Improving Problem Solving Performance by Inducing Talk about Salient Problem Features

PAUL S. STEIF, JAMIE M. LOBUE, AND LEVENT B. KARA
Carnegie Mellon University

ANNE L. FAY
Educational Assessment Consulting

BACKGROUND
Across many domains, research has shown that students often fail to select and apply appropriate conceptual knowledge when solving problems. Programs designed to support monitoring skills have been successful in several domains.

PURPOSE (HYPOTHESIS)
Critical conceptual knowledge in statics appears to be cued by paying attention to the bodies that are present in a problem, as well as to which ones are interacting and how. The research question addresses whether students can be induced to think about the bodies present, and whether focusing on bodies improves problem solving performance.

DESIGN/METHOD
Using a pre-post test design, written and verbal protocols were obtained for students solving problems before and after instruction. During instruction all students saw the same set of examples and corrected answers, but only the experimental group was asked questions designed to promote body centered talk. Solutions and protocols were coded and analyzed for frequency of body centered talk and solution quality.

RESULTS
The experimental group showed statistically significant increases in relevant body centered talk after instruction. Both groups improved their ability to represent unknown forces in free body diagrams after instruction, with the experimental group showing a greater, but not statistically significant, improvement. However, for both groups, the error rate in representing unknown forces at an interaction was significantly lower when a student referred to the bodies in the particular interaction.

CONCLUSIONS
Problem solving in conceptually rich domains can improve if, in addition to acquiring conceptual knowledge, students develop strategies for recognizing when and how to apply it.

KEYWORDS
concept, metacognition, problem-solving strategy, statics

I. INTRODUCTION

A significant component of engineering education and practice involves the activity of modeling. Modeling, in the context of engineering, involves approximating a real engineering system so that its performance or behavior can be estimated through the application of scientific principles and concepts. This paper addresses modeling and problem solving in one such subject, statics. Statics is a foundational course for many engineering disciplines, such as mechanical and civil engineering, and it is generally students' first experience with modeling. In particular, this paper articulates and tests the effectiveness of an approach to problem solving in statics that focuses students on problem features that link to relevant conceptual knowledge, and thereby reduces commonly committed errors.

Very recently, there has been an increased interest from the engineering education community in characterizing the conceptual basis of a number of fundamental engineering subjects. For example, researchers have sought to develop Concept Inventories, patterned after the widely used Force Concept Inventory (Hestenes, Wells, and Swackamer, 1992) in physics education research. These multiple-choice tests address the most important concepts in the subject, and are often the ones with which students routinely have difficulties or harbor misconceptions. However, even students who possess the important conceptual knowledge needed to solve problems may still fail to access that knowledge when applicable. In other words, their knowledge is "inert," present, but not represented in a way that makes it usable when faced with a problem to solve.

The ability to identify and apply appropriate conceptual knowledge when needed may be viewed as a metacognitive skill. Metacognitive skills refer to such mental activities as planning, monitoring comprehension, and evaluating one's progress towards a goal. Several studies across many domains, including physics (Chi, Feltovich, and Glaser, 1981; Larkin et al., 1980), history (Wineburg, 1991), and social sciences (Voss et al., 1984) have shown that a major difference between experts and novices is their metacognitive skills. For example, experts focus considerable time on the development of an appropriate and complete representation of the problem and selecting a strategy for approaching it prior to doing any detailed analysis. By accurately representing the problem, the expert constrains the range of possible analyses or procedures to apply and can quickly and successfully solve the problem. By contrast, novices tend to immediately jump to some type of detailed analysis, often based on a poorly conceptualized and incomplete representation of the problem. As a result, they may omit or mischaracterize key aspects of the problem and fail to successfully solve it.

Another important metacognitive skill of experts that distinguishes them from novices is their ability to monitor their current understanding and determine when that understanding is inadequate (Brown, 1980; Flavell, 1985, 1991). A number of researchers have successfully developed and implemented programs to support students' monitoring skills to improve learning and problem solving. Examples include reading comprehension (Palinscar and Brown, 1986), writing (Scardemalia, Bereiter, and Steinbach, 1984), mathematics (Campione, Brown, and McConnell, 1988; Schoenfeld, 1991), physics (White and Frederickson, 1998; Reif and Scott, 1999), statistics (Lovett, 2001), and computer program debugging (Klahr and Carver, 1988). For example, in Palinscar and Brown's (1986) Reciprocal Teaching method used to support text comprehension, instruction is structured around encouraging students to implement four strategies to monitor their comprehension: summarizing, question generating, clarifying, and predicting. The teacher initially models these comprehension strategies, and then scaffolds students' use of the strategies by prompting them to summarize, predict, etc. Finally, students take turns assuming the role of teacher in leading this dialogue with each other. Although the specific activities in the various instructional programs are domain dependent, they all focus on the generalizable metacognitive procedures or features that address a wide range of problems within their domains, such as diagramming, question posing, and self-explanation, and are not tied to specific problem solution strategies.

