AC 2007-368: INDUCING STUDENTS TO CONTEMPLATE CONCEPT-ELICITING QUESTIONS AND THE EFFECT ON PROBLEM SOLVING PERFORMANCE

Paul Steif, Carnegie Mellon University

PAUL S. STEIF Professor, Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, Pa Degrees: Sc. B. 1979, Brown University; M.S. 1980, Ph.D. 1982, Harvard University. Research area: engineering mechanics and education.

Jamie LoBue, Carnegie Mellon University

Undergraduate Student, Mechanical Engineering

Anne Fay, Carnegie Mellon University

Director of Assessment, Eberly Center for Teaching Excellence, Carnegie Mellon University, Pittsburgh, PA Degrees: B.A. 1983, York University; Ph. D. 1990, University of California, Santa Barbara. Research area: Reasoning and problem solving; applications of cognitive psychology to educational practice

Burak Kara, Carnegie Mellon University

Post-doctoral research associate, Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA Degrees: B.S. 1998, Middle East Technical University; M.S. 2000, Ph.D. 2004, Carnegie Mellon University. Research area: engineering design, computational geometry, sketch-based interfaces.

Steve Spencer, Carnegie Mellon University

Undergraduate student in Departments of Psychology and Industrial Design

Inducing Students to Contemplate Concept-Eliciting Questions and the Effect on Problem Solving Performance

Introduction

In many engineering subjects students learn to solve problems. Problem solving demands the transfer of knowledge from one context to another[1]. This requires that one's knowledge be suitably organized in meaningful patterns, and that one be able to retrieve that knowledge, recognizing its relevance in the context of the problem solving process. This is linked to one widely appreciated dimension of expertise: metacognition or the ability to monitor one's progress in approaching a task and to determine when understanding is inadequate[2-4].

A number of researchers have successfully developed and implemented programs to support students' metacognitive skills to improve learning and problem solving. Examples include reading comprehension[5], writing[6], mathematics[7-8], physics[9-10], statistics[11] and computer program debugging[12]. For example, in Brown & Palinscar's Reciprocal Teaching method, which is used to support text comprehension[4], instruction is structured around encouraging students to implement four strategies: summarizing, question generating, clarifying, and predicting. The teacher initially models these comprehension strategies, asking students to summarize, predict, etc., and then students take turns assuming the role of teacher in leading this dialogue with each other. Although these instructional programs are domain dependent, each focuses on procedures or features that are generally applicable to a wide range of problems within the domain, rather than specific problem solution algorithms. This paper investigates a domain-specific metacognitive strategy that may broadly benefit problem-solving in statics.

Conceptual Framework for Statics and Origin of Metacognitive Strategy

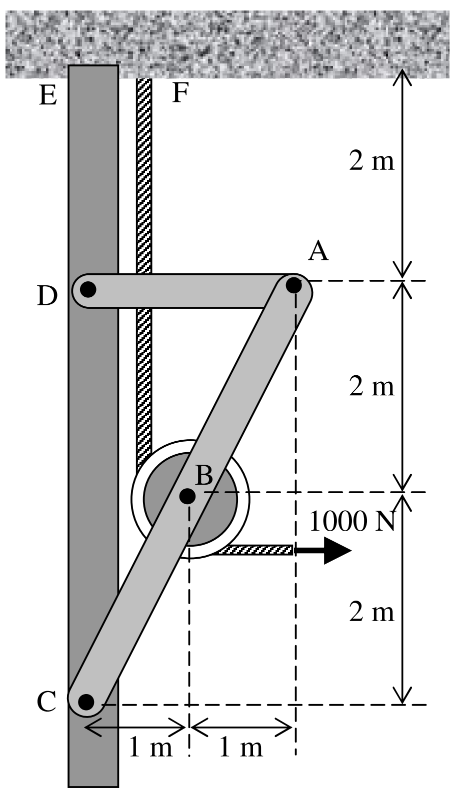
In several branches of engineering, including mechanical and civil engineering, statics forms an important foundation to subsequent courses, such as strength of materials and dynamics. In addition, to the extent that design activities draw upon engineering science knowledge, statics can play a key role 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[13]. Several potential flaws in traditional statics instruction have been catalogued recently[14]. It was argued that students need to learn statics in the context of physical artifacts, and that the concepts of statics need to be presented so they build systematically upon each other. A conceptual framework for statics has been proposed[15], and this has led to the development of a now widely used Statics Concept Inventory[16-17]. Three of the four concept clusters involve bodies and the relations between bodies and forces. This point is quite critical – students and instructors often treat statics as largely an exercise in vectors (long ago statics was taught by mathematicians). This mathematical, rather than physical, approach impedes students in ultimately applying statics to real systems. The centrality of bodies in the concepts of statics is one origin for the metacognitive strategy proposed below.

