Improving problem-solving skills of undergraduates through computerized dynamic assessment

Nagamany Nirmalakhandan a *

aCivil Engineering Department, New Mexico State University, MSC 3CE, Las Cruce, NM 88003, USA

Abstract

A computer-based dynamic assessment (DA) system developed for use in an undergraduate fluid mechanics course is presented. Data collected before and after implementation of this DA indicate significant improvement in student performance after implementation. Student performance is quantified by the % of questions correctly answered in the nationally normed Fundamentals of Engineering (FE) Exam relative to the National average. Since implementation of the DA system, this measure for our students has increased from below National level (mean = 0.942; standard deviation, sd = 0.068) to above National level (mean = 1.04; sd = 0.06). For the same population, performance in fluid mechanics has been higher than that in the other subjects where DA was not used (mean = 1.068; sd = 0.028 vs. mean = 0.837; sd = 0.04). Performance of our students in fluid mechanics has exceeded that of their peers in the top tier programs in the U.S. (mean = 1.068; sd = 0.028 vs. mean = 1.031; sd = 0.03).

Keywords: Problem-solving, dynamic assessment, computer-based assessment, metacognitive understanding;

1. Introduction

Dynamic assessment (DA) is a subset of interactive assessment techniques where, the process of learning and knowledge acquisition are tracked so that instruction could be modified to improve student achievement. It involves planned mediation of teaching and the assessment of effects of that teaching on subsequent performance (Campione & Brown, 1990). DA procedures have been shown to yield different types of information including: more valid measures of student abilities than through static tests; measures of learning ability or “modifiability”; insights into the cognitive processes that students use or fail to use; and clues about instructional methods (Daniel, 1997; Elliott, 2003). Almost all researchers working on DA have found that test performance improves after mediation through DA (Campione & Brown, 1990; Embreston, 1990; Daniel, 1997; Haywood & Tzuriel, 2002; Elliott, 2003). It is in contrast to traditional static tests that test acquired knowledge, without any attempt to intervene in order to change, guide, or improve the students’ ability to learn and potential for achievement (Daniel, 1997; Shepard, 2000; Haywood & Tzuriel, 2002).

Several other benefits of dynamic assessment have been recognized in the cognitive research literature. DA with diagnostic monitoring and context-sensitive prompting and feedback has been found to be an effective...
approach to improve student achievement (Campione & Brown, 1990). DA facilitates near and far transfer of mediated strategies to the solving of new problems (Campione & Brown, 1990; Burns, 1991; Elliott, 2003). Extent of gain in DA tasks has been shown to be a good predictor of later academic accomplishments (Campione & Brown, 1990).

However, a negative aspect of DA is that classroom implementation of DA demands considerable effort and time on the part of the instructor. As such, we have developed a prototype version of a computer-based DA system for use in an undergraduate fluid mechanics course (CE 331) at New Mexico State University (NMSU). The first version of the computerized assessment system initiated in 2000 did not incorporate dynamic assessment. Since its implementation, it has been formatively refined over several semesters incorporating student feedback on its usability and clarity as well as research reports on dynamic assessment. The current version of the DA system has been in use since 2004. Details of this system, its development and refinement, and its validity have been presented elsewhere (Nirmalakhandan et al., 2004; Nirmalakhandan, 2004, 2007, and 2009). In this paper, we present multiple measures collected over several semesters to demonstrate the effectiveness of the computerized DA system in improving student performance.

### 1.1. Performance Index

Recognizing that it is not possible to make direct assessment of student learning and achievement and relate that to specific interventions and remedial actions, we propose the use of the results of the Fundamentals of Engineering (FE) examination as an indirect, external measure. The FE Exam, administered biannually by the National Council of Examiners for Engineering and Surveying (NCEES), is a nationally normed exam that over 6,000 civil engineering graduates take every year during their senior year in college. This exam has two 4-hr sessions, one in the morning and one in the afternoon. The morning session of this test covers 12 subject areas common to all fields of engineering, including fluid mechanics. A summary report of the results of the FE exam showing the % of questions answered correctly in each subject area by the program students as a group is provided by NCEES to the students’ departments. This report also includes corresponding percentages for candidates from three comparator groups- candidates from the Carnegie 1 (Very High Research), Carnegie 2 (High Research), and Carnegie 3 (Masters) institutions, as well as the overall National average.

