Expert-Novice Differences on a Recognition Memory Test of Physics Diagrams

Adam Feil and Jose Mestre

Departments of Physics and Educational Psychology, University of Illinois at Urbana-Champaign, Loomis Laboratory of Physics, 1110 West Green Street, Urbana, IL 61801

Abstract This study used a recognition memory test and specially constructed pairs of physics diagrams to measure differences between physics experts and novices (who have not taken college physics) in the way simple physics diagrams are encoded in memory. Results show that although physics experts encode aspects and features of physics diagrams that novices do not, in some specific cases novices encode features that experts do not. Physics experts were more likely to encode features of diagrams that were more relevant to the physics depicted. This suggests that the knowledge and experience of physics experts influences the way in which they conceptualize physics diagrams, even in the absence of a question prompt.

Keywords: physics education research, expert novice differences, recognition memory
PACS: 01.40.Fk, 01.40.Ha

INTRODUCTION

Many differences between experts and novices have been identified in a variety of domains, including physics [1]. The current study seeks to identify differences between physics experts and novices in the way simple physics diagrams are conceptualized and encoded in memory. In other words, the study attempts to answer the question, "Do experts and novices focus on different things when looking at the same picture?" In addition, the current study used Signal Detection Theory (SDT), a widely-used method of analysis for certain kinds of tasks, such as the recognition memory test used in this study. A secondary goal of this paper is to provide an example of how SDT can be a useful tool in conducting physics education research.

LITERATURE REVIEW

Expert-novice studies reveal the superiority of expert performance in both verbal and non-verbal tasks. For example, in non-verbal tests of memory experts are able to "chunk" bits of strategically related information into a single unit. This allows experts to outperform novices on many kinds of recall memory tests, such as recalling the layout of chess pieces or the components of complex circuit diagrams [1]. Instead of remembering the location of many individual chess pieces, a chess master can chunk groups of pieces that make up a certain strategic arrangement that he or she is familiar with. Similarly, an electronics technician can chunk a collection of electrical components that make up an amplifier. In his study of chess masters, de Groot noted that much of the thinking done while playing chess is non-verbal, and that chess masters are not used to providing verbal explanations for the moves they make [2], suggesting that intuitive, non-verbal behavior comes with expertise.

It is likely that much of the thought-process of physics experts during problem solving is also non-verbal, despite their ability to explain verbally much better than novices the steps made while solving a problem. The current study does not rely on physics experts and novices verbalizing what they are thinking; rather, a memory test is used as a way to implicitly determine which features of physics diagrams are encoded.

Knowledge organization in memory contributes to the non-verbal prowess of experts. Chi, Feltovich and Glaser [3] studied the contents of schemata among expert and novice physicists and found that experts' schemata were organized around physics principles. In contrast, novices' schemata did include physics principles but not as a central component. The better organization and emphasis on physics principles found in the schemata of physics experts may be due in part to factors other than knowledge structure. Physics experts are consciously aware that physics is built on relatively few principles and are used to discussing physics concepts in this way.

This study extends explorations of non-verbal skills of experts by implicitly measuring what experts and novices encode when looking at physics diagrams during short exposures. This approach will ensure that any differences found between experts and novices are not a result of factors such as the superior vocabulary or descriptive ability of physics experts.

Memory tests have been used before to study expert-novice differences in various domains [1,4]. These studies have focused on experts' ability to use their extensive long-term memory, which has been formed over years of exposure to various patterns and situations in their domain of expertise, to "chunk" multiple items in a stimulus into a single item. This study uses a different kind of memory test, namely a recognition memory test. In this type of memory test, subjects must decide if the current stimulus is identical to one shown earlier. This type of test will better show differences in the features of diagrams that are encoded by experts and novices.

METHODS

Subjects included 18 novices and 18 experts. Novices were recruited through an educational psychology subject pool. These subjects were both undergraduates and graduate students enrolled in an educational psychology course. The vast majority of these subjects had never taken a college physics course. Experts were graduate students in physics who had recently been TAs in discussion sections of introductory physics courses.

Twenty pairs of images were created. Ten pairs were designed so that the two pictures were slightly different in appearance, but with no meaningful difference in the physics depicted, and ten pairs were designed to include some difference in the physics. Because only simple drawings were shown, the physics differences were fairly subtle. Figures 1-3 are examples of pairs used in the study.

Procedure

A computer program was written to administer the memory test. Subjects were shown a slideshow of 20 images on the screen. Each image was shown for 10 seconds, and a 2-second blank screen was inserted between each image. Subjects were instructed to provide a verbal description of each image as they saw them. These descriptions were audio-recorded and analyzed later. One image was randomly selected from each pair, and those 20 images were shown in a random order.

After the study-portion, subjects were occupied with a worksheet of addition problems for five minutes.

During the test-portion, subjects were shown 20 images and asked to indicate whether each image was "old", meaning it was exactly the same as an image shown before, or "new", meaning it was similar but not identical to an image shown before. Of the 10 pairs from each of the two types, surface-feature changes and physics changes, 4 of the images that were shown at study were also shown again at test, and 6 alternate images were shown. These images were presented in a random order.