It is clearly necessary for students to have a sound conceptual framework for solving statics problems. However, conceptual knowledge in itself is not sufficient for successful problem solving: students must also know when and how to apply their knowledge. Here we propose and test the effectiveness of a questioning technique in improving students' problem solving. The questions are designed to focus students' attention on features of statics problems that are closely tied to relevant conceptual knowledge, and provide a systematic approach to building a conceptual representation before they pursue mathematical solutions. Like the previous examples of programs, the approach is specific to a domain but applicable to a wide range of problems within the domain.

II. CONCEPTUAL FRAMEWORK FOR STATICS AND ORIGIN OF PROPOSED METACOGNITIVE STRATEGY

In several branches of engineering, including mechanical and civil engineering, statics forms an important foundation to subsequent courses, such as strength of materials and dynamics. In addition, to the extent that design activities draw upon engineering science knowledge, statics can play a key role in design. Indeed, instructors in design courses lament the inability of students to use knowledge from prior courses, such as statics, for practical design purposes (Harris and Jacobs, 1995). Several potential flaws in traditional statics instruction have been catalogued recently (Steif and Dollár, 2005), and alternative approaches that might better address conceptual challenges were recommended. In parallel, a conceptual framework for statics has been proposed (Steif, 2004), and this has led to the development of a now widely used Statics Concept Inventory (Steif and Hansen, 2007). In this inventory, three of the four concept clusters involve bodies and the relations between bodies and forces; these underlie the ability to reason about real systems. The centrality of bodies in the concepts of statics is one origin for the instructional strategy proposed below.

The second origin of the instructional strategy is the observation of the first author as a long time instructor in statics. When students ask for help in solving statics problems, certain body-centered questions posed by the instructor very often appear to provoke productive thought in the student. Such questions include: "Precisely what bodies from the original system are you including in your free body diagram?" or "Which body exerts the force that you have drawn on that free body diagram?" These questions appear to push the student to explicitly address and grapple with fundamental concepts and in effect, support a shift in students' mental model or representation of the problem toward one that places the bodies and the interactions between them at the forefront. These prompts can be viewed as metacognitive because they force students to monitor their process and explain their reasoning to themselves and the instructor in terms of bodies, thus shifting their underlying model of the problem (Hausmann et al., 2009). By focusing students' attention on the critical aspects of the problem, they were more likely to identify and self-correct errors and thus develop a more complete and relevant conceptual representation of the problem.

The apparent success of this questioning strategy of the instructor suggests that students may benefit if they learn to ask themselves similar questions. This study seeks to determine whether metacognitive prompts can induce students to recognize bodies that play important roles and whether this improves their problem solving performance.

