AC 2007-1617: EFFECTS OF CONCEPTUAL UNDERSTANDING, MATH AND VISUALIZATION SKILLS ON PROBLEM-SOLVING IN STATICS

Kelli Higley, Pennsylvania State University

Kelli Higley is a PhD student in Educational Psychology at Penn State. Before working on her PhD, she taught high school mathematics for 3 years. She has worked on diverse projects about learning, including research about discourse, reading, statistics, algebra, and now Statics. Her primary research focus remains improving the quality of mathematics teaching. She can be contacted at kjh262@psu.edu.

Thomas Litzinger, Pennsylvania State University

Tom Litzinger is Director of the Leonhard Center for the Enhancement of Engineering Education and a Professor of Mechanical Engineering at Penn State, where he has been on the faculty since 1985. His work in engineering education involves curricular reform, teaching and learning innovations, faculty development, and assessment. He teaches and conducts research in the areas of combustion and thermal sciences. He can be contacted at tal2@psu.edu.

Peggy Van Meter, Pennsylvania State University

Peggy Van Meter is an Associate Professor of Education within the Educational Psychology program at Penn State where she has been on the faculty since 1996. Her research includes studies of the strategic and meta-cognitive processes that learners use to integrate multiple representations and acquire knowledge that will transfer and be useful in problem solving. She can be contacted at pnv1@psu.edu.

Christine B. Masters, Pennsylvania State University

Christine B. Masters is an Assistant Professor of Engineering Science and Mechanics at Penn State. She received a B.S. in Mechanical Engineering in 1987 and a Ph.D. in Engineering Science and Mechanics in 1992, both from Penn State. In addition to raising four children with her husband of 20 years, she teaches introductory mechanics courses, trains the department graduate teaching assistants, coordinates the Engineering Science Honors Program undergraduate advising efforts and participates in a variety of engineering educational initiatives such as the MechANEX project (software and lab experiment development for statics).

Jonna Kulikowich, Pennsylvania State University

Jonna Kulikowich is a Professor of Education within the Educational Psychology program at Penn State where she has been on the faculty since 2003. Prior to joining Penn State she was an Associate Professor of Education at the University of Connecticut. Her research includes studies of the Academic development in mathematics and statistics, applies statistics, measurement of variables in reading research. She can be contacted at jmk35@psu.edu.

Effects of Conceptual Understanding, Math and Visualization Skills on Problem-solving in Statics

Introduction

Although non-technical skills are increasingly important to successful engineering careers in the global marketplace of today, problem-solving remains a critical skill for most young engineers. In many cases successfully solving problems requires engineers to use their analytical skills. The central importance of problem-solving and analytical skills in engineering motivates the work presented in this paper, which is the first phase of a program aimed at answering two main questions: What are the major difficulties that students encounter when they perform modeling during problem-solving? What are the necessary components of instructional interventions to improve engineering students' modeling during problem-solving?

The work is being done in Statics classes because this is one of the first places that engineering students encounter the engineering problem-solving process. In this study we are paying particular attention to the early steps in problem-solving when students 'model' the system being studied to create a set of equations describing the system. In Statics students typically read a problem statement and then create a model of the system, the free body diagram, that contains all of the salient forces on the body. Then, based on the free body diagram, they create a mathematical model of the system.

Clearly there are many different ways in which students can go wrong as they solve problems in Statics. They may, for example, have inadequate knowledge of the forces and moments for particular types of joints, an inability to visualize forces, or inadequate math skills. Our working hypothesis is that students will cluster into different groups based on their abilities and knowledge, and that these groups will demonstrate differing abilities to solve Statics problems. Therefore, improving the problem-solving skills of these groups will require different interventions.

The work presented in this paper is designed to answer two research questions: can such clusters be identified, and if so, can they be used to identify the specific needs of the students in those clusters? The results presented include a summary of a cluster analysis, which did identify statistically significant clusters, and a comparison of the characteristics of the best and worst performing clusters to illustrate how the data can be used to identify the specific needs of the students in a cluster.