Subjects were instructed to say out-loud any differences they remembered for images that they indicated were "new". These statements were also audio-recorded. For example, a subject may say, "Before, it was a cylinder rolling, and now it's a sphere." After making their "old" or "new" judgment, subjects indicated how confident they were with their choice using a 1-5 scale.

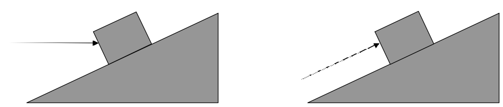

FIGURE 1. This diagram pair contains a physically meaningful difference. The angle of the applied force is an important factor in determining the normal and frictional forces on the block, and the acceleration of the block.

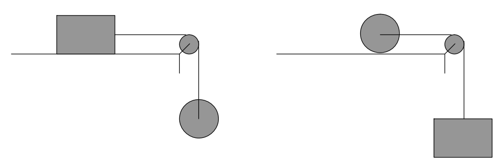

FIGURE 2. This diagram pair contains a physically meaningful difference. The diagram on the right may involve rotational dynamics while the diagram on the left does not.

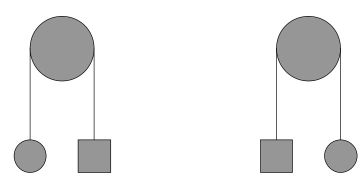

FIGURE 3. This diagram pair differs only in surface features. Switching the square and circle blocks does not have any effect on the physics depicted.

Signal Detection Theory

Data from this type of recognition memory test is best analyzed using Signal Detection Theory (SDT) [5]. There are four possible outcomes from each trial in the memory test, depending on whether the image shown during the test is the same or different as the one shown during the study-portion and whether the subject responds "old" or "new". Figure 4 demonstrates the four possible outcomes for each trial.

		Picture on test:	
		New	**Old**
Subject's answer choice:	**New**	Correct Rejection (CR)	Miss
	Old	False Alarm (FA)	Hit

FIGURE 4. There are four possible outcomes for each trial because the picture shown can be either old or new and the subject's response can be either old or new. For example, if the picture shown at test is old, meaning the same as shown at study, and the subject responds "old", that trial is scored as a Hit.

In a test situation such as the one in this study, a subject's performance on the test can be influenced by his or her tendency to answer "old" or "new" in situations where the subject is uncertain of the answer. Because there may be differences in bias between experts and novices, it is important to use SDT to analyze the results of the test to help ensure the validity of the findings.

SDT provides a measure of performance that takes into account Hits and False Alarms. This measure, d', indicates how well a subject can distinguish between signal and noise. This measure does not depend on each subject's tendency to answer "old" or "new", which can differ significantly among individuals. In the following sections, d' will be used as the measure of performance on the memory test.

A d' value of 0 represents no ability to distinguish between signal and noise. In this experiment, a subject would have a d' of 0 if he or she had a Hit Rate equal to his or her False Alarm Rate. As other examples, a subject with a Hit Rate of 75% and a False Alarm Rate of 33% would have a d' of 1.1, and a subject with a Hit Rate of 75% and a False Alarm Rate of 17% would have a d' of 1.6.

RESULTS

T-tests were used to determine the significance of the results reported. Results (see Figure 5) show that novices do not perform significantly better on pairs that include physics differences than they do on pairs that differ only on surface-features. Physics experts, however, perform significantly better on pairs that include physics differences than they do on pairs that differ only on surface-features ($p = .024$). Experts performed marginally better than novices on surface-feature only pairs ($p = .062$). Unsurprisingly, experts performed significantly better than novices on pairs that include physics differences ($p = .005$).

In addition to these main findings, an examination of performance on particular image pairs showed two cases where physics novices outperformed physics experts. Both of these diagram pairs depicted blocks on either side of a double-incline, connected to each other by a string that went through a pulley at the top of the incline. In one diagram, the blocks were at equal heights, and in the other diagram, the blocks were moved so that one hung lower than the other. Figure 6 shows one of these image pairs.

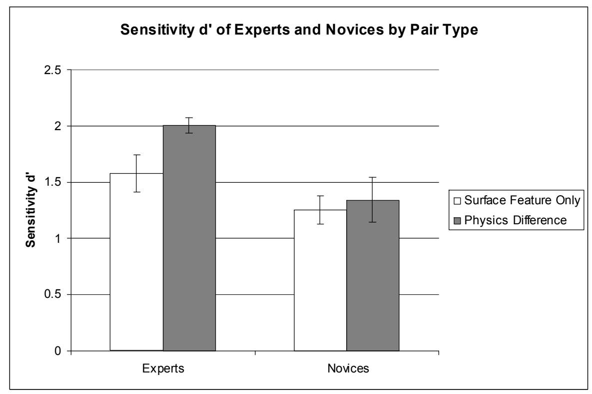

FIGURE 5. Memory test results for physics experts and novices.

