

ACTIVITY 10: CRITICAL EVALUATION OF A QUANTATIVE STUDY



COLLABORATIVE LEARNING ACTIVITY: #2

037659 – Emilio Cuauhtémoc Ruiz Esparza Flores, Saltillo Coahuila.

229155 - Linda Luz Licón Gavaldón, Cd. Obregón Sonora.

449762 – Alejandra de Jesús Magaña de León, México DF.

986258 – Verónica Zapata Mociños, Veracruz Ver.

1001319- Chelsea Olson, Phoenix Az.

ED 4002– CRITICAL THINKING STRATEGIES FOR
EDUCATIONAL RESEARCH

Professor:

Dr. Ricardo Valenzuela

Tutor:

Dr. Armando Lozano

March 9, 2005

The following paper is critical analysis of:

Hickey, D. T., Moore, A. L. and Pellegrino, J. W. (2001). The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference? *American Educational Research Journal*, 38, 611– 652.

Within the context of this paper our team has attempted to analyze and critique both the structure of the paper as well as the content of the research conducted.

Introduction

In “The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference?” (Hickey, Moore and Pellegrino, 2001) the authors examined the influence of a video-disc series focused on mathematical problem-solving, “The Adventures of Jasper Woodbury” (Jasper), in combination with a district wide, constructivist curriculum reform and their influence on the academic motivation of fifth-grade students in 19 of the district’s classrooms. The authors revealed that Jasper represents “one of the few applications of the constructivist principles to technology-based learning, (which) simultaneously (is) adhering to mathematics curriculum reform guidelines” (Hickey et al., 2001).

This research was further justified by the 1997 President’s Committee of Advisors on Science and Technology [PCAST] (as cited in Hickey et al., 2001) and the 1989 and 1991 sessions of the National Council of Teachers of Mathematics [NCTM] (as cited in Hickey et al.). By taking a quasi-experimental approach to this study, the researchers were able to “... not simply attempt to show better outcomes on some dependent measure when students used Jasper materials,” but rather to employ “the statistical power and unique aspects presented in one district’s implementation to study several theoretically derived assumptions about the specific consequences of the Jasper activities, as well as the broader constructivist inspired curricular reforms in which they were implemented” (Hickey et al.). Based on an analysis of the present

research the positive outcome calls for additional research in the field as well as further review of curriculum format for implementation in the classroom.

Article Assessment and Presentation of Goals:

Early in the review of the article readers are introduced to the present problems, the research on which this study was based, and details as to why each aspect was being analyzed. The article outline is clear, hypothesis to method, conclusions to further research possibilities. Although originally difficult to follow by a novice reader because of the multiple references to external data, the authors presented a clear case. Woven throughout the writing are suggestions for future and responses to prior research.

The article covers a vast subject area in the field of mathematics curriculum reform, but the authors are quick to point out that the purpose of this work is not just to test out innovative tools and reforms, but more importantly to examine and expand upon the differences they make on a student's motivational and academic experience in mathematics. The authors reviewed prior research, largely primary sources, in these areas in order to present information which would supplement prior findings. However, nowhere in this current paper do the authors specifically point out deficiencies in these sources which would call for an updated study. Update is a key word in the previous sentence; for although the research is thorough, based on the references cited, it seems outdated, many sources published six years prior to the current study, in comparison to the technologies and reforms under review. Citing contemporary research would further validate the current study.

It is interesting to point out that little work has been done in the area of constructivist reform; taking into consideration the success of the Jasper exercises in prior studies provides the base in which to expand the evaluation. So, although the prior research relates to the current study, the span of this study's implications goes beyond Jasper and into constructivist reform innovations,

focused on the motivations and achievements of each student, for use in the general mathematics classroom. The following is a review of how the current study was formulated and subsequently completed.

Problem under Review and Hypothesis:

Based on the complex, yet non-declarative hypotheses, the study assumes that: a) the actual subjective motivation of students participating in Jasper activities is greater than that of students who participate in traditional mathematical teaching and that this difference can be attributed to the interesting subject matter of the jasper activities which reduces the expectations of working with minimum effort; b) classroom curriculum which is more consistent with the constructivist reforms leads to a stronger goal orientation; c) only minor gains on standardized mathematical achievement testing are attributable to the Jasper activities, while significant gains in the areas mathematical problem-solving ability and conceptual mathematical knowledge are obtainable when the Jasper activities are utilized in an environment more consistent with the broader constructivist reforms at a particular school; d) these gains are independent of previous academic achievement or socio-economic status of the students involved; e) large-scale implementation of the Jasper activities, and other similar innovations, is feasible and effective when paired with similar curriculum reform. (Hickey, et al., 2001)