III. METHOD

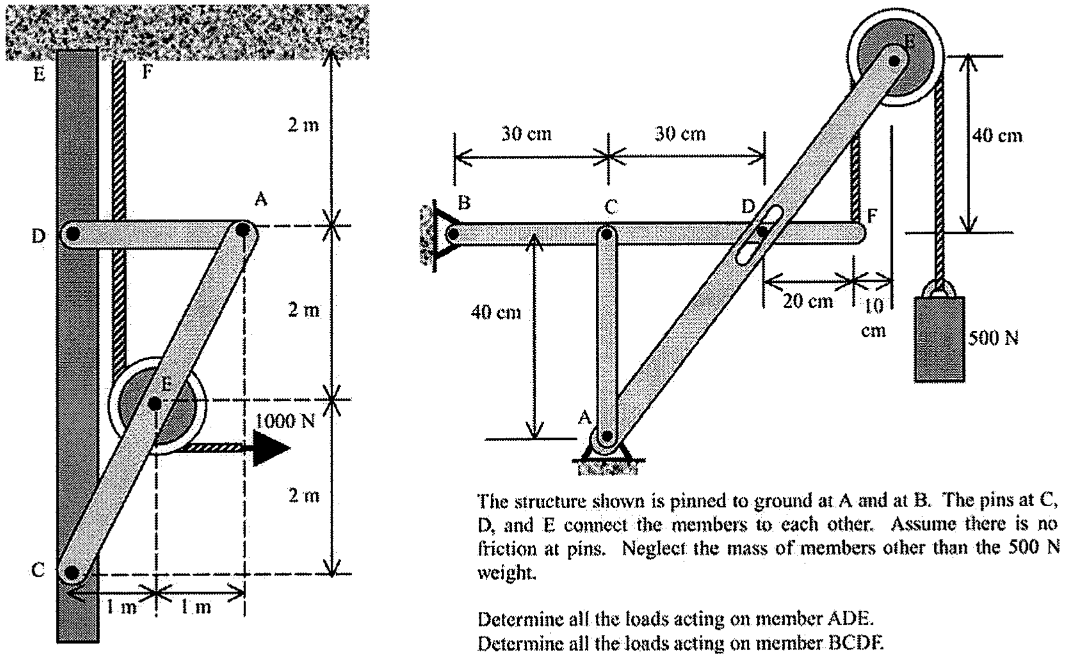
A. Participants

Twenty-one college students who previously completed and passed an introductory statics course participated in the study. Students were enrolled in either the mechanical or civil engineering program of the University of Pittsburgh or in the civil engineering program of Carnegie Mellon University (none were students of the authors). Students were randomly assigned to the experimental ($N = 10$) and control groups ($N = 11$), respectively.

In order to examine whether body-centered thinking is a more expert-like approach to solving statics problems, we also recruited two mechanical engineering graduate students. Both students had served as statics TAs for one semester for the first author two and four years prior to the study, respectively, but neither had been students of the author in a statics course. Graduate students were chosen because they represent an intermediate level of expertise, characterized by strong conceptual knowledge and proficient problem solving skills while still maintaining an explicit awareness of the individual steps in their solution process, the latter being a characteristic often lost at higher levels of expertise (Berry, 1987; Ericcson and Simon, 1984).

B. Pre-test and Post-test Problems

For all pre-test and post-test problems students were given a diagram and asked to determine the loads (interactions or forces) acting on various bodies. All problems involved multiple bodies connected in various ways and required many critical concepts in statics. The second and third pre-test problems are those shown in Figure 1. The post-test problems were constructed to be conceptually identical permutations of these problems so that performance on the post-test problems could be meaningfully compared with the second and third pre-test problems. No feedback was given to participants on the solutions to these problems.



The structure shown is pinned to ground at A and at B. The pins at C, D, and E connect the members to each other. Assume there is no friction at pins. Neglect the mass of members other than the 500 N weight.

Determine all the loads acting on member ADE.
Determine all the loads acting on member BCDF.

Figure 1. Typical problems used in study.

C. Instructional Sessions

Instruction for both groups centered on the analysis of three statics problems (comparable to those in Figure 1). The goal of the instruction was to improve students' ability to draw free body diagrams, and in particular to represent the unknown forces at various points, by having them analyze diagrams and determine whether the representation of forces was correct or incorrect. The control group was shown representations of forces for various subsets of the problem (for example, one part or two parts together). Some representations were correct and others were incorrect, and students were asked if the representation was correct via written instructions next to the diagram. Students answered orally, and when the students signaled readiness, the system would give feedback in the form of a recorded voice as to whether the representation was correct, and if incorrect, displayed correct representations.

Students in the experimental group were ultimately shown the same set of problems, subsets, and correct and incorrect representations, but their instruction included a set of intermediate questions that were designed to promote a more body-centered approach to statics problems. Specifically, after being shown the original problem diagram (e.g., in Figure 1) and prior to being shown any free body diagrams, students were asked via written instruction:

- Name the distinct parts.
- Where exactly do the parts contact each other; that is, on what surfaces do they touch?

Students answered orally and when ready heard a recorded response to the questions. Next, students in the experimental group were shown a subset of bodies that would form a free body diagram, without representations of forces, and were asked via written instruction:

- Name all of the parts that are included in the free body diagram(s).
- Name all of the external parts that directly contact the parts of the free body diagram(s).

Students answered orally and when ready heard a recorded response to the questions. Then using the same subset of bodies, students were shown a completed free body diagram with representations and were asked via written instruction:

- For each load drawn, name the part that exerts that load.
- Think about the forces that the parts can exert on one another: is each load drawn correctly?
- Is a contacting body exerting a force that is not shown on the diagram? If so, where and to what load would it apply?

Students answered orally and when ready heard a recorded response to the questions.