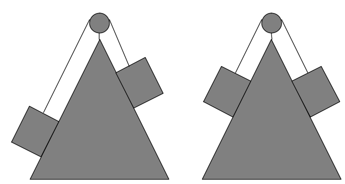

FIGURE 6. Novices were better able to distinguish between these two diagrams than experts, suggesting that the symmetry of the hanging blocks may be a feature considered relevant to novices but not experts in this case.

The descriptions provided by experts and novices during the study-portion of the experiment, in general, were very similar. Because no question prompt was provided, physics experts did not use language describing physics principles or concepts. In most cases, both experts and novices simply gave a description of the surface features of the picture. Some novices did not recognize pictures of circuit diagrams; one subject described a simple resistor circuit diagram as a "heart monitor" because the symbol for a resistor resembles the display of a heartbeat monitor.

Verbal descriptions were coded on a fairly objective criterion: Does the description given for the picture shown during study distinguish between the two pictures of the pair. That is, a description was coded as discriminating if it is only an accurate description of one of the diagrams from a pair. If the description could also be used to correctly describe the other diagram in the pair, it was considered non-discriminating. Figure 7 shows the fraction of discriminating descriptions given by experts and novices. Both groups gave more discriminating descriptions for "physics difference" pairs than for "surface-feature only" pairs ($p < .0001$ for experts, $p = .013$ for novices). Experts gave more discriminating descriptions on "physics difference" pairs than novices ($p = .012$).

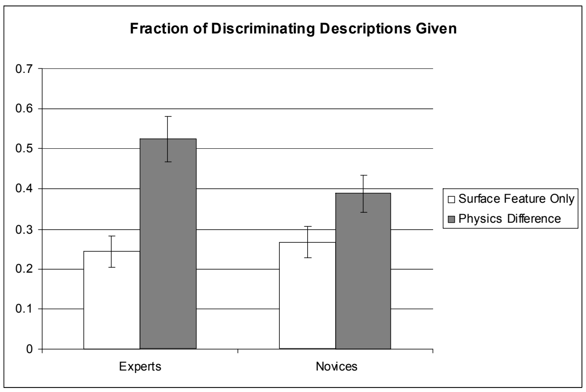

FIGURE 7. The pattern of discriminating descriptions among experts and novices for the two types of diagram pairs is similar to performance on the memory test.

DISCUSSION

The experience and knowledge of individuals causes them to see physics diagrams in a certain way, focusing on different features and eliciting different knowledge depending on expertise. The main difference between experts and novices is that experts have learned which features and relationships shown in a diagram are relevant to the physics and which are not. Because the diagrams used in this study were very simple and no question prompt was given to subjects, it is not sufficient to describe the difference between experts and novices in terms of experts focusing on physics principles and novices focusing on surface features. To a large degree, surface features were all that any of the subjects described while looking at each diagram, so the differences observed between the two groups in this study are due to more subtle mechanisms. The activation of relevant deep structure from viewing surface features in diagrams likely happens automatically for experts, consistent with their hierarchically organized knowledge that bundles together chunks of related contexts, concepts/principles, and problem solving strategies.

Even though both experts and novices encoded surface features when viewing the diagrams, the two groups apparently encoded different relationships among the objects in the diagrams. For example, when describing the diagram on the left shown in Figure 1, one novice subject said, "A force pushing on a block, pushing it up the wedge." One expert subject said, "A block on an inclined plane with a horizontal force." The only difference between these descriptions is that the expert specified the exact direction of the force and novice did not. Our data indicate that when a subject's description for the diagram viewed during the study-portion can distinguish between the two diagrams in the same pair, the subject is almost always able to correctly reject the similar but different diagram shown on the test. This is true of both experts and novices. In the particular example shown in Figure 1, the direction of the applied force appeared to be an important characteristic for an expert to encode since it has implications for the behavior of the system.

The finding that novices outperformed experts on specific instances, such as the diagram pair shown in Figure 6, suggests that some relationships among objects in a specific context are considered irrelevant by experts and hence not encoded in memory, which is perfectly reasonable since attention and memory are limited. The ability to disregard irrelevant information makes any learning or problem-solving task much easier. On the other hand, the symmetry of the situation in Figure 6 was an attractive characteristic for the novice to encode, despite its irrelevance to the behavior of the system.

REFERENCES

1. *How People Learn*, edited by J. Bransford, A. Brown, and R. Cocking (National Academy Press, Washington, D.C., 1999).
2. A. de Groot, *Thought and Choice in Chess* (Mouton, The Hague, the Netherlands, 1965).
3. M.T.H. Chi, P. Feltovich, and R. Glaser, *Cognitive Science* **5**, 121-152 (1981).
4. D.E. Egan and B. J. Schwartz, *Memory and Cognition* **7**, 149-158 (1979).
5. T. D. Wickens, *Elementary Signal Detection Theory* (Oxford University Press, New York, 2002).