Relationship to Previous Work

This study has been influenced by a number of studies of problem-solving in general and of problem-solving in engineering specifically. The relationship to past work was discussed at some length in a previous paper[1] and therefore it is only briefly summarized here. Three subsets of the literature have had the most influence on our work: Problem-solving processes, translations between symbol systems, and domain knowledge.

Since Polya's seminal work in mathematics,[2] the utility of learning and using a sequence of steps during problem-solving has been widely accepted. Although several specific models exist, a generic 4-step model captures most: (1) Represent the Problem, (2) Goal Setting and Planning, (3) Execute the Plan, and (4) Evaluate the Solution. In the first step, problem representation, the student must read the problem statement and discern the objective. There are instructional interventions for engineering education that are grounded in this theoretical model of problem-solving. For example, Gray *et al*.[3] developed a systematic approach to solving Statics and Dynamics problems. In this intervention, it is recommended that students be taught the sequence of: Road Map (Planning), Modeling (Representation), Governing Equations (Representation), Computation (Execution), and Discussion and Verification (Evaluation). Don Woods completed some of the most thorough work that has been done in this area while developing the McMaster Problem-solving program.[4] In his most recent work,[5] Woods has focused on the processes of problem-solving and has developed a model to describe ideal problem-solving.

Without a doubt, the quantity of prior domain knowledge affects problem-solving.[6] It is also widely accepted that qualitative aspects of knowledge matter. Prior knowledge is believed to act as an important scaffold for problem-solving. The structure provided by the knowledge base can, for example, act as a constraint during analogical reasoning,[7] support strategic processing during reading,[8] and contribute to positive motivational states during problem-solving.[9] In short, the effects of prior knowledge are wide-reaching and powerful. Within the domain of Statics, Paul Steif closely examined the role of misconceptions[10] and developed a concept inventory in collaboration with Dantzler[11] to determine the effect of these misconceptions on problem-solving. Mehta and Danielson have developed and used a Statics skills and knowledge inventory.[12, 13]

A final approach to understanding problem-solving in engineering focuses on the symbol system translations inherent in the analysis process. By symbol system, we refer to the semiotic system used to understand and express elements and their relations. Mathematical expressions are an example of a semiotic system in which numbers and operators act as elements. How these elements are configured in relation to one another communicates the full meaning of the expression. Translations are required when problem solvers move between symbol systems. McCracken and Newstetter[14] developed the Text-Diagram-Symbol (TDS) model to capture the transformations that take place during analysis. This model includes verbal (Text), visual (Diagram), and mathematical (Symbol) semiotic systems through which the student must pass to complete an analysis task, with each phase corresponding to a different symbol system. The importance of visualization in transforming from a problem statement to a free-body diagram and the well documented gender effects on visualization skills, see for example,[15, 16, 17,] led us to include spatial reasoning instruments in the study.

Methodology

In order to identify clusters of students, data was collected on three types of measures: mathematics, spatial reasoning and conceptual knowledge related to Statics. A secure web site was created to provide participants with easy access to the measures. Upon completion and testing of the website, participants were recruited from ongoing Statics classes. Participants were offered 1% extra credit on their course grade for the completion of the measures. (Completion of the measures also made students eligible to participate in the second phase of the study using think-aloud protocols, which would garner them an additional 4% extra credit on their course grade.) Students were able to log in and out of the web site, enabling them to take the three measures in any order they chose and in multiple sittings if desired. During their first visit to the website, students read and indicated agreement with the informed consent and answered basic demographics questions, such as gender, race, SAT scores, major, and GPA. They then were brought to new page containing a separate link to each measure.

Ward's method of cluster analysis[18] was applied to the data to identify clusters whose members performed similarly on the measures. Ward's method forms groups by considering all possible pairs of participants, seeing which set has the least difference in their set of responses. After the first group is created, the mean of their responses are considered one group, and all possible sets are again considered. This iterative process is repeated until all participants are combined in one group. In the method used, the squared Euclidean distances are the measure of the differences between the groups. Participants are grouped so that within-group differences are minimized and differences between groups are maximized. Solutions with three to six clusters were studied using SPSS 14.0. The quality of the solutions was judged by their ability to predict an external criterion not used in the cluster analysis, which in this case was students' performance on problems from the mid-semester Statics examination.

