issue. Consistency with position was tested as a within- subjects variable. Thus, the design was a 3x2x2 (Expertise x Text x Consistency) mixed design with repeated measures on the last variable.

## Results

### Domain Knowledge

A two-factor analysis of variance (ANOVA) using level of expertise (Less-educated nonexpert, More-educated nonexpert, Expert) and text (Abortion, Persian Gulf) performed on the results of the 10-item domain knowledge test revealed no main effect for text,  $F(1,56)=.01$ ,  $MSE=1.90$ ,  $p=.93$ , but a highly significant main effect for knowledge,  $F(2,56)=49.67$ ,  $MSE=1.90$ ,  $p<.0001$ . Tukey's HSD test of pairwise comparisons indicated that less-educated nonexperts answered significantly fewer items correctly ( $M=3.65$ ) than the more-educated nonexperts ( $M=6.50$ ,  $p<.01$ ), who, in turn, answered significantly fewer items correctly than the experts ( $M=7.85$ ,  $p<.01$ ). Thus, the expertise groups represented three distinct levels of domain knowledge.

### Recall

A 3x3x2 mixed ANOVA (Knowledge Level, Text, Consistency with Attitude) on accurate recall of arguments from the texts revealed a highly significant main effect for knowledge level,  $F(2,56)=10.84$ ,  $MSE=4.65$ ,  $p<.0001$ . Post-hoc comparisons indicated that experts ( $M=9.6$ ) recalled significantly more arguments than either the more-educated nonexperts ( $M=7.9$ ),  $t(38)=2.70$ ,  $p<.01$ , or the less-educated nonexperts ( $M=5.15$ ),  $t(38)=7.08$ ,  $p<.0001$ .

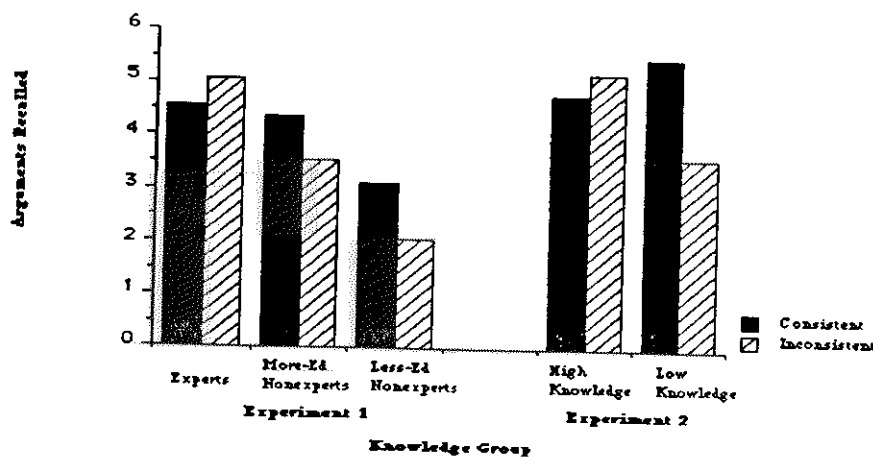


Figure 1: Mean number of consistent and inconsistent arguments recalled by knowledge group in Experiments 1 and 2

In addition, the more-educated nonexperts recalled more arguments from the text than the less-educated nonexperts,  $t(38)=4.38$ ,  $p=.0001$ . Thus, as domain knowledge increased so did the number of arguments recalled. A significant main effect was also found for text,  $F(1,56)=5.82$ ,  $MSE=4.65$ ,  $p=.02$ , as subjects recalled more arguments from the Abortion text ( $M=8.50$ ) than from the Persian Gulf text ( $M=6.60$ ).

The most interesting result, however, as predicted by the first hypothesis, is the significant interaction between knowledge level and consistency with position,  $F(2,57)=3.51$ ,  $MSE=1.98$ ,  $p=.04$ . Figure 1 presents the mean number of attitude-consistent and attitude-inconsistent arguments recalled for each knowledge level. Planned comparisons indicated that the less-educated nonexperts exhibited a significant consistency bias, recalling more information supporting their position than opposing it,  $t(19)=2.36$ ,  $p=.02$ . More-educated nonexperts also tended to recall more supporting than opposing information, but this bias was only marginally significant,  $t(19)=1.80$ ,  $p=.08$ . Experts, on the other hand, recalled slightly less supporting than opposing arguments but this difference was not significant,  $t(19)=-1.13$ ,  $p=.27$ . Planned comparisons on the difference of consistent versus inconsistent arguments recalled between groups showed that experts were significantly less biased than either less-educated nonexperts,  $t(38)=-2.47$ ,  $p=.02$ , or more-educated nonexperts,  $t(38)=2.07$ ,  $p=.04$ . There was no difference in bias of less-educated and more-educated nonexperts,  $t(38)=.40$ ,  $p=.69$ .

Thus, in support of the first hypothesis, subjects with the least domain knowledge had a consistency bias in their recall of the text, while the most knowledgeable subjects had none. Further, the experts were significantly less biased than either nonexpert group. Support was not found for the second hypothesis, however, as no differences were found between the biases of the nonexpert groups.