D. Procedure

Participation for the control and experimental groups consisted of two, individually run sessions, both approximately 1.5 to 2 hours in length. In the first session students were asked to solve three pre-test problems followed by an instructional unit. In the second session students began with a second, similar instructional unit, followed by two post-test problems. All students solved the identical set of problems. The first pre-test problem was intended to acclimate students to the system and was not used in the analysis.

While students worked on the pre-test and post-test problems, they were asked to talk-aloud about what they were thinking. Think-aloud protocols have been used extensively in problem solving research, and an extensive review of many studies found no evidence that performance or process changed as a function of the think aloud procedure (Ericsson and Simon, 1993). In addition, for the pre-test and post-test problems, all written work generated by the students (e.g., diagrams, equations, notations) was captured with a large digitizing tablet and cordless stylus. A computer program recorded the time of each pen stroke and participant's speech was recorded digitally and transcribed with time stamps, allowing written work and verbalizations to be played back and coded. Thus, for each student a large, multi-faceted, and dense set of data was produced that was transcribed, coded, and analyzed in multiple ways.

Only the data generated during the pre-test and post-test problems were coded and analyzed. To measure problem-solving performance, free body diagrams and associated equations of equilibrium for the pre- and post-instruction problems were graded for conceptual errors. In particular, we coded superfluous and missing forces, the correctness of representations of unknown forces, the presence or absence of equal and opposite pairs of forces, and the inclusion of all relevant equations of equilibrium. To assess the degree to which participants engaged in body-centered talk, each verbal protocol was transcribed and coded into several categories, as described in the Results section. Relative frequencies of different coding categories were computed. In addition, for key portions of the written solution, the verbal protocol was searched to determine whether body-centered talk specifically relevant to that portion was used.

The graduate student sessions were run individually in a single session. Both students solved the same two statics problems (depicted in Figure 1) and their verbal and written protocols were analyzed in the same way as those of the student participants. For the first problem, the instruction was: "Imagine you are solving this problem for yourself because you wanted to verify the answer given in the textbook. Simply talk aloud whatever you are thinking as you solve the problem." For the second analyzed problem, the instruction to the graduate student was: "As you are solving this problem, talk as if you were explaining your solution process to a student."

E. Data Coding and Reliability

To establish the reliability of the coding, all pre-test and post-test session protocols were coded independently by two individuals with expertise in statics, and the agreement between the raters was measured. There is no unique measure of inter-rater reliability; the task here is further complicated by the presence of multiple types of body-centered talk. The κ statistic was used to capture the agreement of the raters in categorizing each utterance of the protocol as body-centered talk or non body-centered talk. The κ statistic can be used (Landis and Koch, 1977) when each of two raters makes a binary decision (yes or no); κ utilizes the frequencies of each possible pair of answers (no-no, yes-no, no-yes, and yes-yes). For this study we calculated $\kappa = 0.77$, which is considered a significant degree of agreement. In addition, in 94 percent of cases where the raters agreed that an utterance should be coded as body-centered, raters also agreed on the sub-category of body-centered talk.

IV. RESULTS

A. Talk

While participants' talk was coded into several categories, for our analyses we focus on four types of talk:

1. Overall number of utterances
2. Identification of body isolated in free body diagram
3. Identification of body exerting a force
4. Identification of equilibrium condition imposed

Category 1 provides a baseline measure of the overall amount of talk. The length of an utterance was determined by the location of pauses, which were indicated by the timestamps created by the capture software. Categories 2 and 3 represent two types of body-centered talk, and category 4 is a type of talk that is likely common during statics problem solving, but is not body-centered talk. The following examples of Talk 2 through 4, extracted from the verbal protocols of the student participants, are typical:

> Talk 2:
>
> "Ok, so drawing a free body diagram of member EDA"
>
> Talk 3:
>
> "There is a pin at D, so that is D_x and D_y"
>
> Talk 4
>
> "Sum of forces in the y direction is T plus D_y, plus C_y, plus B_y equals zero."

We monitor the first type to determine the overall amount of talk and changes with instruction. The second and third types of talk are those that instruction was intended to promote, and the last type was completely unaddressed in the instruction (which did not proceed to the solution phase where equilibrium is imposed).

B. Student Talk

There was no significant difference between the two groups of students in terms of overall number of utterances (Talk 1) before or after instruction, $F(1,19) = 2.7, p > 0.1$ and no significant change in number of utterances from pre-test to post-test, $F(1,19) = 0.55$, $p > 0.1$. Thus the two groups were similar in their overall tendency to verbalize, and this did not change significantly over time.