Measures

The mathematics test was created from the ten math questions of the inventory developed by the Mehta and Danielson,[12] which is intended to measure students' knowledge of the prerequisite mathematics for a Statics course. Problems include solving basic equations for one- and two-variables, finding triangle characteristics through trigonometry and similarity, basic integration, and vector multiplication. Students were assigned a total score based on the number correct. The internal reliability based on Cronbach's coefficient alpha for this measure was 0.40.

Spatial reasoning was measured by two well-accepted measures in the field, Card Rotation and Paper Folding from the Factor-Referenced Cognitive Tests.[19] Both tests are timed, limiting the students to three minutes for each set of items (12 minutes total). The original tests were developed in paper and pencil format and were adapted for online use. The online versions were designed to be as much like the paper and pencil version as is possible. In the Card Rotation task, participants are asked to observe a target image, then determine whether eight other images are planar rotations of the figure, or other transformations such as mirror-image. Students indicate which of the images are equivalent to the original image. Scores are assigned by subtracting the number of incorrect responses from correct responses. The reported reliability for this measure is 0.80[19]; the reliability for our delivery was 0.97. In the Paper Folding task, a series of two to four folds are indicated through diagram, and various holes are punched into the folded paper. Participants are to choose which of five options has the correct hole configuration on the unfolded piece of paper. This score is found by awarding one point for an accurate response, and subtracting ¼ point for an incorrect response. The reported reliability for this measure is 0.84[19]; the reliability for our delivery was 0.72

Knowledge related to Statics was measured using the Statics Concept Inventory,[11, 20] which is a 27-item measure of the concepts that have been identified as key in Statics comprehension. The inventory is intended to only tap conceptual errors, so very little math is involved, and what math is used is trivial. The inventory measures nine areas of conceptual understanding, forces on collection of bodies, Newton's 3rd law, Static equivalence, roller forces, slot forces, negligible friction, representation, friction, and equilibrium. The reported reliability of this test is 0.83 for students who have completed a Statics class.[20] The reliability for the administration of the test in this study, which occurred midway through the Statics course, was 0.70.

Data that would serve as the external criterion were scores on two questions on the mid-semester examination. The first required students to create a free-body diagram and then solve for forces on the arm of the seat in an aircraft. The second required them to create a free body diagram of a suspended sign. For the first problem sub-scores were given for the free-body diagram, distributed load equivalent forces, and the accuracy of the equilibrium equations.

Sample

Of the approximately 480 students enrolled in the class, 390 students registered on the website. 62% of the students completed all three measures during one session; 38% used multiple sessions. Time spent on the website ranged from 10 minutes to 2 hours, with a mean of 58 minutes (SD = 16 minutes). Of the 390 students who registered on the website, only 367 completed all three measures. Because testing was done in an online environment, the reasons that students did not complete all measures could not be determined. However, a comparison of those who did and those who did not complete the measures indicated that they had similar demographics. In addition, means on the web measures were also compared for students who completed all measures and those who had not. An ANOVA indicated that the means were not significantly different (all F's less than 2.9), so we can assume that the missing data does not represent a non-completion bias.

As would be expected, not all students took the mid-semester exam the day it was scheduled. Because the data was analyzed immediately after the test was given, not all student tests were available. Thus, the exams scores that would serve as the external criterion to judge the quality of the clusters solutions were available for only 336 students. The demographic characteristics of the participants for whom the exam scores were available are summarized in Table 1. The majority of the participants were white (87%), male (88%), and sophomores (88%). The participants had an average SAT verbal score of 583, and SAT math score of 653 (all self-report).

Table 1 Sample demographics

	Count	Percent		Count	Percent
Gender			**Ethnicity**		
Male	297	88	African-American	3	1
Female	39	12	Asian	19	6
			Caucasian	292	87
Year in School			Hispanic	14	4.2
Freshman	3	1	Indian	2	0.6
Sophomore	298	89	Pacific Islander	1	0.3
Junior	26	8	Other	5	1.5
Senior	7	2			

Results and Discussion

Figure 1 compares the average scores on the inventories and the scores for the exam questions for the female and male students. In general the average scores for the male students were higher than those for the female students with two exceptions; female students had higher sub-scores for the accuracy of their equilibrium equations for exam problem 1, and their scores for calculations related to the distributed load were essentially equal to those of the male students. However, only the difference in Statics Concept Inventory score is statistically significant, with the female students scoring below the males. The fact that the female students scored higher for the accuracy of their equilibrium equations, but had lower scores of the free body diagrams (FBD 1 and FBD 2 in Figure 1) may reflect something significant, but more analysis, and a larger sample of female students, are required before any statistically significant conclusions can be drawn.