We have used a performance index, PI, defined as follows to assess the improvement of our students:

$$PI_{j,k} = \frac{\% \text{ of questions correctly answered by group } j}{\text{National average of questions correctly answered in subject } k}$$

In this paper, the PI is used in the following three ways:

1. Comparison of PI of our students in fluid mechanics before and after implementation of the DA system ($j = \text{NMSU}; k = \text{fluid mechanics}$)
2. Comparison of PI of our students in fluid mechanics against their PI in other subjects ($j = \text{NMSU}, k = \text{fluid mechanics vs. other subjects}$)
3. Comparison of PI of our students against PI of peers in Carnegie institutions, in fluid mechanics ($j = \text{NMSU vs. Carnegie Institutions}; k = \text{fluid mechanics}$)

### 1.2. Pre- and post-test

To further validate the improvements in the students’ problem-solving skills, another computer-based pre- and post-test system (P-P Test) has been developed and implemented during the past 4 years. The intent was to assess the improvement in students’ conceptual understanding that is essential for problem solving; and to assess the students’ metacognitive understanding of the concepts. The students’ confidence in the correctness of their answers to each problem was taken as a measure of their metacognitive understanding.
This computerized P-P Test consisted of 10 conceptual problems designed around the Continuity Equation. Each problem had three surface modifications, so that at run time, a randomly chosen version is presented to the student. The problems are multiple-choice questions with four choices; none of them require any calculations. Upon making a selection, the students were requested to indicate their confidence level in the correctness of their selection, from a list of five levels of confidence: clueless, unsure, somewhat sure, pretty sure, and 100% sure. Once the confidence level is selected, immediate feedback is given about the student’s response to the problem including the correct answer, if the selection was not the correct answer. No explanation is provided.

The system kept track of the students’ response to the problems as well as their confidence level associated with their selection for each of the 10 problems. The system produces a report summarizing the results by problems and by students. During the use of the system, students were not told what the concepts were that were associated with each problem. They were also not told that there would be a post-test, and the problems were not specifically discussed during regular class. For the purposes of analysis, correct and incorrect items on each problem were scored as 1 and 0, respectively. Confidence measures allowed were clueless, unsure, somewhat sure, pretty sure, and 100% certain. These metacognitive judgments were converted to the numerical values, 0 to 4.

2. Results

2.1. Comparison of PI: before and after implementation of DA

Prior to initiation of the computerized system in 2000, the performance of our students in the FE exam was significantly below the National level, with average PI of 0.872 (sd = 0.148). During the initial stages of the implementation of the system (2000 to 2003), PI increased to 0.964 (sd = 0.081). Since implementing DA in Spring 2004, PI has increased further to 1.068 (sd = 0.028), exceeding the National performance (Nirmalakhandan et al. 2004). Since the instructor and the teaching methods have remained almost the same since 2000, the gradual increase in the performance in the FE Exam is attributed primarily to the computerized DA system. This claim is corroborated further by additional analysis of the FE Exam results as discussed next.

2.2. Comparison of PI: fluid mechanics versus other subjects

Figure 1 compares the performance of our students in fluid mechanics against the average of their performance in the other 11 subjects covered in the morning session of the FE exam. The following two observations can be noted from this figure. First, there is a step increase in the performance in fluid mechanics, coinciding with the implementation of the DA system; the performance (as quantified by mean PI) increased from below National level of 0.942 (sd = 0.068) pre-DA to above National level of 1.068 (sd = 0.028) post-DA. It has to be noted that the students take the FE exam 3 semesters after they take this course and the instructor had remained the same over pre-DA and post-DA periods.

