Contrasting Evaluations of Problem Solutions: Assessing Problem Solving With Diana

Kathleen Andre Harper
Physics Education Research Group, The Ohio State University

Experts and novices completed an interview task where they evaluated a student's solution to a mechanics problem. Instructors were more likely to make specific criticisms, where the students spoke more in generalities. Additional evidence indicates that novice problem solving knowledge consists of both conscious and tacit pieces.

Background

Previous physics problem solving research includes many expert-novice comparisons. Some results include: 1) Experts classify problems based on the deep structure (or underlying physics) while novices rely on surface features.[1] 2) Experts begin with general principles and work most of the problem algebraically before inserting numbers; novices start with specific equations and insert numbers earlier.[2,3] 3) Experts view problem solving as a process, where novices see it as a recall task.[4] This paper describes another such comparison.

Description of the Task

This task was inspired by Elby's work.[5] In studying student epistemology, he created a fictitious student, Diana. She was described to each subject as being "just like you," but was taking physics with the main objective of learning the material more deeply. His survey then asked questions related to Diana's study habits.

In the interviews described here, the student was given a problem solution and told to imagine that Diana had just worked it and handed it to him. In the instructor interviews, subjects were told that Diana was a student in their introductory physics course who was taking it pass/fail with the primary goal of learning physics more deeply. Diana had worked out the solution and brought it to the instructor. In both cases, the subject was told that Diana wanted to know whether the solution was a good one or not and asked how he would answer Diana. (Although subjects were of both genders, all will be referred to in the masculine to avoid confusion with Diana.)

Student subjects were volunteers from the Freshman Engineering Honors (FEH) program at The Ohio State University. They took two quarters of nontraditionally taught physics. "Lectures" were extremely interactive, laboratories consisted primarily of design labs, and recitations were mostly cooperative learning exercises. Students participated in a variety of qualitative and quantitative problem solving activities. The student interviews took place roughly five weeks after the completion of the second quarter of physics. Nineteen students participated, representing a full range of ability level, as indicated by final course grade. The majority of the students seemed comfortable with the physics concepts involved in the interview. The thirteen instructor subjects all had earned Ph.D.'s, and all but one were professors. They were a diverse sample with respect to both research area and primary teaching duties.

Diana's solution was a copy of an actual solution written by an FEH student to a problem on a final exam. There were two sections of the mechanics class, but all the subjects for this study were from one section. This problem appeared on the other section's final, so the subjects had not seen it prior to the interview. The problem, along with Diana's solution, is shown in Figure 1.

This particular solution was selected for two main reasons. First, it is quite sparse, leaving some of the solution steps unclear.

Second, it obtains the correct numerical answer; it was thought that an incorrect solution might cause the subjects to focus unduly on that aspect of the solution.

Analysis and Results

The comments and behaviors of the subjects were categorized; the categories relevant to this paper are described below.

Ambiguous notation: a comment that Diana's notation was not accurate enough, usually related to subscripts

Ask Diana: phrased a specific question for Diana about her solution

Hard to follow: indicated Diana's solution was difficult to understand, unclear, or confusing

Looks good: said after some analysis that Diana's solution looked good

Missing work: stated that Diana had not shown much work on the paper or that she needed to show more work

No words: spoke about the lack of words in Diana's solution

Right/wrong: made a statement about the correctness of Diana's final answer

Two answers: said it appeared as if Diana had calculated two answers

What principles: wanted Diana to indicate what principles she was using

Worked problem: worked out a solution of his/her own while evaluating Diana's

For each interview, the number of times an event occurred which fit into one of the categories described above was tallied. The percentages of each sample exhibiting a particular behavior were computed. Figure 2 summarizes these results. The two samples were compared using a chi-squared independence test. Highly significant findings are further discussed below.

The most significant difference ($p<<.005$) stemmed from the lack of words in Diana's solution. As the "No words" category shows, nearly 40% of the instructors commented on the lack of words in Diana's solution, but no students did. Although not statistically significant, instructors were also more likely to comment that Diana had not identified the physical principles she used.