In order to test the hypotheses the authors identify three independent variables for review: Instructional group, Socio-Economic Status (SES), and Reform orientation, which are further defined in the coming paragraphs. Each of these variables is studied in relation to the following dependent variables: Subjective motivational experience, the emotional or motivational state of the students while engaged in Jasper or Non-Jasper mathematical activities; Motivational beliefs, the student's motivation and disposition while utilizing the Jasper materials; and Mathematical achievements, based on the achievement scores. To test the hypotheses these variables were

examined within a representative sample of mathematics classrooms.

Housed in a major metropolitan area of the southeastern United States is a school district made up of approximately 40, kindergarten to fifth-grade, elementary schools. Each year a certain amount of schools were chosen by the district for the Jasper material upgrade. This upgrade was mandated by the district, not voluntary. District administrators assisted in the preparation for this study by identifying for the researchers two pairs of schools for evaluation. Each pair of schools was closely matched based on the independent variables, with a “Jasper” school set to utilize the Jasper materials in the coming school year, which is when the study was to be administered, and a “non-Jasper” school scheduled for Jasper implementation the following year after the study was completed. Overall there were 19 fifth-grade classrooms studied, 10 Jasper classrooms and 9 non-Jasper classrooms. Therefore, based on the details listed above, the sample is finite and utilized the quota sampling method.

In figure one (Hickey, et al., 2001) provided within the article the authors illustrate the breakdown, based upon the independent variables. Defining the Instructional Group, whether or not the school used the Jasper materials, was the simplest of the three variable descriptions. In order to identify the SES groups the researchers further investigated the schools, and separated them into either high-SES (fewer racial minorities, reduced-price lunches, student transfers) or low-SES (many racial minority students, reduced-price lunches, student transfers) schools. Also notable in the SES segments is math achievement. In general, the high-SES schools registered high mathematical achievement scores, where as the mean mathematical achievement for low-SES schools remained below average.

In order to attempt control over the stated variables and avoid extraneous factors, researchers conducted informal interviews with the district mathematics coordinator as well as the Instructional Resource Teacher (IRT) at each of the schools involved in the study to qualify the

final variable, Reform Orientation. It seems this categorization was necessary for the third variable in order to prevent a confounding by the other two. In order to define the variable segments researchers conducted surveys with the IRTs in which they were asked to rate each of their fifth-grade classrooms based on its consistency with the curriculum reforms implemented by the district. To qualify the responses, the IRTs were also asked to assess their familiarity with the reforms and their confidence level in rating the classroom environment. Based on the IRT surveys the classroom was categorized as either “less-consistent” or “more-consistent” with the NCTM curriculum reforms.

Surveys were also utilized in this study in order to measure the dependent variables. For Motivational Experiences measurement the researchers applied the Motivational Experiences Survey (MES) three times throughout the year. Students completed the survey upon the conclusion of an exercise or learning activity. This 21-item survey follows the “on-line” motivation assessment method which was advanced by Boekaerts (Boekaerts, 2003) and assesses task appraisals and task-specific motivation orientation while involved with the specific mathematics exercises. A total of 397 students in 18 of the classrooms completed at least one of MES. More specifically 46 non-Jasper students and 116 Jasper students completed all three MES. In order to create a balanced design “scale scores” or mean scores were input for students who completed two of the three surveys. Because more than 30% of the students did not complete two of the three surveys and nearly 50% missed one, there is a possibility for sample bias or a sample error. In order to assess Motivational Beliefs students completed the Motivational Beliefs Survey (MBS), once at the beginning of the school year and once at the end. This 50-item survey was the result of various resources, including published survey instruments, prior Jasper materials and refinements made during pilot studies. In total 331 students completed both MBS, 183 were Jasper students and 148 were non-Jasper. This survey carried a high reliability measure for each

scale. In both the MES and MBS the final items were derived using a factor analytic method, then the authors conducted a four way variance analysis with a within subject comparison (instruction type and motivational changes through time) and a between-group comparison (socioeconomic status and consistency of reform orientation).