The second origin is the observation of the first author as a long time instructor in statics. When students come for help in solving statics problems, certain 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 grapple directly with fundamental concepts.

The apparent success of this questioning strategy of the instructor as tutor suggests that students may benefit if they learned to ask themselves similar questions. This present study seeks to determine whether students can be induced to contemplate statics problems with heightened regard for the bodies present and whether this improves their problem solving performance.

Research Design and Methods

The investigation uses a pre-post design that monitors participants solving statics problems, both before and after instruction. The experimental group receives instruction featuring questions that promote a more body-centered approach to statics problems; the control group receives instruction with the same examples as the experimental group, but without the questions that induce body-centered thinking. Problems used for this study include those shown in Figure 1. The participant is asked to determine the loads (interactions or forces) acting on various bodies. All problems involve multiple bodies connected in various ways and require many critical concepts in statics.



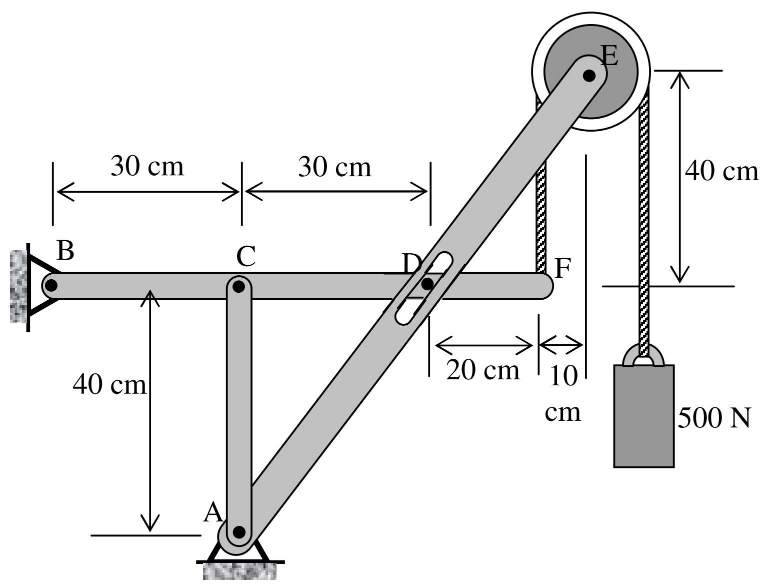


Figure 1. Examples of problems used in study.

Each participant first completes the Statics Concept Inventory on-line, and then attends two sessions approximately one and two weeks later, respectively. Each session lasts from 1.5 to 2.5 hours and involves a single participant. In the first session, a participant solves three such problems (designated problems A, B, and C) and receives one dose of instruction (designated problem D). In the second session, the participant receives a second dose of instruction (designated problems E and F), and then solves two more problems (designated G and H). The two problems of the second session (G and H) are conceptually identically to problems B and C of the first session, which are shown in Figure 1. Subsequent analysis is based on comparisons between these pairs of problems, not problem A of the first session.

While solving problems, participants are asked to think-aloud. The written solutions are captured with a large digitizing tablet and cordless stylus; a computer program records the time of each pen stroke. The participant's speech is recorded digitally and transcribed with time stamps; this allows the written solution and words to be played back in synchrony. Figure 2 shows a played back solution with the protocol in the text window at the right.