Another highly significant difference was in the "Ask Diana" category ($p<.025$). Only one student phrased a question for Diana, but 38% of the instructors had specific questions for Diana about her solution. One of the experts even asked the interviewer to play the part of Diana, so that he could have a conversation like one he would have with a student. Most questions regarded Diana's thoughts, indicating that these expert subjects were aware that problem solving is a process that is not completely captured by what is written on paper. None of the students explicitly displayed this awareness. This relates well to previous research.[4]

The further differences outlined in Figure 2 indicate that students tended to speak in generalities about the solution. In contrast, the instructors, although sometimes making general comments, would elaborate further, pointing out specific areas of the solution that lacked information or aspects making it hard to follow. Although some of these results may be due to the instructors' greater familiarity with the material, these results also seem to be in agreement with previous work: instructors looked at the deeper structure of the solution while students focused on its surface features.[1]

Before the study, it was thought that novices might base their answer primarily on the final answer's correctness. This was not the case; about 60% of each population commented on whether the answer was right or wrong. Also, roughly 30% of each group said at the end of their analysis that the solution "looked good." The two populations also had similar percentages of subjects who worked a solution of his own to the problem, with just 6% more of the student sample exhibiting this behavior.

Although the raw numbers make it appear that the populations were similar in these two respects, fundamental differences in behaviors related to these measurements were observed. Most instructors who said Diana's solution looked fine were those who had little difficulty understanding it. They stated this view after just a few other comments. It should also be noted that three instructors stated directly that Diana's solution was NOT good, one even saying he would not call it a solution. In contrast, no students directly stated that the solution was poor. Some students who said the solution was good or right did so after a long struggle of trying to comprehend the solution.

Although similar fractions of each sample wrote a solution of his own, the way in which this solution was written and utilized differed greatly between the two groups. Most instructors who wrote one decided to do so either immediately or after looking at Diana's solution for a short while and worked the solution quickly. Two of them then used the self-generated solution to illustrate points about what was lacking in Diana's. Most students who wrote solutions did so alongside Diana's, comparing theirs to hers frequently. This also seems consistent with earlier work.[6] Three students obtained a wrong answer and decided *Diana's* solution was wrong. In contrast, the 2 instructors who made initial errors quickly found and corrected them. Although this difference is related to the greater ease with which physicists read the mathematical language of physics,[7] it also resonates with previous self-monitoring findings.[8]

In another aspect of this study, the novices were compared to themselves at a prior time. All the students had solved a similar problem approximately 15 weeks earlier. The observations each subject made about Diana's work were compared to his solution of this similar problem. One result is especially striking: any student who singled out a particular aspect of Diana's solution exhibited a favorable version of that behavior in his own similar solution. For example, 4 students said it was difficult to follow Diana's solution because her number insertion was unclear; all of these students did an excellent job in their solutions of showing exactly which numbers were being substituted for which variables. This same pattern was observed for students who commented on subscripts, displaying initial equations, and writing a complete solution. An analysis of additional problems from the mechanics final found that this was true for those problems as well.

The predictive power of the Diana exercise is that it can reveal some aspects of good problem solving which the subjects have been utilizing for a while. The areas of Diana's problem solving which were described as poor by the students were the ones that those students had been using well for over a quarter. However, failing to mention deficiencies in Diana's problem solving process told the researcher nothing about whether the subject had those same deficiencies or not. Also, it was not possible to predict a subject's responses to the Diana question by looking at his solution to the similar problem.

These comparison results indicate that novices' problem solving knowledge and strategies consist of both conscious and tacit components. The conscious pieces are those related to the noticed deficiencies in Diana's solution; the tacit knowledge is that which the students exhibited in their own solutions, but did not find lacking in Diana's. Along with giving new insight into students' knowledge structures, this indicates that researchers must evaluate a student's own solutions to reliably measure his problem solving behavior.

Acknowledgements

Thanks to Michael Wittmann for suggesting I try this approach and to one referee for his substantial input. I also acknowledge the feedback and support of the OSU Physics Education Research Group. This work was supported in part by NSF grants DUE-9653145 and GER-9553460.

References

[1]M. T. H. Chi, P. J. Feltovich, & R. Glaser, "Categorization and Representation of Physics Problems by Experts and Novices" Cog. Sci. **5**, 2, 121-152 (1981).

[2]J. H. Larkin, "Processing Information for Effective Problem Solving," Engineering Education **70**, 3, 285-288 (1979).