### Recognition

The mean number of arguments that were correctly recognized as being in the text was 16.05 for the experts, 16.10 for the more-educated nonexperts, and 15.15 for the less-educated nonexperts. A 3x2x2 mixed ANOVA (Knowledge Level, Text, Consistency with Attitude) on recognition of items from the text revealed that no main effects or interactions approached significance ( $F_s < 1.59$ ). These results support the third hypothesis and suggest that the bias observed in the recall of the less-educated nonexperts was not due to selective attention or exposure on the basis of attitude or knowledge, as all subjects were comparable in their ability to recognize the majority of both consistent and inconsistent items from the text.

### Conclusions

The results of Experiment 1 suggest that recall consistency bias in favor of attitude-supporting information is moderated by the possession of domain knowledge. Subjects with the least amount of domain-specific knowledge had a significant bias in their recall due to attitude, while no bias due to attitude was observed among the expert subjects. Meanwhile, more-educated nonexperts fell between the experts and less-educated nonexperts both in domain knowledge as well as in recall bias.

It must be noted that these results may be less than

convincing due to the fact that subjects' position on the issues was not controlled and there were differences among the knowledge groups in positions taken on the Persian Gulf issue. In relation to the abortion issue, the knowledge groups were fairly similar. On the Persian Gulf issue, however, the nonexperts were somewhat for the war while the experts were mostly against it.

## Experiment 2

In Experiment 2, position on issue was controlled such that at each knowledge level, half the subjects agreed with each side of the issue. Also of interest was whether the effect observed in Experiment 1 would generalize to a new topic.

### Method

#### Subjects

20 undergraduates at the University of Pittsburgh participated in this experiment as part of the Introductory Psychology subject pool.

#### Procedure

This experiment was a partial replication of Experiment 1 with a new topic, "Is OJ Simpson guilty", done in the fall of 1994. A 545-word text, 30-item recognition test and 10-item domain knowledge test were created. Subjects with 7 or more correct answers on the domain knowledge test were considered "high" knowledge; subjects with less than 7 were considered "low" knowledge. In each knowledge group, 5 subjects agreed and 5 disagreed that OJ Simpson was guilty beyond a reasonable doubt.

### Results

#### Recall

The mean number of arguments recalled was 10.0 for the high knowledge group and 9.1 for the low knowledge group. No significant difference due to level of knowledge was found,  $t(18)=-.88$ ,  $p<.39$ . As in Experiment 1, there was a significant effect of knowledge on recall bias. Figure 1 presents the mean number of attitude-consistent and inconsistent arguments recalled for each knowledge level. Paired  $t$ -tests indicated that the low knowledge group had a significant consistency bias, recalling almost 2 more arguments in favor of their position than opposing it,  $t(9)=4.38$ ,  $p<.01$ . The high knowledge group, however, recalled slightly fewer supporting than opposing arguments, but this reverse bias was not significant,  $t(9)=-.77$ ,  $p<.46$ . Further, a two sample  $t$ -test between knowledge groups indicated that the low knowledge group was significantly more biased than the high knowledge group,  $t(18)=3.40$ ,  $p<.01$ .

#### Recognition

The mean number of arguments that were correctly recognized as being in the text was 16.2 for the high knowledge group and 16.0 for the low-knowledge group. No differences were found in recognition due to knowledge or position.

### Conclusions

The results of Experiment 2 clearly replicated the effect found in Experiment 1: high knowledge subjects had no bias

in their recall, while low knowledge subjects did. Further, these results were obtained on a new issue and with an equal number of subjects on either side of the issue on each knowledge level.

### General Discussion

The results of the two experiments clearly indicate that consistency bias in recall due to attitude may be moderated by domain knowledge, with the most knowledgeable subjects showing no bias, and less knowledgeable subjects showing a significant bias. It is suggested that recall consistency bias is the result of the less knowledgeable subjects relying primarily on their evaluation of a topic as a basis for encoding, retrieving, or reconstructing the text, while more expert subjects have the advantage of using their domain knowledge as the principal basis for text processing and retrieval.

The observed qualitative difference in the recall of the experts and nonexperts has implications for research in Cognitive Science on the modeling of expert performance or processing, and the use of those models to train nonexperts. The present results suggest that expert models may not be appropriate tools in cases where attitudes are involved, as nonexperts may have difficulty using their knowledge in a manner similar to experts. On the other hand, the demonstration of qualitative differences in the memory performance of experts and nonexperts serves to inform models of expertise, as it suggests which aspects of knowledge representation or processing may be changing as one gains expertise.

### References

- Bartlett, F. C. (1932). *Remembering*. Cambridge, England: Cambridge Univ Press.
- Bedard, J. & Chi, M. T. H. (1992). Expertise. *Current Directions in Psychological Science*, 1(4), 135-9.
- Chase, W. & Ericsson, K. A. (1981) Skilled Memory. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition*. Hillsdale: Erlbaum.
- Ericsson, K. A. & Stasewski, J. (1989) Skilled memory and expertise. In D. Klahr & K. Kotovsky (Eds), *Complex Information Processing*. Hillsdale: Erlbaum.
- Pratkanis, A. R. (1989). The cognitive representation of attitudes. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.) *Attitude Structure and Function*. Hillsdale: Erlbaum.
- Roberts, J. V. (1985). The attitude-memory relationship after 40 years: A meta-analysis of the literature. *Basic and Applied Social Psychology*, 6(3), 221-241.