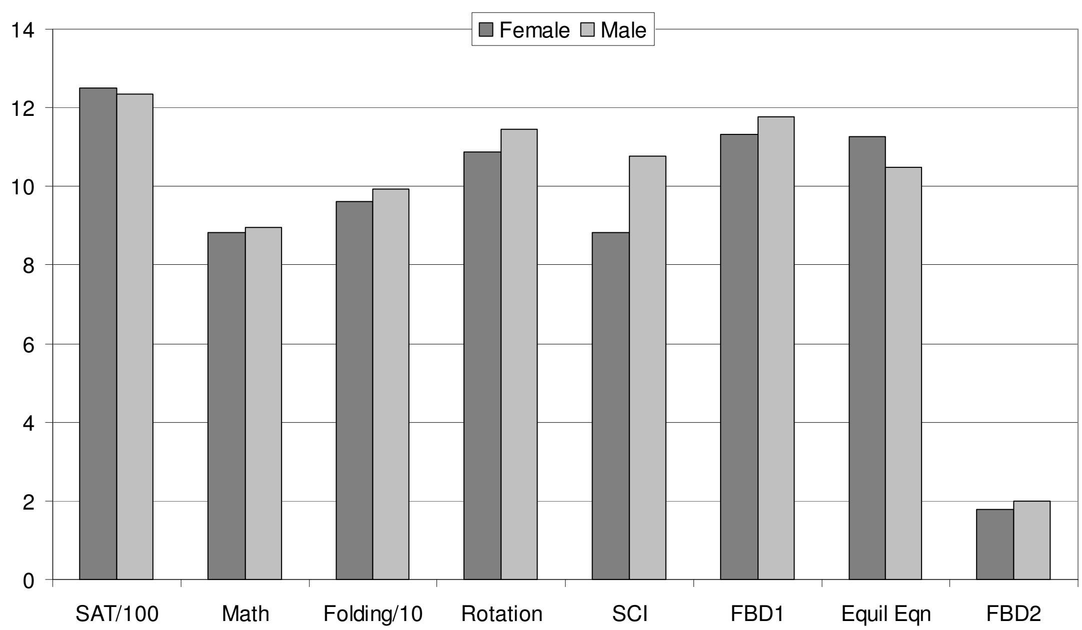

Figure 1: Comparison of inventory and external criterion scores for female and male participants

The cluster analysis was run using three of the four measures, and the students' total SAT score. Because of the low variability associated with the math baseline test (M = 8.92, SD = 1.09), that score was not included. The results of the cluster analysis for three to six clusters are shown schematically in Figure 1; the N values are the number of students in each cluster. As expected, the three cluster solution has a high performing cluster, a mid-performing cluster, and a low-performing cluster. As the number of clusters is increased, the mid-performing cluster splits into three groups, and the high cluster splits into two groups. For the four cluster solution, the mid-performing group splits based on SAT scores and scores on the Card Rotation task, with one group having a higher SAT and the other with a higher score on the Card Rotation task. In the five cluster solution, the lower group of the original mid-performing group splits based on scores on the Paper Folding test and the Statics Concept Inventory, with one group doing much better than the other on both measures. Interestingly the poorest performing cluster, CR6, actually emerges from the mid-performing group. In the six cluster solution, the high-performing group splits based on the total SAT scores and Statics Concept Inventory, with one sub-cluster excelling in SAT and the other excelling in the Statics Concept Inventory.