Talk 2 measures the frequency of identifying the body drawn in a free body diagram, normalized by the number of free body diagrams drawn. Identifying the body in a free body diagram is emphasized in instruction for the experimental group, but not in the control group. Nonetheless, as Table 1 indicates, even prior to instruction, students have a strong tendency to identify the body in any free body diagram and this tendency did not change significantly from pre-test to post-test overall, $F(1,19) = 0.48, p > 0.4$. It may be that instructors normally demonstrate problem solving by identifying the bodies depicted in the free body diagrams, and indeed, we found that the graduate students named virtually all the bodies in their problem solving session.

On the other hand, it may be that instructors do not tend to emphasize the naming of bodies exerting a given force. Naming the exerting body was one goal of the instruction offered to the experimental group and was measured in Talk 3. Analyses revealed a significant interaction effect, $F(1,19) = 6.0, p < 0.03$, with both groups rarely naming the exerting body before instruction and only the experimental group significantly increasing its naming of the exerting body after instruction (Figure 2). The frequency of articulating the imposition of equilibrium (Talk 4) did not differ between the groups, $F(1,19) = 1.37, p > 0.2$, and there was no significant change in the tendency to articulate from pre-test to post-test, $F(1,19) = 2.10, p > 0.1$. This was expected since neither group received instruction pertaining to the phase of imposing equilibrium.

Group	Pre-Instruction	Post-Instruction
Control	81% (9.1)	80% (7.1)
Experimental	73% (9.9)	82% (7.1)

Table 1. Mean percentage (with Standard Error) of free body diagrams in which bodies are identified (Talk 2).

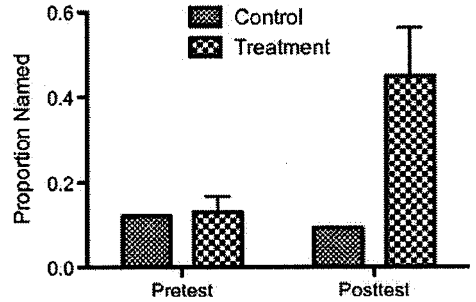


Figure 2. Proportion of forces for which the exerting body is named (Talk 3).

Subject	GS Directed to Explain to Self (Total Utterances)	GS Directed to Explain to Student (Total Utterances)
GS 1	0.56 (217)	1.00 (221)
GS 2	0.20 (68)	0.67 (123)

Table 2. Fraction of interactions in which the body is named (Talk 3) and total utterances for each of two problems solved by graduate student subjects.

C. Graduate Student Body-centered Talk

To further verify the importance of a body-centered approach to solving static problems, verbal and written protocols of the two graduate students were collected for two problems. Both graduate students named virtually all the bodies that they drew for free body diagrams (Talk 2). Since the greatest change in students that we seek to induce is Talk 3, naming the body exerting a force, this feature of the graduate student protocols was of most interest. It can be seen (Table 2) that graduate student GS 1 generated more talk and had a higher proportion of named interactions than GS 2. In the first problem, where they were asked to solve the problem for themselves', both students were more likely to name interactions than the experimental and control groups on the pre-test problems and the control group on the post-test problems. When they were asked to solve the problem as if explaining the process to a student, the fraction of named interactions increased markedly for both graduate students. The high proportion of interactions named by the graduate students when explaining to themselves and to students suggests that they recognized the importance of focusing on this aspect of the problem in order to solve it.

D. Performance

While performance on various aspects of the solution was monitored, here we report on two Performance Measures:

1. Correctness of interaction representations
2. Identification of Newton's 3rd Law interaction pairs

Instruction was expected to affect some aspects of performance but not others. In particular, our expectation was that Talk 3, concerning which bodies exert forces, would affect Performance Measures 1 and 2. Performance Measure 1 is expected to improve with instruction because the representation is tied to the nature of the connection between the bodies, which is strongly influenced by the two bodies. Performance Measure 2 is also expected to improve with the experimental instruction. For example, consider the situation where we name the body that exerts a force: if we later draw a free body diagram of the exerting body, its role as one half of a Newton 3rd Law pair is already acknowledged. But, if we merely note that a force is exerted at a point, with the second (exerting) body of the pair left unspecified, the force on the second body is more likely to be overlooked later. Other aspects of statics problem solving, such as the imposition of the equations of equilibrium, which are not addressed by the instruction tested, were, as expected similar for the two groups and not affected by instruction.

