The Motivational and Academic Consequences of Elementary Mathematics Environments: Do Constructivist Innovations and Reforms Make a Difference?

Daniel T. Hickey
University of Georgia
Allison L. Moore
Vanderbilt University
James W. Pellegrino
University of Illinois at Chicago

This study examined the effects of a videodisc-based mathematical problem-solving series known as The Adventures of Jasper Woodbury, as implemented by one school district within a constructivist-inspired reform of its math curricula. The motivational and academic consequences of both the specific innovation and the broader reforms were examined in 19 fifth-grade classrooms in two pairs of closely matched schools. One pair of schools served higher-achieving high-socioeconomic status (SES) students while the other pair served relatively lower-achieving low-SES students. Significantly larger gains on the Mathematical Problem-solving subtest of the ITBS were documented in the 10 classrooms where the Jasper activities were implemented, and in the 10 classrooms that were ranked as relatively more consistent with the broader curricular reform goals. The largest relative gains were found in the five classrooms that both used the Jasper activities and were ranked more consistent with the broader reforms. The positive conse-
quences of both the Jasper activities and the broader reforms were documented in both pairs of schools. The implications of these results are discussed relative to current proposals for curricular reform and research on educational innovations.

A recent report to the President's Committee of Advisors on Science and Technology (PCAST, 1997) advocated dramatically increasing the proportion of K–12 spending devoted to computer and information technologies. The report also urged refocusing the use of classroom technology away from teaching about technology itself, and toward achieving the academic content goals in existing educational reform plans. The report expressed cautious optimism that contemporary constructivist models of learning and teaching will ensure that such an investment yields worthwhile returns in educational outcomes. As most recently summarized in a report by the National Research Council, entitled How People Learn (Bransford, Brown, & Cocking, 1999), these models focus on higher-order skills, meaningful problem-solving activities, and the active construction of knowledge. These models stand in contrast to more conventional behaviorist and associationist models that emphasize the assimilation of isolated, factual knowledge, as embodied in conventional drill-and-practice applications of educational technology.

Reflecting the immaturity of constructivist practices and opposition from some quarters, the PCAST report recommended sharply increasing the amount of research conducted on constructivist applications of technology and broadening such research beyond a focus on formative questions and interpretive methods. While acknowledging the essential role of "early-stage" research, the report calls for more research that "exposes well-explicated falsifiable hypotheses... to potential refutation through the execution of well-designed, carefully controlled experiments having sufficient statistical power to distinguish genuine effects of a relatively modest size from differences that can easily be explained as chance occurrences" (p. 94).

The report further advocates such research "be conducted under conditions more typical of actual classrooms, using ordinary teachers, and without access to unusual financial or other resources, for example, or to special outside support from university researchers" (p. 95). This paper describes an initial study of one technology-supported constructivist innovation carried out in the spirit of the goals spelled out in the PCAST report. We highlight the challenges such efforts face, as well as the possibilities when innovation, coordination, and preparation coincide.

Our study focused on a set of videodisc-based mathematical problem-solving materials known as The Adventures of Jasper Woodbury (hereafter referred to as "Jasper") (Cognition & Technology Group at Vanderbilt, CTGV, 1991, 1992, 1997). These commercially distributed materials are noteworthy because they represent one of the few applications of constructivist...
principles to technology-based learning, simultaneously adhering to mathematics curriculum reform guidelines advanced by the National Council of Teachers of Mathematics (NCTM, 1989, 1991). In addition, the Jasper materials are now being used in numerous classrooms nationwide.

Within calls for rigorous empirical field-based studies of educational technology, it is worthwhile to note the shift beyond the earlier focus on merely demonstrating that new technology can be used to teach as effectively or more effectively than conventional methods. The probability that elementary and secondary education will prove to be the one information-based industry in which computer technology does not have a natural role would at this point be appear to be so low as to render unconscionably wasteful any research that might be designed to answer this question alone (PCAST, 1977, pp. 93–94).

In other words, the challenge of researching innovative applications of K–12 technology are far too complex to be distilled into simple calculation of the innovation’s “effect size.” Rather, research is needed to test theoretically driven assumptions about the consequences of this fundamentally new class of instructional innovations. A companion report to the NRC’s How People Learn report, entitled Bridging Research to Practice (Donovan, Bransford, & Pellegrino, 1999), argues that studying effective instructional techniques is analogous to asking which tool is best suited for accomplishing specific outcomes. Thus, we did not simply attempt to show better outcomes on some dependent measure when students used Jasper materials. Rather, this study employed 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.

The Jasper Materials, Prior Research, and Study Overview

The Jasper materials were developed according to a set of theory-based design principles (described in more detail in CTGV, 1992, 1997) and incorporate a broader approach known as anchored instruction (CTGV, 1990). Each of the “adventures” centers on a short (approximately 17 min) video and was designed to support fifth- to eighth-grade students’ collaborative problem solving across four or five class periods. In each adventure, the main character confronts a compelling problem that calls for a mathematical solution. Each problem is quite complex, requiring 15–20 steps to solve. The materials were designed to be consistent with the standards advanced by the NCTM and other educational reform advocates and organizations. One examination of available mathematical curriculum products concluded that the Jasper materials were the one product “that most fully incorporates the ideas expressed in the NCTM standards, including mathematical sophistication, stress on group work, and real-world relevance” (Eisler, 1993, p. 58).