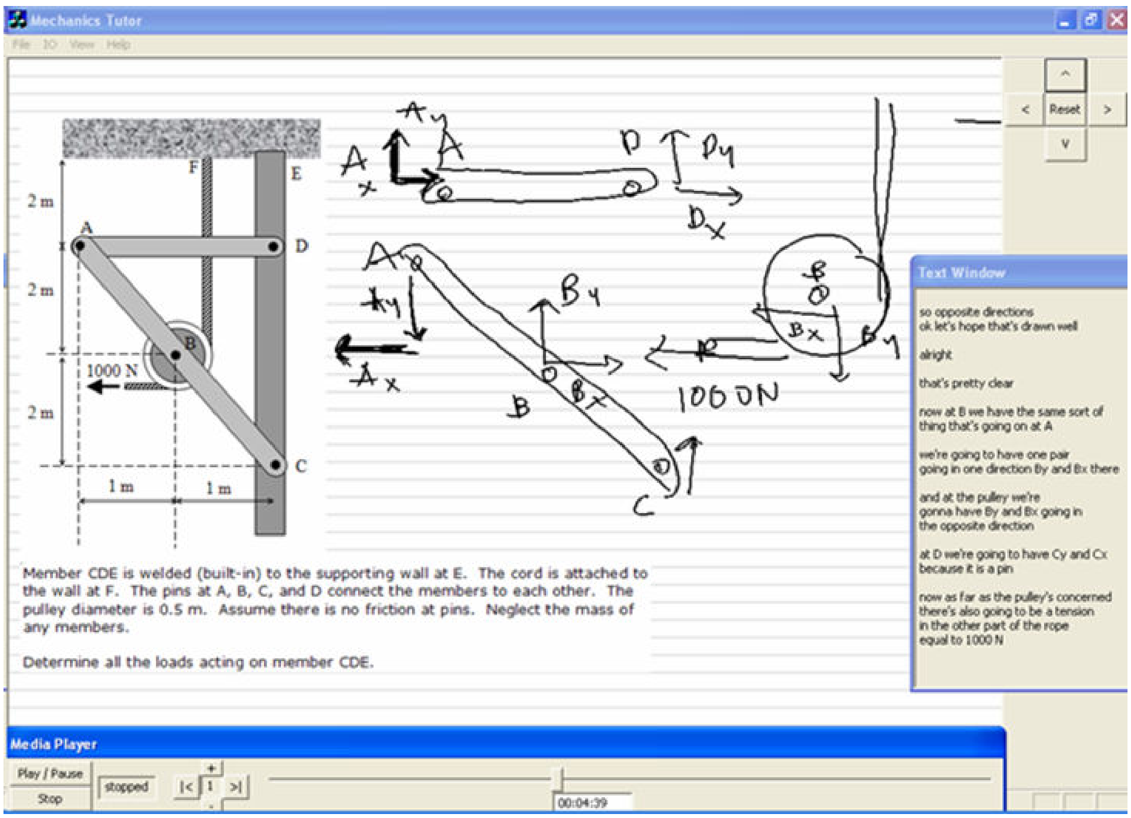


Figure 2. Solution and transcript of the verbal protocol played back.

Instruction is offered on the computer through a series of Flash Movies which are controlled by the participant; before each movie, a question appears on the monitor which the student is supposed to answer. More specifically, the participant is shown a problem similar to those solved, as well as a series of candidate free body diagrams associated with the problem displayed; the diagrams contain correct and incorrect elements. Participants in the control group are asked whether each free body diagram is correct. When the participant so requests, a flash movie is played in which correct and incorrect portions of free body diagrams are identified. Participants in the experimental group see the same sequence of free body diagrams, but, in addition, are asked questions relating bodies and forces. Questions involve naming parts that contact a given body, naming the part that exerts a drawn force, and determining whether the unknown drawn forces are consistent with the exerting bodies. A flash movie is then played in which an expert answers the questions posed. Both groups see the corrected diagrams and receive instruction aurally. Thus, the differing conditions seek to test whether thinking about bodies and forces offers benefits beyond those of merely seeing correct and incorrect examples.

Participants

Participants were solicited by email announcement to undergraduates in the Department of Mechanical Engineering and the Department of Civil and Environment Engineering at the University of Pittsburgh, and in the Department of Civil and Environment Engineering and the Department of Architecture at Carnegie Mellon University. In addition, participants were solicited by posted notices at Carnegie Mellon University. (In no instance, had participants taken statics from the first author, the only statics instructor among the co-authors.) Note that it would be infeasible to get sufficient participants who were at the same stage in the middle of a statics course, and infeasible to run the experiments in a short enough period of time to maintain students at the same stage of learning. Therefore, we opted for participants who had completed a statics course. Since statics courses are fairly standardized in terms of content, we could presume all participants had been exposed to roughly the same material (although they may have had different amounts of additional exposure to mechanics through dynamics or strength of materials courses). If the metacognitive strategy were found to benefit students who had completed statics, it would be fair to assume that its inclusion within statics instruction would be beneficial. Participants were paid $50 for completing the study; their participation did not involve any academic credit.