Also important for review was the individual mathematics achievement of each student, Jasper and non-Jasper. For this review the researchers utilized percentile achievement scores from the Iowa Test of Basic Skills (ITBS). The ITBS, like “typical standardized math achievement tests emphasize speeded recall of arithmetic facts to the neglect of the higher-level mathematical understanding targeted by the Jasper activities and by the NCTM standards” (The University of Iowa, n.d.). Form K of the ITBS, was introduced “in response to the NCTM's call for continuing attention to problem solving in the mathematics curriculum with additional emphasis on the interpretation and analysis of data”. (as cited in Hickey. et al.) Researchers compared the scores from the ITBS (Form K) which was administered two-weeks after the final MBS and the ITBS scores received by the same students after completion of their third-grade year. The school district provided the fifth-grade scores for each student in the current study and 293 scores from the third-grade exam. Using the means and standard deviations for percentile scores obtained, the researchers then applied a four-way analysis of variance for within subjects (third grade versus fifth grade score) and between group comparisons (reform orientation, instructional group and socioeconomic status).

Evaluation and Construct Validity:

The detail provided by the authors regarding the use of the study evaluations was minimal because the MES, also known as the “Online Motivation Survey”, and ITBS are widely utilized in the field of educational research. In order to support or negate the construct validity, and in general, the accuracy and consistency of instruments used by the authors, this team chose to

broaden the search beyond the references provided and utilize the power of the Internet.

The ITBS is a well known and widely used test throughout the United States. Created at Iowa State University, the ITBS is a voluntary, non profit cooperative testing program for Grades K through 8. Nationwide, in the U.S., schools access the ITBS through the Educational Testing Service (ETS). Important to note, the ITBS is the only measuring instrument employed in this study which depends on “external” reference criteria. Although an objective approach of basic-skills, this test has been criticized multiple times for its reliance on the "drill and practice" teaching method, rather than the constructivist reform currently being advocated. So, although standardized examinations have their faults, as mentioned throughout the article, the continual revisions and appraisals made by ITP and the ETS support the idea that this exam is reliable.

For the review of the Motivational Experience Survey, the home page of Monique Boekaerts proved useful (Monique Boekaerts), providing general information and a description of this test. Through further searching we encountered a book called “New Directions in Measures and Methods”, edited by Pintrich and Maehr, on the Elsevier website. This book provided the tools utilized in the creation of the “On-Line Motivation Questionnaire: A Self-Report Instrument to Assess Students' Context Sensitivity” published by Dr. Monique Boekaerts and forming the basis of the MES. Based on the broad availability and usage of these resources it can be concluded that this survey is also reliable.

The only exception of survey reliability is the Motivational Beliefs Survey, which was developed by the authors several years before the present study. Because this survey was created using a combination of similar surveys there is no direct evidence of broader usage. And although the researchers continued to refine the survey during a pilot study there is no analytical tool available for a reader to measure the survey’s reliability.

For practicality reasons the researchers did not utilize random subject selection. A possibility for

bias entered the research because of this independent variable manipulation. Also an opportunity for bias, students who deliberately “falsify” the answers of the MES and the MBS, creating a situation where the participants “tell them (the researchers) what they want to hear,” rather than take the time to evaluate the experience. Similarly, students may misrepresent their answers with the intention of causing chaos, possibly introducing the idea that the instructors have influenced their answer choices. The authors have taken this into account, and proposed serial measurements for each of these two surveys, which at least in theory, should reduce the risk of adopting false answers as true measures of the variable under review. There appears to be no selection-bias, since the sample seems to be well planned, and no score bias, since evaluators seem to be independent from the schools authorities. Researchers in this study utilized a statistically supported format for analysis of the MES and MBS survey results. By means of four-way analysis of variance (ANOVA) each of the survey responses was open to comparison. Different criteria were used for each survey, and the authors provide multiple tables and figures which graphically represent these comparisons.

Validation Analysis and Findings:

Even if the researchers did take into consideration the possible falsification mentioned above, there are many other external factors, some mentioned in the research, that are out of their control. For instance, levels of acceptance of the materials and the curriculum reform on the part of the teachers and students, levels of utilization of the suggested Jasper activity delivery methodology, and possibly more telling, the prior intelligence levels of the students studied. Being a quasi-experimental design, readers of the study realize the researchers, as already mentioned, may not have had total control over the variables. The results presented, therefore, required additional verification and validation, both internally and externally. Internal validity is explored through sophisticated statistical tests like the four way ANOVA. External validity is

pre-qualified by using footnote commentaries regarding the results. These statistical tests appear carefully selected so as to accomplish the kind of analysis that all the variables demand.