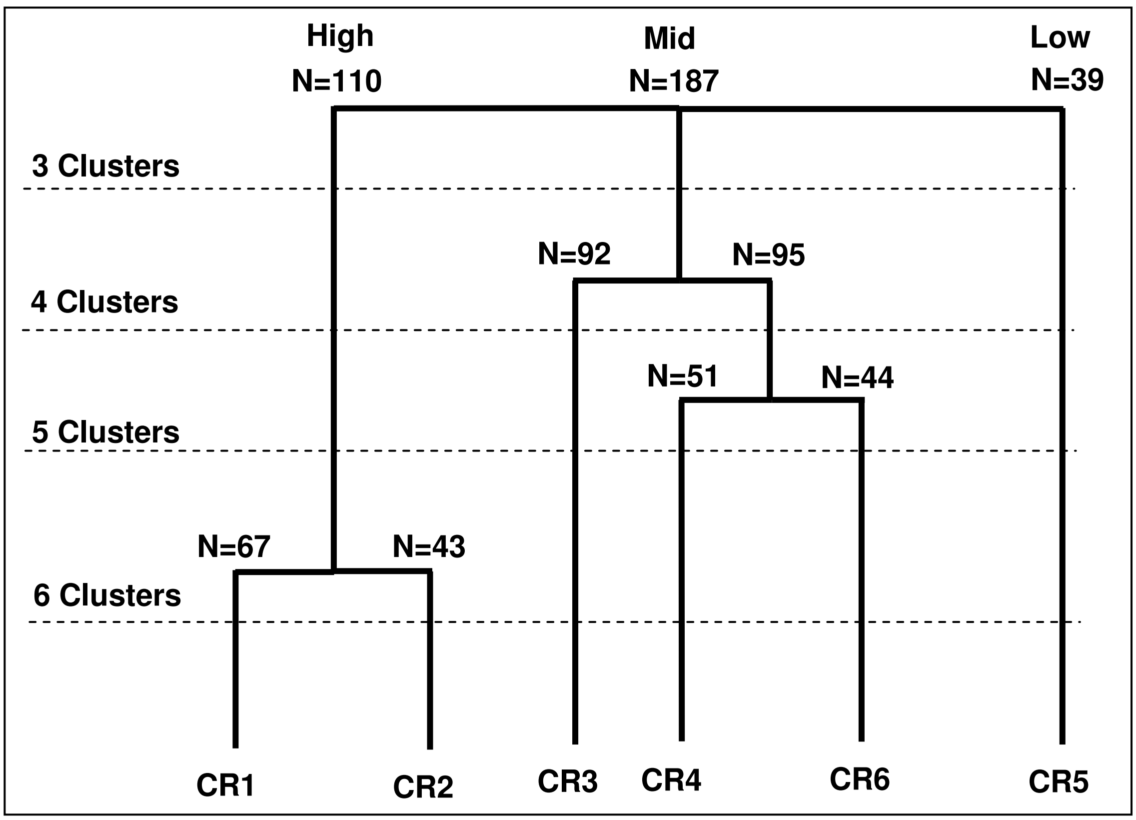

Figure 2: Cluster solution tree for 3 to 6 clusters

The splitting of clusters tends to happen somewhat along gender lines as well. The (Female/Male) composition of the six clusters are CR1 (2/65), CR2 (6/37); CR3 (15/77), CR4 (3/48), CR6 (8/36); and CR5 (5/34). The clusters with a higher representation of females tend to have higher SAT scores. It is not possible to draw conclusions about this information because of the small number of females in the sample.

With each division of a cluster, a more exact description of that cluster is possible. Yet, as the number of people within a cluster decreases, the stability of those characteristics also decreases. To judge which solution was most appropriate, each solution was considered as the predictor variable to the external criterion, which was the total score on the two exam questions. The solutions with five and six clusters were the most predictive. Within the General Linear Model (GLM) using the five cluster solution to predict the exam score, the F-statistic for the criterion was 11.45 ($p < .001$) and had an effect size, partial-eta-squared, of 0.122. A GLM based on the six cluster solution had a lower F-statistic of 9.19 ($p < .001$), but an equivalent effect size of 0.122 (partial-eta-squared). These are both high-level effect sizes, representing differences between the scores of members of those groups of approximately 0.8 standard deviations. Ultimately, the six cluster solution was chosen because the clusters were more informative. Thus the first research question for this portion of the study was answered positively; statistically significant clusters of students were found. The next challenge was to interpret them.

