Views from the Classroom: Teachers' Opinions of Statewide Testing Programs

This article addresses the question: How do teachers perceive the effects of high-stakes testing programs, particularly in the area of teaching and learning? The authors chose to not collect their own data for this article, but rather use existing data from a survey conducted by the National Board on Educational Testing and Public Policy (Pedulla et all, 2003). The authors of this article were involved with the collection of data for that survey and the write up titled, *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers,* which was published in a Boston College paper.

The article is logically organized by being divided into three sections. The first section gives an overview of the literature available on teachers' perceptions of state-mandated testing programs and previous or similar research. The previous research mentioned generally covers smaller areas of the topic. This research is different because a national survey was conducted, not just a focus on a single state.

The second section presents the findings from the nationwide survey of teachers mentioned above. This section explains how the survey was developed, topics that were focused on, and who took part in the survey. There was not a great deal of information as to how the survey was conducted or specific questions that were used. Survey participants were divided into two groups depending on if there state was considered to be a high/high stakes state or a moderate/low stakes state. The first descriptor refers to the severity of consequences for districts, schools, and/or teachers; the second descriptor refers to the severity of consequences for students based on the statewide testing based on a high/moderate/low scale.

Survey results were listed in the final section of this article. Emphases was given to the impact statewide testing has on classroom instruction and assessment, pressure to raise test scores and prepare students for the state test, impact on students and teachers motivation and morale, and teacher views on accountability. Without being able to see the actual data collected, the interpretation of the results seems accurate based on my limited knowledge of the effects statewide testing. However, the authors attempted to make a case that statewide testing has caused teachers to teach in ways that contradict

their own notions of sound educational practice, but did not explicitly back this with the results from the survey. In my opinion, this area could have used a little more attention by the authors.

This article caught my attention because it addressed assessment in the classroom, but not in the way I had hoped. If the authors had finished addressing the idea that state testing forced teachers to contradict sound educational practices or addressed the types of testing used in statewide testing, this article would have been a tremendous asset to my research proposal. As it did not address these issues, the article will only be a satisfactory asset to my collection. Overall the authors answered their research question, but this study leaves the reader wanting to know more about areas that were not addressed.

Can Computer-Based Testing Achieve Quality and Efficiency in Assessment?

This article attempts to address the question: Can computer-based testing (CBT) achieve quality and efficiency in assessment? It is based on a two-year study conducted at a university in the United Kingdom. The article is logically divided into several sections that flowed well from one to the next, including data gathering, method used, results interpretation, and implication for further research. However, there is no mention of studies previously conducted or similar research problems.

Data collection from the 1,000 students in undergraduate classes took place over a period of nineteen months. Questionnaires, observation, interviews, and an analysis of student scores on both CBTs and paper-based exams were all used to gather data. In my opinion, the write-up of the data collection makes the process very confusing and therefore would be very difficult to duplicate.

Results were deciphered using spreadsheets and statistical programs. The research provided student input on attitude, fairness, assessment of skills, perceived impartiality, perceived difficulty, and perceived performance level in regards to computer-based assessment. This was then compared to similar aspects of paper-based assessment. Overall, the authors found that computer-based assessment is a better option for students in terms of comfort, relaxation, anxiety, time, and scores. Based on the data provided, I was not as easily convinced of these findings, but agree that a well-planned CBT has potential.

It was not clear until the end of the article that the CBTs involved were used only to give multiple-choice questions. The main difference to the CBTs used and paper-based testing was the opportunity for immediate feedback and confirmation before submitting an answer. This study did not address the opportunity of alternative methods of assessment with computers, however the article is still an acceptable addition to my research topic because it pointed out that this type of CBT operated at only the knowledge level of Bloom's Taxonomy of Educational Objectives. My favorite part about the article was the warning label in the first paragraph advising against those who may try to implement a computer-based assessment system without careful planning.

Every research article should come with such a warning label to remind us that a quick fix is not the best solution.

The Impact of Computer-Based Testing on Student Attitudes and Behaviors

This article seeks to research the impact that computer-based testing has on student attitudes and behaviors, specifically their anxiety, readiness, convenience, and control. It is logically set up, starting with an introduction, followed by discussing the preliminary assessment, then the methods used, followed by the results and summary. I found this article relevant to my study because of its insight into computer-based testing and student's attitudes.

The introduction mentions previous studies; one study encourages increasing the frequency of exams as one way to increase student success. The other studies mentioned are about negative correlations between class size and grades, however, this did not relate to the remainder of the article and therefore seemed irrelevant. Also included in this section is the setup for the study. A small, proctored, computer-based testing (PCBT) facility at Ball State University is where this study took place. The study then examines the effects that the PCBT has on student's perceptions and success.

A preliminary assessment was conducted using students who had previously taken a class that used the PCBT for outside classroom testing. The researcher then decided to look into student attitudes concerning the PCBT and the nature of cheating. The participants for the study (908 volunteers) were from 25 different classes, almost half using PCBT and the rest not. There was a great difference between the amounts of tests given in the PCBT classes versus the paper-based exam classes. An online survey was used to collect the data, which was included in the article. The methods procedures are fairly straightforward and therefore it would not be too difficult to recreate this study, especially since the online survey is available. However, the survey had several questions that should be included in the bad survey questions handout given in class.