Analysis of Data

To gauge problem solving performance, written solutions to the pre- and post-instruction problems were graded for conceptual errors. In particular, we noted superfluous and missing forces, representations of unknown forces, and the presence or absence of equal and opposite pairs of forces. To assess the degree to which participants engaged in body-centered talk, the verbal protocol was coded into categories: (1) body-centered talk, (2) general mechanics reasoning, (4) other metacognitive statements, (4) mathematical reasoning, and (5) restatements or paraphrasing of the problem statement. Body-centered talk was broken down into 7 sub-categories:

1.1. Naming parts in system
1.2. Identifying a subsystem to focus on
1.3. Describing relevant feature of body
1.4. Identifying a body that contacts the subsystem of interest
1.5. Ascribing a force at some location to a contacting body
1.6. Representing an unknown force, for example indicating it is in particular direction or has x- and y-components
1.7. Explicit mention of an equilibrium condition

For each unknown interaction between bodies, the verbal protocol was searched to determine whether there was discussion of the relevant bodies. For the comparisons below between conceptual errors found in the free body diagrams and the participants' talk, we counted sub-categories 1.1 through 1.5 that explicitly involve reference to bodies. (Categories 1.6 and 1.7 were tracked for future studies.)

Reliability

To establish the reliability of the coding, the 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 sub-categories of body centered talk. The κ statistic was used to capture the agreement of the raters in placing each utterance of the protocol into category 1 versus another category[18]. The κ statistic can be used 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). In addition, for all cases where the raters agreed that an utterance merited a coding of 1 (body-centered), the frequency at which there was agreement on the sub-category, 1.1, 1.2, and so forth, was captured.

Results

Preliminary experiments had been conducted prior to completing the flash instruction modules, and difficulties in recording left the protocols of several additional subjects unintelligible. Here we show results for the subsequent 9 participants for which analysis has been completed. Since the flash instruction modules were completed for the experimental group first, more participants (7) were run under the experimental condition as compared to the control condition (2). Once the numbers are approximately equal, subsequent subjects will be assigned randomly. The experimental group consisted of 5 males and 2 females; the control group consisted of 1 male and 1 female. The average Statics Concept Inventory score was 39% for the experimental group and 19% for the control group. While the Statics Concept Inventory has been shown to correlate well with other measures of statics performance, such as exam scores in a given statics class[17], the scores for this small group did not predict performance on the problem solved. As an alternative measure of the initial state of the two groups with respect to this set of problems, the average fraction of forces that were represented correctly in the pre-test problems was 74% for the experimental group and 82% for the control group.

For the full set of utterances for all problems of all subjects, the codings of raters were as shown in Table 1, with a reliability of κ =.772, which is considered to be significant[18]. In 94% of the 415 agreed upon instances of body centered statements, the raters agreed on the sub-category (1.1 through 1.7). After rating protocols independently, each utterance is given an agreed upon coding based on discussion between the two raters. The agreed upon coding is used for subsequent analysis.

Table 1. Codings of raters and inter-rater reliability κ.

Agreed body centered statement	Rater 1: body-centered; Rater 2: not	Rater 2: body-centered; Rater 1: not	Agreed not body centered statement	Reliability κ
415	124	68	2289	0.772

Since the analysis has been completed for relatively few subjects, statistically sound comparisons between the experimental and control group cannot yet be made. However, graphical display of the results for all subjects offers a preliminary indication of trends in the data. In Figure 3, we show the fractional increase in body centered talk in problems G and H relative to problems B and C) in sub-categories 1.1 to 1.5 that involve explicit mention of bodies. The average number of body centered statements per problem before instruction was comparable for the two groups: 3.9 for the experimental group and 3.5 for the control group. For example, subject 1 had an average of 1.5 body centered statements on the first two problems, and an average of 10 body centered statements on the last two problems. For the control subjects (8 and 9), the change in body centered talk was relatively small; one increased and the other decreased. In the experimental subjects (1 to 7, two had relatively small increases, and the remaining five had more substantial increases. This suggests that the instruction in body centered talk did induce students to engage in such talk.