Figure 3 and 4 present the means for the six cluster solution on the measures used in the cluster analysis and for the performance on the exam items. Substantial differences are present in many cases; however, not all are statistically significant. A series of Mann-Whitney U tests were conducted to determine significant differences at $\alpha = 0.05$. This test is similar to a t-test and is interpreted the same way. Thus, if a significant difference is found, it can be assumed that the two clusters being compared are different, supporting the cluster solution. The results of that comparison are presented in Table 3 in which the clusters 2 through 6 are compared to cluster 1, the highest performing cluster.

The entries in the table indicate those measures for which that cluster had lower performance than cluster 1. This table begins to show the most important differences between the clusters, which will guide the selection of the interventions that are required to raise the level of performance of that cluster. Clusters 5 and 6, the two poorest performing clusters, have lower free body diagram scores than cluster 1. They both have lower scores on the Statics Concept Inventory and on the math baseline test. These differences indicate that basic Statics knowledge and math skills may be causing the students to perform poorly. More detailed investigation is required, however, to understand the types of issues that are causing the clusters to perform more poorly than cluster 1.

Pair-wise comparison of clusters is used to determine the types of errors the students in the lower performing clusters are making. A comparison of clusters 1 and 6 is provided to illustrate how this comparison is conducted and the types of results that it yields. Cluster 6 has lower performance on all four of the measures that were used in the cluster analysis. The analysis concentrated on the Statics Concept Inventory and the math baseline test because a regression analysis showed that the Paper Folding and Card Rotation scores were not significant predictors of the exam scores.

For the math measure, the only item with an effect size greater than 0.1 was one requiring differentiation to find a minimum. This particular skill was not required in either of the Statics problem used as the external criterion so it does not shed much light on the differential performance of the two groups. For the Statics Concept Inventory, however, the process yielded much more meaningful results.

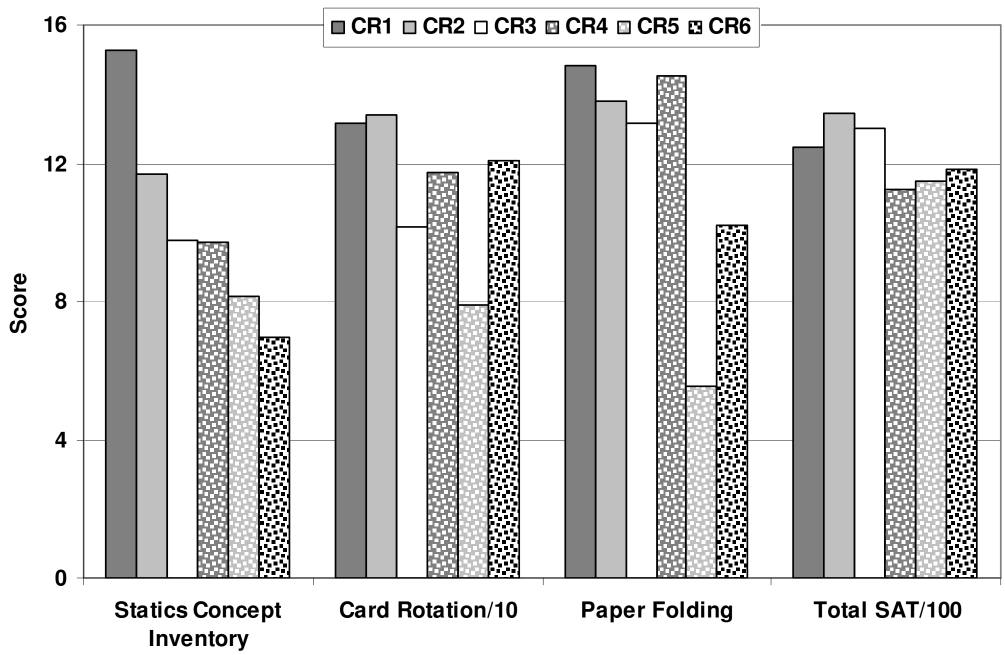

Figure 3: Means of SCI, visualization measures and self-reported SAT for the six clusters