The main effect for instruction on Performance Measure 1 was significant, $F(1,19) = 14.25, p = 0.0013$ (Figure 3). Prior to instruction, both groups correctly identified only 65 percent of the interactions. After instruction, the control group correctly identified 76 percent and the experimental group correctly identified 90 percent. This is not surprising: instruction included the identical (large) set of examples of free body diagrams and representations; representations that were incorrect were corrected in the same way for both groups. Instruction for the experimental group differed in that the exerting bodies were named, and the relation between the bodies, their connection, and the representation was articulated. Although the experimental group did show a greater improvement than the control group (pre-post difference = 0.25 vs. 0.09, respectively) this difference was not significant, $F(1,19) = 3.03, p = 0.098$.

An alternative means of capturing the effect on Performance Measure 1 is found if we note that the majority of interactions were represented properly by most subjects (all of whom had passed statics), and that most of the interactions were not named. Clearly, failing to name the exerting body does not necessarily mean that a student has not identified it, nor does failing to name it guarantee an error. However, explicitly naming the exerting body is an indication that the body has been identified; so naming the exerting body should reduce the likelihood of error. To quantify this effect, we considered all errors, and determined how many were associated with named and unnamed bodies. So for each student we computed two fractions: (i) fraction of errors when naming the body, and (ii) fraction of errors when not naming the body. Figure 4 shows means for these two fractions, computed over students in each of the two groups, both before and after instruction. As shown, the fraction of errors committed having named the body is very small, whereas the fraction of errors committed having not named the body is considerably higher. Furthermore, this difference was significant for both experimental and control groups before ($F(1,19) = 170, p < 0.0001$) and after ($F(1,19) = 27, p < 0.0001$) instruction. This supports the idea that the practice of explicitly identifying the exerting body, which the instruction is intended to foster, reduces errors. Furthermore, the relationship between naming and errors regardless of group suggests that some students develop this strategy

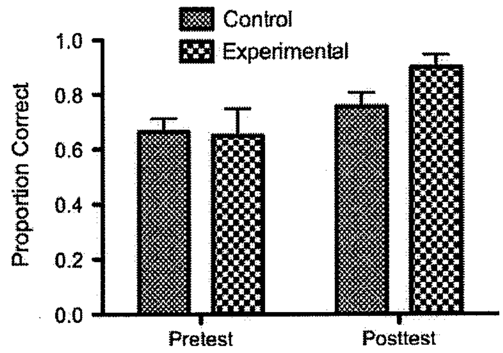


Figure 3. Proportion of correct interactions represented.

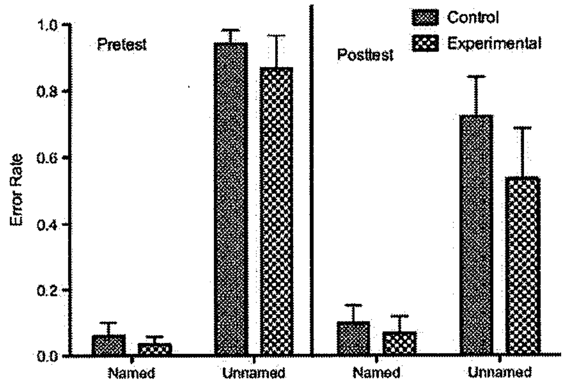


Figure 4. Error rates as a function of naming and not naming the interactions.

without instruction but others do not and thus may benefit from explicit instruction of this kind.

The effect of experimental instruction on Performance Measure 2, identifying equal and opposite interaction pairs, was less strong. In this case, we counted the number of opportunities to draw equal and opposite pairs, and then determined the fraction identified and drawn. (This method does not include the instances in which pairs of forces were falsely identified as equal and opposite). The two groups identified nearly the same fraction before instruction (0.39 vs. 0.44 for control and experimental, respectively), and both groups improved significantly after instruction, $F(1,16) = 5.6, p <= 0.03$. Although the experimental group showed a greater improvement (increase of 0.05 vs. 0.20 for control and experimental, respectively), this difference was not significant, $F(1,16) = 1.1, p > 0.3$.

V. SUMMARY AND CONCLUSIONS

Solving of problems in statics, as in other engineering courses, requires students to draw upon relevant conceptual knowledge and apply it appropriately. Accordingly, students must recognize salient elements that relate to relevant concepts in the problem statement and diagram. It is hypothesized that salient elements in statics are the distinct bodies in the diagrams, and their contacts with one another that are represented by forces. This study examines the hypotheses that a questioning strategy that focuses attention on and thinking about the bodies present in a problem can improve problem solving performance.