Figures 4 and 5 capture the changes in two measures of performance. For each problem we consider all the unknown interactions that should be represented and that the student tries to represent, and we determine the fraction that is correct. The difference between the average fraction correct in the post-instruction problems G and H and the average fraction correct in the pre-instruction problems B and C is shown for each subject in Figure 4. In this respect, differences between the experimental and control groups cannot yet be discerned. We also determined the number of instances in which the equal and opposite interactions were correctly recognized in each problem; the increase in the absolute number of such pairs is shown in Figure 5. There appear to be greater improvements in the case of the experimental subjects. Quantification of the difference and its statistical significance will be carried out once sufficient numbers of subjects have been analyzed.

The results presented thus far tracked total change in relevant body-centered talk (categories 1.1 to 1.5) and the overall improvement in various measures of performance. Also of interest is whether body centered talk regarding any specific interaction is correlated with an increasing likelihood of that individual interaction being represented correctly. Further, we sought to determine whether this correlation held for participants in both experimental and control groups. Then, differences between the groups could be attributed to the difference in the amount of body centered talk. This was tracked by considering each unknown interaction to be represented and noting the presence of an utterance in either sub-categories 1.4 or 1.5 pertaining directly to that interaction. An utterance in sub-category 1.4 asserts that a particular body contacts the body of interest (the body of the free body diagram). An utterance in sub-category 1.5 asserts that a particular body was responsible for the force acting on the body of interest.

In Table 2 we show the numbers of interactions correctly and incorrectly represented. In addition, we show the number of the correctly and incorrectly represented interactions for which the exerting body was explicitly named via an utterance in sub-categories 1.4 or 1.5. It can be seen that the rate at which errors of representation are made drops from 18% to 2% when bodies are named. There is not quite as dramatic a drop in the rate of errors involving equal and opposite forces; when one or two of the two interacting bodies are named, the error rate is 21% compared to a rate of 43% when no bodies are named.

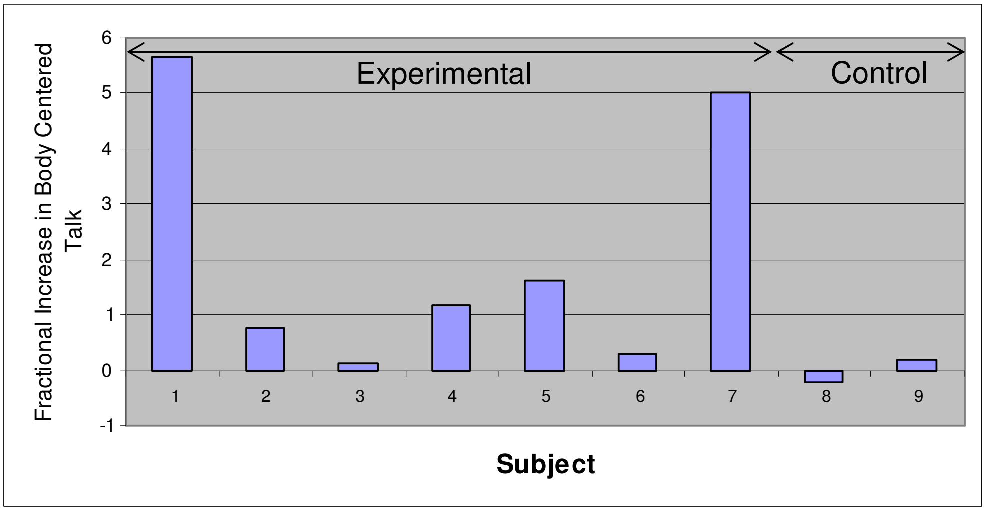


Figure 3. Pre- to post instruction increase in body-centered talk.