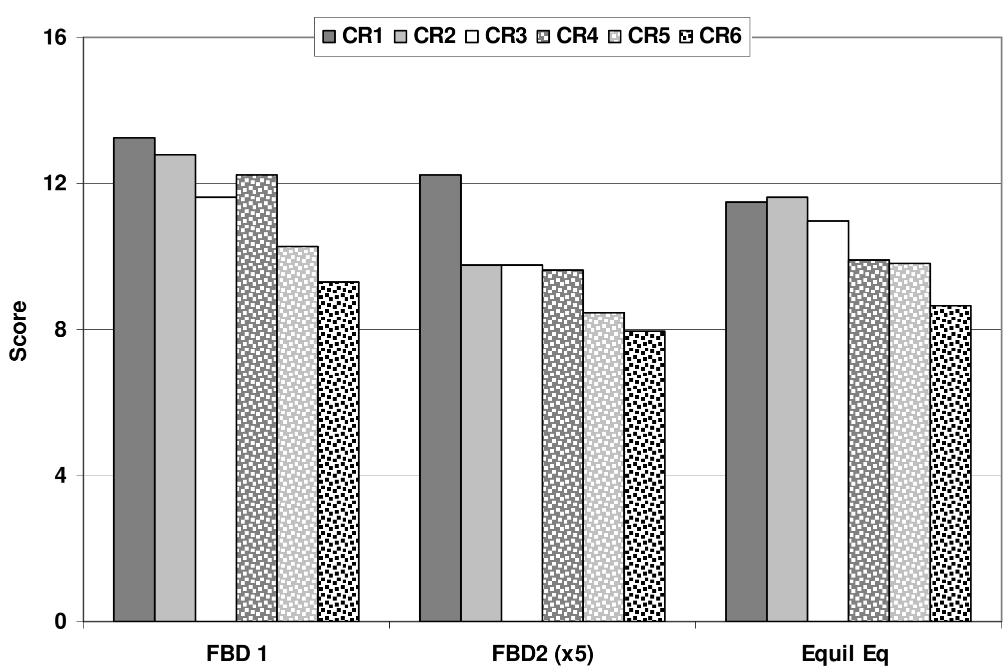

Figure 4: Means of exam items for the six clusters

Table 3 Comparison of performance of clusters to that of highest performing cluster
(An x indicates a significant difference between the cluster and cluster 1 at $\alpha = 0.05$)

	Cluster				
Variable	CR2	CR3	CR4	CR5	CR6
Statics Concept Inventory	x	x	x	x	x
Math Baseline				x	x
Card Rotation		x	x	x	x
Paper folding	x	x		x	x
Free body diagram1		x		x	x
Distributed load analysis					x
Equilibrium equations			x	x	x
Free body diagram 2				x	x

Seven items from the SCI were identified with effects sizes greater than 0.1 between clusters 1 and 6. Three of the seven items involved establishing the direction of a reaction force between a curved body and a flat surface. Three other items involved the modeling of reactions at slotted or non-slotted pin joints. The last item involved creating a free-body diagram of a body with multiple interconnected masses. Most compelling are the curved body meeting a flat surface and the modeling of pin reactions. The answers to these problems by students in cluster 6 indicate that they lack very basic knowledge of how a particular support restricts motion and subsequently generates reactions. Our previous work using a think-aloud methodology to study the process of creating free-body diagrams indicates that, when student lack such basic knowledge, they will fail to accurately complete free body diagrams.[1] These results indicate that students in cluster 6 would benefit substantially from some form of intervention that required them to internalize the modeling of basic supports.

Conclusions

The work presented in this paper has demonstrated that clusters can be identified and used to predict the exam performance of students in Statics. These clusters have predictive capability equivalent to large effect sizes. Further the work has shown that the data collected for the cluster analysis can be used to determine the types of errors that the students are making. These errors in turn can be used as a starting point for identifying the interventions that are required. More insight into the differences among the clusters and the types of interventions required to address them will be obtained through ongoing analysis of the cluster results and through the think-aloud portion of the study that is currently underway.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant EEC- 0550707. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The authors would like to thank Paul Steif, Sudhir Mehta, and Scott Danielson for allowing us to use their inventories in our study.