The study employed a pre-post design with experimental and control groups. Both groups solved the same set of problems before and after instruction, and both groups received instruction based on the same set of problems and partially correct solutions. Instruction differed in that the experimental group was prompted to comment on the bodies in the problem and the correctness of the various free body diagrams, whereas the control group only commented on the latter. Students' written work while problem solving was captured on a computer tablet, while the think-aloud protocol was captured by digital recording; the development of the solution and verbal protocol could then be played back in synchrony and analyzed. While the sample size may be considered small compared to studies that only use performance measures, it is typical for studies that use quantitative data, such as verbal protocols, that are time intensive to collect and analyze. Furthermore, these types of data (real-time pen-stroke analysis, verbal protocols and accuracy), provide a rich picture of both the process and outcome of problem solving performance. Since a major focus of the study was to examine how students' thinking affected their solution strategies, these measures were the most appropriate data to collect.

It was found that the amount of body centered talk was comparable for the two groups before instruction, and increased significantly only for the experimental group. For one prime measure of performance, representation of unknown forces on free body diagrams, it was found that both groups improved; the experimental group improved more although the difference was short of significant. Since a large fraction of the forces are represented correctly, we focused on those which were represented incorrectly. Although the control group displayed less body centered talk, both groups had a substantially lower frequency of errors when the relevant body was discussed. This result points to the specific benefit of body centered thinking, as measured by the amount of body-centered talk, regardless of instruction. This is consistent with research on self-explanation, where students who generate more explicit and deep explanations of the process perform better than students who generate few or shallow explanations (Chi et al., 1994). These results suggest that student performance in statics may benefit from instruction that promotes a more systematic discussion of bodies and that the use of metacognitive prompts may help students develop a body-centered representation and help them monitor their problem solving process. Perhaps studies of conceptual knowledge in other engineering subjects might be expanded to identify common, salient features in problems which prompt thinking about conceptual knowledge.

ACKNOWLEDGMENT

The authors thank Stephen Spencer for his assistance in initially setting up the experiments and Victoria Pagos for her assistance in conducting some of the experiments. Support by the National Science Foundation under grant REC-0440295 and by the Department of Mechanical Engineering at Carnegie Mellon University is gratefully acknowledged.

REFERENCES

Berry, D.C. 1987. The problem of implicit knowledge. *Expert Systems* 4 (3): 144–51.

Brown, A.L. 1980. Metacognitive development and reading. In *Theoretical issues in reading comprehension: Perspecitves from cognitive psychology, linguistics, artificial intelligence and education*, eds. R.J. Spiro, B.C. Bruce, and W.F. Brewer. Hillsdale, NJ: Lawrence Erlbaum Associates.

Campione, J.C., A.L. Brown, and M.L. McConnell. 1988. Metacognition: On the importance of understanding what you are doing. In *Research agenda for mathematics education: The teaching and assessing of mathematics problem solving*, eds. R.I. Chales and E.A. Silver, 93–114. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M.T.H, N. DeLeeuw, M. Chiu, and C. Lavancher. 1994. Eliciting self-explanations improves understanding. *Cognitive Science* 18 (3): 439–77.

Chi, M.T.H., P. Feltovich, and R. Glaser. 1981. Categorization and representation of physics problems by experts and novices. *Cognitive Science* 5 (2): 121–52.

Ericsson, K.A., and H.A. Simon. 1984. *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Flavell, J.H. 1985. *Cognitive development.* Englewood Cliffs, NJ: Prentice-Hall.

Flavell, J.H. 1991. Understanding memory access. In *Cognition and the symbolic processes: Applied and ecological perspectives*, eds. R. Hoffman and D. Palermo. Hillsdale, NJ: Lawrence Erlbaum Associates.

Harris, T.A., and H.R. Jacobs. 1995. On effective methods to teach mechanical design. *Journal of Engineering Education* 84 (4): 343–49.

Hausmann, R.G.M., T.J. Nokes, K. VanLehn, and S. Gershman. 2009. The design of self-explanation prompts: The fit hypothesis. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society.* Amsterdam, Netherlands.

Hestenes, D., M. Wells, and G. Swackhamer. 1992. Force concept inventory. *The Physics Teacher* 30: 141–58.

Klahr, D., and S.M. Carver. 1988. Cognitive objectives in a LOGO debugging curriculum: Instruction, learning and transfer. *Cognitive Psychology* 20 (3): 362–404.

Landis, R.J., and G.G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1): 159–74.

Larkin, J.H., J. McDermott, D.P. Simon, and H.A. Simon. 1980. Experts and novice performance in solving physics problems. *Science* 208 (4450): 1335–42.