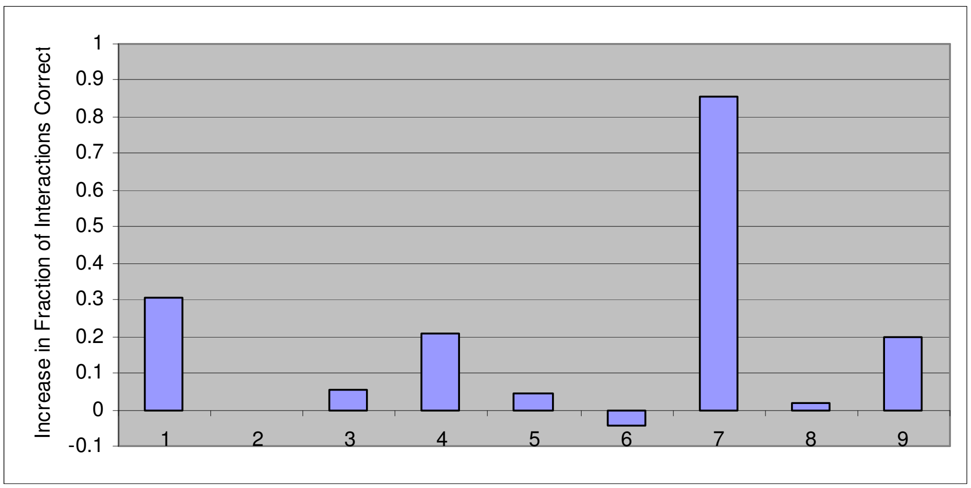


Figure 4. Pre- to post instruction increase in fraction of interactions represented correctly

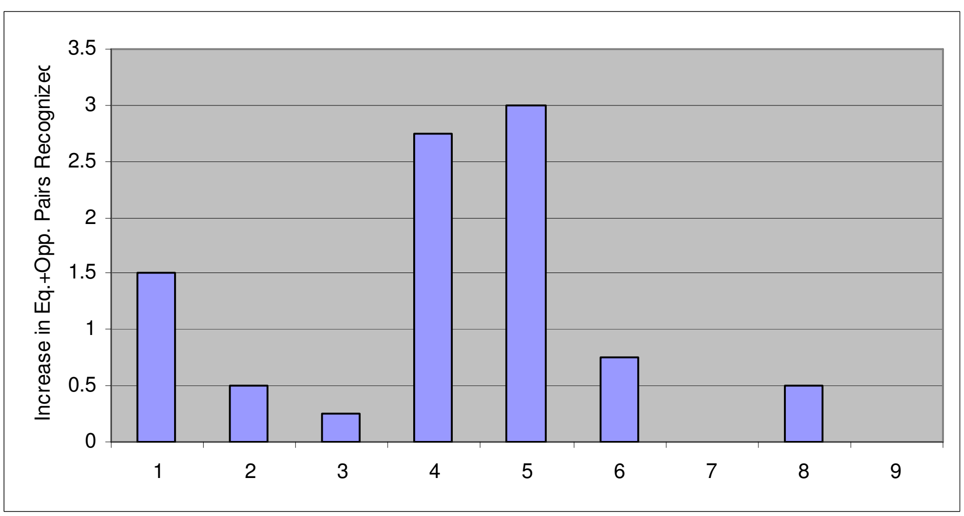


Figure 5. Pre- to post instruction increase in number of equal and opposite pairs recognized.

Table 2. Error rates for two types of errors and the effect of naming the body associated with the interactions.

	Number Correct	Number Correct (Body Named)	Number Wrong	Number Wrong (Body Named)	Error Rate (Body Unnamed)	Error Rate (Body Named)
Representing Interactions	260.5	65	45	1	18%	2%
Acknowledging Equal & Opposite	62.5	24.5	35	6.5	43%	21%

Summary

Problem solving in an engineering subject such as statics demands many skills, including for example, conceptual knowledge and analytical skills. It has been recognized in other domains that metacognitive skills are also required: the ability to monitor one's progress in solving problem. Based on the conceptual structure of statics, and on the experience of tutoring students, we propose an approach to help students guide themselves through the solution of multi-body statics problem. The approach focuses on the bodies in the problem, their contact with each other, and their relation to forces.

The proposed approach is investigated experimentally. Participants solve problems, receive instruction, and then solve additional problems. Written solutions are captured by stylus and digitizing tablet, and think aloud protocols via audio recording. By virtue of the transcription of the verbal protocol including time stamps, the solution and protocol can be replayed in synchrony. Instruction for the experimental group includes questions that the participant is to answer regarding bodies in the problem and their relation to forces; after responding to the questions the participant hears an expert responding to the same questions. All the examples used for the experimental group are also shown to the control group, but without the discussion of bodies. Solutions are graded via a rubric that focuses on conceptual errors; the verbal protocol is coded and body centered talk is identified.