Second, the performance in fluid mechanics post-DA has been significantly higher than that in the other subject areas, which has remained consistently below National level from 2000, at a mean PI of 0.841 (sd = 0.029). In fact, there is no significant change between the mean PI for the other subjects, pre-DA (mean PI = 0.834; sd = 0.028) versus post-DA (mean PI = 0.844; sd = 0.029). This comparison for the same population of students supports the claim that their higher performance in fluid mechanics is probably due to the DA system that was used only in the fluids mechanics course.
2.3. Comparison of PI: our students versus Carnegie peers

A comparison of the PI of our students in fluid mechanics against that of the three comparator groups over the past seven administrations of the FE Exam is shown in Figure 2. As can be seen from this figure, pre-DA performance of our students had been below the performance of Carnegie 1 and Carnegie 2 peers. However, post-DA performance has been consistently and significantly above that of Carnegie 1 peers: mean PI = 1.068 (sd = 0.028) versus mean PI = 1.022 (sd = 0.020). This comparison affirms that the step improvement in performance of our students in Spring 2004 is not due to any fluctuations in the standard of the fluid mechanics section of the FE exam, but due to the DA system used at NMSU that helped improve students’ achievement.
It is worth noting that our students take the fluid mechanics course (CE 331) during the junior year and take FE exam about 3 semesters later in their senior year. They do not take any further courses in this area beyond CE 331. Yet, the FE exam results indicate that the skills developed and the knowledge gained using the computerized DA system were long lasting for successful far transfer. This benefit of the DA approach is also in agreement with similar findings reported in the literature (Campione & Brown, 1990; Burns, 1991; Elliott, 2003).

The box-and-whisker plot in Figure 3 summarizes the above comparisons in terms of 10th percentile, 25th percentile, mean (o), median, 75th percentile, and 90th percentile of the PI values for the different groups. These comparisons validate the notion that the computerized DA system presented in this paper is beneficial to the students in improving their problem-solving skills and achievement in the FE exam. The system enables students to learn the material by working problems individually, with help provided by the DA system. In contrast to traditional homework assignments where students tend to work on the problems in groups, this system helps students to solve problems individually and learn from their errors by themselves, with immediate feedback and prompting. This feature of the system that cultivates individual competence could be a reason for the increased performance of the students in the FE exam, which is designed to measure individual competency rather than group effort.

![Figure 3](image)

Figure 3. Percentage of questions correctly answered in the morning section of the FE Exam: NMSU students relative to National peers: s- number of semesters; n- number of students taking the FE test

2.4. Results of pre- and post tests

The number of students who answered each of the 10 problems correctly in the pre-test in a typical semester are compared against those who did so in the post-test in Figure 4. The averaged results (pre-test = 15.2 versus post-test = 18.3) suggest that the students’ conceptual understanding has improved over the semester.
The confidence levels expressed by the students in the pre-test are compared against those in the post-test in Figure 5. As can be observed in Figure 5, the confidence levels of the students had improved in the post-tests. The above improvements are probably due to the implementation of the dynamic assessment system.
3. Conclusions

The results presented in this paper indicate that the proposed computer-based dynamic assessment system has the potential to improve student achievement. While these results are in agreement with the previous findings about DA, this study demonstrated the advantages of implementing a computer-based DA system, whereby the instructor’s time commitments are significantly reduced. The use of FE Exam results as an external measure of students’ problem solving skills enabled long-term assessment of the computer-based DA system. The analysis of the results from the computer-based pre- and post-test approach proposed here needs to be refined further to relate them more rigorously to metacognitive learning.

Acknowledgements

Funding for this study was provided by the CCLI Program of the National Science Foundation.

References

Campione, J. C. and Brown, A. L. (1990) “Guided learning and transfer: Implications for approaches to assessment” in Diagnostic Monitoring of Skill and Knowledge Acquisition, Ed. Frederiksen, N.,