Statistical analysis is presented by using the letter “F”, which presents a ratio similar to the variance. It seems the authors used this test in order to evaluate the effect of three independent variables over three dependent variables (Hernández, Fernández, and Baptista, 2003). The “F” test requires the independent variables be expressed in categories, which is exactly what the authors did in this study (low-high, yes/no, less/more). The value F represents the calculated factor, and the parentheses enclose the number of values in the final calculation that are free to vary. The term “p” designates the probability of incurring in a type I error when assuming that the statement is true; usually a “p” lower than 0.05 is necessary to accept the premise (Johnson and Christensen, n.d.). Thus, the statistical results show statistical significance, using the F-test, but some lack practical significance. In reality the authors were making non-quantitative decisions, and thinking about many different factors such as the size of the relationship, whether an intervention would transfer well to the real world and the costs of using a statistically significant intervention in the real world, etc (Johnson and Christensen). For example, the authors mentioned that there was a lower subjective competence in the Jasper group. Although statistically significant, the statement is contradicted by mentioning that the complexity of Jasper activities is responsible for this difference. Another contradiction stated that Jasper activities were found less relevant than comparison activities, yet the Jasper activities actually improved mathematical achievement.

Critical Analysis Conclusions:

A major challenge confronted by the authors was to take a qualitative issue and migrate it to a quantitative one. In fact, one of the resonating complaints of past studies was the integration of results into the real world; this task was diligently attempted through the conceptualization of a

methodology which clearly states its goals and covers the possible variables. It was difficult to qualify the multiple parts of this study, when at times both Jasper and Constructivism were under review. Yet, after analyzing the structure and the content, the constructivist reform represents a broader effect than just introducing a new technological format with the Jasper activities. This finding illustrated that the structure of the document meets all the standard requirements of a quantitative method, which span from the background research for the study to the experiment results and conclusions. Step one in the analysis was complete. However, upon evaluation of the content it is important to highlight some conclusions. First, the limitations of the study mainly overview methodological problems; the deficient implementation of Jasper is something that was not thought of completely by the researchers and finally was a confounding variable that could not be controlled with their methodological design. Second, the main conclusion of the study was not confronted with pros or cons from other researchers (external validity). And if the authors believed that mathematical achievement was better in a broader constructivist reform, the concept was not made apparent as such and the conclusion should be supported by other authors or other references about where this idea is defined should be cited. Third, the authors stated the importance of their findings about the larger gains in math achievement for American children with weak mathematical and computational skills, but there is no clear definition of this statement on this study. So, we may conclude that, as an early study in the field, the results and procedures show very impressive results given the problematic variables. The effort is valid although not soundly acceptable as a premise since we may not compare it with others and more research is needed to validate the objectivity of this research.

References

- Boekaerts, M. (2003). Monique Boekaerts Web Site. Retrieved March 6, 2005, from <http://www.boekaerts.com/>
- Educational Testing Service (2004). ETS at a Glance: Listening. Learning. Leading. Retrieved March 6, 2005, from <http://www.ets.org/aboutets/index.html>
- Hernández, R., Fernández, C. & Baptista, P. (2002). Metodología de la investigación (3a. ed.). Distrito Federal, México: McGraw–Hill.
- Hickey, D. T., Moore, A. L. & Pellegrino, J. W. (2001). The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference? *American Educational Research Journal*, 38, 611–652.
- Johnson, B & Christensen, L. (n.d.). Educational Research: Quantitative, Qualitative, and Mixed Approaches. Retrieved March 5, 2005, from http://www.southalabama.edu/coe/bset/johnson/dr_johnson/lectures/lec16.htm
- National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics. Reston.
- National Council of Teachers of Mathematics (1991). Professional standards for teaching mathematics. Reston.
- Pintrich, P & Maehr, M (Eds.). (n.d.). New Directions in Measures and Methods, 12 [Contents]. Retrieved March 6, 2005, from http://www.elsevier.com/wps/find/bookdescription.cws_home/622452/description#description
- President’s Committee of Advisors on Science and Technology, Panel on Educational Technology (March, 1977). Report to the president on the use of technology to strengthen K-12 education in the United States.

Riverside Publishing (1994). Iowa Test of Basic Skills technical summary I. Riverside 2000.
Chicago, IL.

The University of Iowa (n.d.). ITBS - Iowa Testing Program - The Iowa Test of Basic Skills.

Retrieved March 6, 2005, from <http://www.education.uiowa.edu/itp/itbs/>