The data from the survey was given in multiple charts where one item (preferred frequency of exams, anxiety, readiness for the exam, convenience to take the exam, and perceived control) was plotted with the actual number exams taken. This was then graphed on two lines: those using the PBCT and those taking in class, paper-based exams. In many charts, the PBCT seemed to be more favorable with the students. However, in some charts (for example, readiness) the data jumped and was not consistent

for the PBCT but was for the paper-based. As for the cheating element, it was determined that students who took there tests in the PBCT were more likely to talk generally about the test, whereas paper-based exam students were more likely to share specific questions/answers.

The data was interpreted to mean that a greater number of exams, but not too great, in the PBCT were more favorable for students in terms of attitudes. My biggest concern with this study was the huge factor of the difference in the amount of exams given in the PBCT versus the classroom. The researcher pointed out that students were generally less stressed, for example, about their grades because they had more tests to average a the test scores out. Also, students generally do better when there are more tests, because less is covered in each test. This idea was not even considered in the study. It seems to me that in order to have a truly helpful study, these underlying factors need to be considered.

Interactive Assessment and Course Transformation Using Web-Based Tools

This study was conducted by a college professor who desired to know the effects that web-based tools could have on assessment and the course as a whole. The article does not mention any previous studies conducted on a similar topic, however there are many entries in the reference section with titles similar to the topic. It is difficult to tell if the researcher actually took previous studies into consideration, but I am not convince that she did.

The article arrangement is coherent enough, but the other articles I read were much better. The sections after the introduction were as follows: theoretical framework, techniques, data source, and the conclusion. Data interpretation and results was included mostly in the conclusion, but a little was mentioned in the data source section, as well.

The data was generated from students' opinions in a multimedia production course using an online course evaluation/survey and the professor's opinion. The survey addressed issues about the class, both the technological side and non-technological side of it. While many data generating examples were included in the article, I believe that this study would be very difficult to reproduce based on the individuality of the study. This study was solely developed for this particular course and would require much transformation to be applicable to other courses.

Another negative aspect of this article was the minor attention to the results of the survey. This article primarily focused on the techniques and methods rather than the interpretations. It is difficult for me then to make a formal opinion on the interpretations. On a positive note, the article did have a single major strength. The research topic was personally meaningful to the researcher. As all instructional designers know, this is key to a memorable learning experience. Before this article, I don't think I would have seen the value a personally meaningful experience incorporated into a research project quite so clear.

I had hoped that this article would fit into my research study by giving insight into alternative assessment methods by way of computer-based testing. However, it did not. Unfortunately, not every research article is helpful in the way that you had hoped. Then

again, the article did make apparent items to consider not doing when conducting my own research, as mentioned above, and the relevance of personally meaningful research.

Assessment and Grading in High School Mathematics Classrooms

I selected this article for two reasons. The first is because it is about assessment strategies in the classroom, which is what my research paper will focus on. The second reason is that it is about the mathematics classroom, an environment that I have just recently left. The main focuses of the study is the question what are teacher's perspectives as they assess their students' performance and translate the results of their assessment into grades? There are also minor, more detailed questions: what types of instruments do high school mathematics teachers use for assessment in their classrooms, how do teachers use these instruments to determine students' grades, to what extent are teachers' assessments and grading practices consistent with the recommendations about assessment in the *Curriculum and Evaluation Standards* from the National Council of Teachers of Mathematics, and what factors account for teachers' current assessment and grading practices. This article is by far the best I have seem in terms of structure, content, and research.

Previous similar studies, as well as similar studies occurring at the same time, are brought up in the introduction and then referred to throughout the article. This research study collected its own data from nineteen mathematics classrooms, in five high schools, in three states by questionnaires, interviews with teachers, and a compilation of assessment tools used. Categories and coding schemes were invented to compare all the data collected. Examples of such are included in the article and most procedures are explained. The one negative aspect of this article is that the study may be difficult to reproduce because of its complexity in data organization.

Unfortunately, the study showed that the majority of teachers employed tests and quizzes, mostly multiple-choice questions, addressing lower-level knowledge. Very few teachers incorporate open-ended items or other similar methods into their assessment. Twelve out of the nineteen classes surveyed also integrated other forms of assessment, such as written projects or student interviews along with the standard type of assessment. Lastly, questions that encourage the use of technology for more advanced reasoning were not used.

From reading the article, I am inclined to agree with the interpretations as stated. While this article was very well written and the study very well executed, I am sad to find out the result that there are not more teachers making use of the wonderful technological advances out there related to mathematics assessment and mathematics in general. Hopefully in time the numbers will increase.

References

Abrams, L., Pedulla, J., & Madaus, G. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, v. 42 no. 1,* p. 18–29.

Bocij, P. & Greaskey, A. (1999). Can computer-based testing achieve quality and efficiency in assessment? *International Journal of Educational Technology, v. 1 no. 1.* September 13, 2003. http://www.ao.uiuc.edu/ijet/v1n1/bocij/index.html

Butler, D. (2003). The impact of computer-based testing on student attitudes and behavior. *The Technology Source, v. January/February.* September 14, 2003. http://ts.mivu.org/default.asp?show=article&id=1013.

Byers, C. (2002). Interactive assessment and course transformation using web-based tools. *The Technology Source, v. May/June.* September 14, 2003. http://ts.mivu.org/default.asp?show=article&id=9388.

Senk, S., Beckmann, C., & Thompson, D. (1997). Assessment and grading in high school mathematics classrooms. *Journal for Research in Mathematics Education, v. 28, no. 2,* p. 187 – 215.