Lovett, M.C. 2001. A Collaborative convergence on studying reasoning processes: A case study in statistics. In *Cognition and instruction: Twenty-five years of progress*, eds. S. Carver and D. Klahr, 347–384. Mahwah, NJ: Lawrence Erlbaum Associates.

Palinscar, A.S., and A.L. Brown. 1986. Interactive teaching to promote independent learning from text. *Reading Teacher* 39 (8): 771–77.

Reif, F., and L.A. Scott. 1999. Teaching scientific thinking skills: Students and computers coaching each other. *American Journal of Physics* 67 (9): 819–31.

Scardemalia, M., C. Bereiter, and R. Steinbach. 1984. Teachability of reflective processes in written composition. *Cognitive Science* 8 (2): 173–90.

Schoenfeld, A.H. 1991. On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In *Informal reasoning and education*, eds. J.F. Voss, D.N. Perkins, and J.W. Segal. Hillsdale, NJ: Lawrence Erlbaum Associates.

Steif, P.S. 2004. An articulation of the concepts and skills which underlie engineering statics. In *Proceedings of the 34rd ASEE/IEEE Frontiers in Education Conference.* Savannah, GA.

Steif, P.S., and A. Dollár. 2005. Reinventing the teaching of statics. *International Journal of Engineering Education* 21 (4): 723–29.

Steif, P.S., and M.A. Hansen. 2007. New practices for administering and analyzing the results of concept inventories. *Journal of Engineering Education* 96 (3): 205–12.

Voss, J.F., T.R. Greene, T.A. Post, and B.C. Penner. 1984. Problem solving skills in the social sciences. In *The psychology of learning and motivation: Advances in research theory*, ed. G.H. Bower, 17. New York: Academic Press.

White, B.Y., and J.R. Frederickson. 1998. Inquiry, modeling and metacognition: Making science accessible to all students of reflective processes in written composition. *Cognition and Instruction* 16 (1): 3–117.

Wineburg, S.S., 1991. Historical problem solving: A study of the cognitive processes used in evaluation of documentary and pictorial evidence. *Journal of Educational Psychology* 83 (1): 73–87.

AUTHORS' BIOGRAPHIES

Paul S. Steif is professor of Mechanical Engineering at Carnegie Mellon University. He received a Sc.B. in engineering from Brown University (1979) and M.S. (1980) and Ph.D. (1982) degrees from Harvard University in Applied Mechanics. He has been active as a teacher and researcher in the field of engineering mechanics. In particular, Dr. Steif develops and implements new approaches and technologies to measure student understanding of engineering and to improve instruction.

Address: Department of Mechanical Engineering, Carnegie Mellon University, Schenley Park, Pittsburgh, PA, 15213; telephone: (+1) 412.268.3507; fax: (+1) 412.268.3348; e-mail: steif@cmu.edu.

Jamie M. LoBue is a graduate of Carnegie Mellon University (2008) with a degree in Mechanical Engineering. During her undergraduate studies, she conducted research with Professor Paul S. Steif into student learning of engineering concepts. She currently lives in Southern California and works at an aerospace company.

Address: 4981 Pearce Street, Unit B, Huntington Beach, CA, 92689; telephone: (+1) 914.318.2381; e-mail: jlobue1231@gmail.com.

Anne L. Fay is an independent educational consultant, specializing in assessment and evaluation. She received a B.A. from York University (1983) and a Ph.D. (1990) from the University of California, Santa Barbara in cognitive psychology. She has conducted research on reasoning and problem solving and has conducted educational assessments and evaluations for such groups as the National Science Foundation, Society of Automotive Engineers, The Educational Program of the Supercomputing Conference and The Children's Television Workshop.

Address: 718 Maryland Avenue, Pittsburgh, PA 15232; telephone: (+1) 412.661.7186; email: annelfay@me.com.

Levent Burak Kara is an assistant professor in the Department of Mechanical Engineering at Carnegie Mellon University. His research interests include intelligent design tools, pen computing, geometric modeling, human-computer interaction, and artificial intelligence. Kara has a BS in mechanical engineering from the Middle East Technical University, and an M.S. and Ph.D. in mechanical engineering from Carnegie Mellon University.

Address: Department of Mechanical Engineering, Carnegie Mellon University, Schenley Park, Pittsburgh, PA, 15213; telephone: (+1) 412.268.2509; fax: (+1) 412.268.3348; e-mail: lkara@cmu.edu.