[3]F. Reif, J. H. Larkin, & G. C. Brackett, "Teaching General Learning and Problem Solving Skills," Am. J. Phys. **44**, 3, 212-217 (1976).

[4]R. Good, "Scientific Problem Solving by Expert Systems," J. of Res. in Sci. Teach. **21**, 3, 331-340 (1984).

[5]A. Elby, "Another Reason that Physics Students Learn by Rote," Am. J. Phys. **67**, 7, S52-S57 (1999).

[6]D. P. Simon & H. A. Simon, "Individual Differences in Solving Physics Problems" in *Children's Thinking: What Develops?*" ed. by R. S. Seigler. (Lawrence Erlbaum Associates, Hillsdale, NJ, 1978).

[7]A. Van Heuvelen, private communication.

[8]S. K. Reed, *Word Problems: Research and Curriculum Reform* (Lawrence Erlbaum Associates, Mahwah, NJ, 1999).

Figures

A 0.10-kg bullet is fired into at 1.90-kg block. The block is attached to a spring of force constant 1000 N/m. The block slides for 0.40 m while compressing the spring after the bullet runs into the block. Determine the bullet's speed before it hit the block. Assume that the gravitational constant is 10 m/s^2. You must show all of the work supporting your answer or no credit will be given.

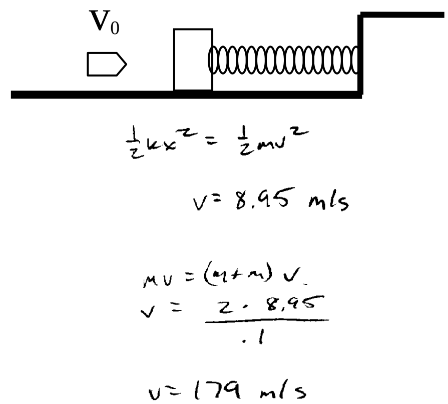

Fig. 1. The problem and Diana's solution

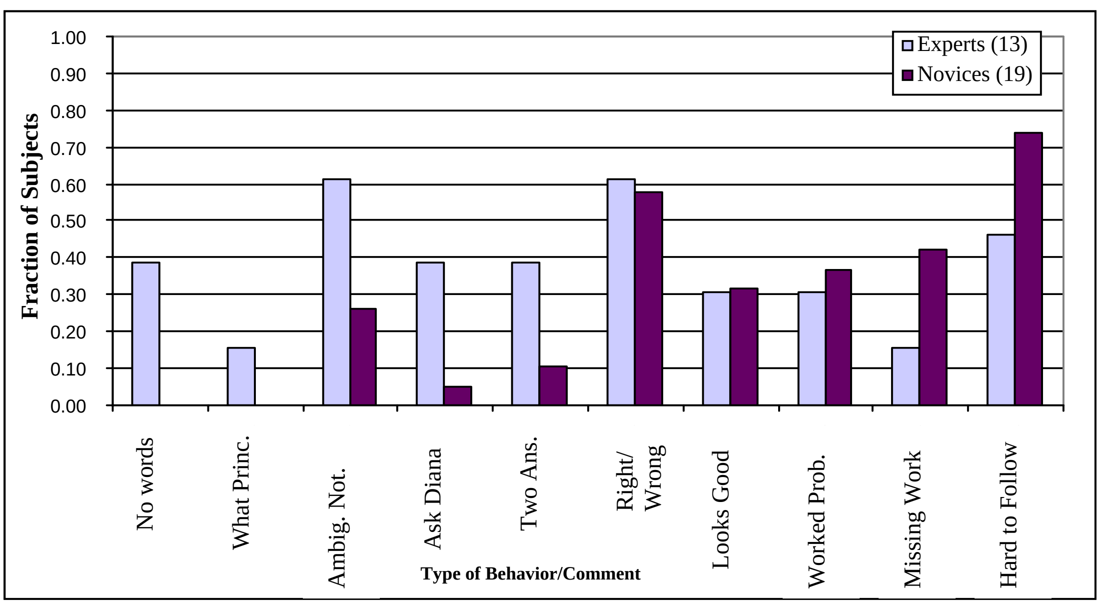

Fig. 2. Summary of notable expert and novice behaviors on the Diana task