Analysis has been completed for 7 experimental and 2 control subjects. While the results cannot be analyzed statistically due to the small numbers, the results suggest that the instruction can induce an increase in the amount of body-centered talk. In some respects, the experimental group appeared to have a greater performance gain than the control group; in other respects, the effect is not yet clear. In addition, detailed analysis showed that the rate at which errors were committed was less for those specific elements in the solution in which bodies were cited. With data from additional subjects, we will seek to determine more definitively if a strategy based on body-centered talk offers significant benefits.

Acknowledgments

Support by the National Science Foundation under grant REC-0440295 and by the Department of Mechanical Engineering, Carnegie Mellon University is gratefully acknowledged.

Bibliographic Information

1. National Research Council, (1999), How People learn: Brain, Mind, Experience and School, Committee on Developments in the Science of Learning, Bransford, J.D., Brown, A.L., Cocking, R.R. (Eds.), Washington, D.C., National Academy Press.
2. Brown, A.L. (1980). Metacognitive development and reading. In Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence and Education, (Spiro, R.J., Bruce, B.C. and Brewer, W.F., eds.), Hillsdale, NJ: Erlbaum.
3. Flavell, J.H. (1985) Cognitive Development. Englewood Cliffs, NJ: Prentice-Hall.
4. Flavell, J.H. (1991). Understanding memory access. In Cognition and the Symbolic Processes: Applied and Ecological Perspectives, (Hoffman, R. and Palermo, D. eds.) Hillsdale, NJ: Erlbaum.
5. Palinscar, A. S., & Brown, A. L. (1986). Interactive teaching to promote independent learning from text. Reading Teacher, 39, 771-777.
6. Scardemalia, M., Bereiter, C. & Steinbach, R. (1984). Teachability of reflective processes in written composition. Cognitive Science, 8, 173-190.
7. Campione, J.C., Brown, A.L. and McConnell, M.L. (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 (R.I. Chales and E.A. Silver, eds.) (pp.93-114). Hillsdale, NJ: Lawrence Erlbaum.
8. 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 (Voss, J.F., Perkins, D.N. and Segal, J.W., eds.). Hillsdale, NJ: Erlbaum.
9. White, B.Y and Frederickson, J.R. (1998). Inquiry, modeling and metacognition: Making science accessible to all students of reflective processes in written composition. Cognition and Instruction, 16, 3-117.
10. Reif, F. and Scott, L.A., (1999), Teaching scientific thinking skills: students and computers coaching each other, American Journal of Physics, Vol. 67, pp. 819-831.
11. Lovett, M. C. (2001). A Collaborative convergence on studying reasoning processes: A case study in statistics. In S. Carver, & D. Klahr (Eds.) Cognition and instruction: Twenty-five years of progress (pp. 347-384). Mahwah, NJ: Erlbaum.
12. Klahr, D. and Carver, S.M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, Learning and Transfer. Cognitive Psychology, 20, 362-404.
13. Harris, T.A. and Jacobs, H.R., 1995, On Effective Methods to Teach Mechanical Design, J. Eng. Educ., Vol. 84, p. 343.
14. Steif, P. S., and Dollar, A., (2005), Reinventing the Teaching of Statics, International Journal of Engineering Education, Vol. 21, pp.723-729.
15. Steif , P. S., “An Articulation of the Concepts and Skills Which Underlie Engineering Statics,” Proceedings of the 34th Annual ASEE/IEE Frontiers in Education Conference, Savannah, GA, Oct. 2004, pp. F1F-5-10
16. Steif, P. S., and Dantzler, J. A., “A Statics Concept Inventory: Development and Psychometric Analysis,” Journal of Engineering Education Vol. 94, No. 4, October 2005, pp. 363-371.
17. Steif, P.S. and Hansen, M.A. “Comparisons Between Performances In A Statics Concept Inventory and Course Examinations”, *Int. Journal of Engineering Education*, Vol. 22, 2006, pp. 1070-1076.
18. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159-174.