References

[1]Van Meter, P., Litzinger, T., Wright, M., and Kulikowich, J., “A Cognitive Study of Modeling during Problem-solving: The integrated problem-solving model”, Proceedings of the 2006 ASEE Annual Conference.
[2] Polya, G. (1957). How to Solve It (2nd Ed.), Princeton University Press.
[3] Gray, G., F. Costanzo, and M. E. Plesha, “Problem-solving in Statics and Dynamics: A Proposal for a Structured Approach,” Proceedings of the 2005 ASEE Annual Conference and Exposition, Portland, OR, 2005.
[4] Woods, D.R., Hrymak, A.N., Marshall, R.R., Wood, P.E., Crowe, C.M., Hoffman, T.W., Wright, J.D., Taylor, P.A., Woodhouse, K.A., and Bouchard, C.G.K., “Developing Problem-solving Skills: The McMaster Problem-solving Program”, Journal of Engineering Education, Vol. 86, No. 2, 1997, pp. 75-91.
[5] Woods, D.R., “An Evidence-based Strategy for Problem-solving,” *Journal of Engineering Education*, Vol.89, No.3, pp. 443-459, 2000.
[6] Gelman, R., & Greeno, J. G., On the nature of competence: Principles for understanding in a domain. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 125-186). Hillsdale, NJ: Erlbaum.
[7] Gentner, D. “Structure-mapping: a theoretical framework for analogy” *Cognitive Science*, Vol. 7, pp. 155-170, 1983.
[8] Hegarty-Hazel, E. & Prosser, M. “Relationship between students’ conceptual knowledge and study strategies – Part 2: Student learning in biology” *International Journal of Science Education,* Vol. 13, No. 4, pp. 421-430, 1991
[9] Alexander, P.A., Jetton, T.L., & Kulikowich, J.M. “Interrelationship of knowledge, interest, and recall: assessing a model of domain learning.” *Journal of Educational Psychology*, Vol. 87, pp. 559-575, 1995.
[10] Steif, Paul S., “An Articulation of Concepts and Skills Which Underlie Engineering Statics,” Proceedings of the 34th ASEE/IEEE Frontiers in Education Conference, Savannah, GA, 2004
[11] Steif, P.S. and Dantzler, J. A., “A Statics Concept Inventory: Development And Psychometric Analysis”, Journal of Engineering Education, Vol. 33, pp. 363-371 (2005)
[12] Mehta, S. and S. Danielson, “Math-Statics Baseline (MSB) Test: Phase I,” Proceedings of the 2002 ASEE Annual Conference and Exposition, Montreal, CA, 2004.
[13] Danielson, S., Kadlowec, J., Mehta, S., Masters, C., Magill,M., and Steadman, S. “Work in Progress – A Statics Skills Inventory,” Proceedings of 2005 FIE Conference
[14] McCracken, W.M. and W.C. Newstetter, “Text to Diagram to Symbol: Representational Transformation in Problem-Solving,” Proceedings of the 31st ASEE/IEEE Frontiers in Education Conference, Reno, NV, 2001
[15] Sorby, S. A., “A Course in Spatial Visualization and its Impact on the Retention of Female Engineering Students,” *Journal of Women and Minorities in Science and Engineering*, Vol. 7, pp. 153-172, 2001
[16] Sorby, S.A., and B.J. Baartmans, “The Development and Assessment of a Course for Enhancing the 3-D Spatial Visualization Skills of First Year Engineering Students,” *Journal of Engineering Education,* Vol. 89, No. 3, 301-307, 2000.
[17] Devon, R., R.Engel, and G. Turner, “The Effects of Spatial Visualization Skills Training on Gender and Retention in Engineering,” *Journal of Women and Minorities in Science and Engineering,* Vol. 4, No. 4, pp371-380, 1998.
[18] Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58,* 236-244.
[19] Ekstrom, R. B., French, J. W. and Harman, H. H. (1976) *Manual for Kit of factor-referenced cognitive tests*. Princeton: Educational Testing Service.
[20] Steif, P.S. and Hansen, M.A, Comparisons Between Performances In A Statics Concept Inventory and Course Examinations, International Journal of Engineering Education, Vol. 22, pp.1070-1076 (2006)