As described in CTGV (1997), the relatively unique combination of features in the Jasper materials include (a) the use of authentic real-world
narrative contexts, (b) the integration of mathematical concepts within a problem and the integration of mathematics and other academic subject areas, (c) the use of video and other interactive technologies, (d) the explicit modeling of real-world inquiry relevant to school mathematics, (e) the need to generate the problems and identify relevant data prior to actually constructing a solution, (f) the support of extended collaborative problem solving across multiple days and multiple activities, (g) the opportunity to develop even deeper understanding of mathematical concepts through various extensions and analogs, and (h) the provision of multiple positive role models.

The Jasper series consists of 12 adventures, plus a variety of support materials. The series includes three increasingly complex adventures in four problem areas (distance/rate/time, statistics and probability, geometry, and algebra). Although originally developed for videodisc delivery, the series is also being made available on CD-ROM. The present research was carried out while the series was still in production, and considered three of the initial adventures. The first two are time/distance/rate problems that involve trip planning. In Journey to Cedar Creek, Jasper decides to buy an old boat with no running lights and a temporary fuel tank, and students must determine if Jasper can make it home before dark without running out of fuel. In Rescue at Boone's Meadow, Jasper encounters a wounded eagle and students must consider a range of alternatives for getting the eagle to the veterinarian. The third adventure, The Big Splash, is the first of the three statistics and probability adventures, and requires helping one of Jasper's young friends develop and evaluate a business plan involving a "dunk tank" to raise money for a student-run TV station.

Prior Jasper Evaluation Research

Prior research of the Jasper materials is summarized in CTGV (1997). This includes a wide range of formative investigations, as well as several implementation studies carried out in the "design experiment" tradition (e.g., Brown, 1992). These latter studies generally used teachers who were already affiliated with the research project and who were provided with extensive technical, curricular, and instructional support throughout the implementation. Although the design experiments yielded some impressive outcomes, inherent research design and implementation issues qualify the generalizability of these findings to traditional research-driven implementations as well as school-initiated and district-initiated implementations.

One large-scale multistate study of the Jasper materials utilized selected teachers who were offered more modest (i.e., 2 weeks) training and less support throughout the implementation, and compared gains in implementation classrooms to gains in matched non-Jasper classrooms (CTGV, 1992; Pellegrino, Hickey, Heath, Rewey, & Vye, 1992). Encouragingly, students in the classrooms that used the Jasper materials showed slightly larger gains on traditional math achievement tests (despite the reduced time for traditional
math instruction), significantly enhanced performance on more complex multistep word problems and problem-solving tasks, and significantly enhanced attitudes toward mathematics.

Goals of the Present Research

The present research extends the previous studies of the Jasper learning environment (and such innovations in general) by (a) considering students' subjective motivational experiences, (b) studying a large-scale implementation that was initiated and carried out by the school system, (c) using newer, ostensibly more appropriate standardized achievement measures, (d) comparing consequences in classrooms that are more consistent and less consistent with the broader curricular reforms, and (e) comparing consequences in higher-achieving, high-socioeconomic status (SES) classrooms and lower-achieving, low-SES classrooms. Following is an overview of the how this study addressed each of these goals.

The study was initiated primarily to explore the consequences of the Jasper activities for students' subjective motivational experiences. This line of inquiry roughly follows from Csikszentmihalyi's pioneering studies of "flow" experiences and the associated experience sampling methods (e.g., 1990), as well as subsequent considerations of situated motivation (Paris & Turner, 1994) and situational interest (Hidi & Anderson, 1992). Whereas the prior studies focused on the longer-term consequences of the Jasper activities for students' more general motivational beliefs, the present study also examined students' subjective motivational experiences while they were actually participating in the activities.

Regarding the context in which the implementation was initiated, the present study examined a district-initiated implementation where all teachers in the implementation schools were trained by district personnel and required to implement the materials in their classrooms in an ostensibly similar fashion. Because teachers in the present study received relatively typical implementation support, and because the research design minimized the likelihood that the outcomes could be attributed to the types of teachers who choose (or were chosen) to use the materials, the present study's findings more readily generalize to other schools that independently purchase and use these materials.

Although the initial goals of the study included assessing consequences of the Jasper activities on standardized measures of math achievement, we anticipated finding very modest gains attributable to the activities. The PCAST report reiterates well-known concerns that standardized achievement tests were designed to measure students' prior acquisition of low-level skills and factual knowledge; as such they overlook the higher-level conceptual understanding and problem-solving ability that are the focus of many constructivist innovations and reforms, including the Jasper materials and the NCTM Standards. As with the prior large scales studies of Jasper, we initially assumed that standardized achievement tests would be relatively insensitive to conceptual, reasoning, and problem-solving abilities targeted by the Jas-
per activities, and wanted to ensure that diverting math class time to the activities did not lead to lower performance on the standardized tests. Serendipitously, the district in the present study was using a popular achievement test (ITBS) that had just been revised (Form K), ostensibly in response to such concerns. Thus, we were presented with one of the first opportunities to examine the academic consequences of these innovative practices using a standardized test that included specific subtests purported to be more sensitive to the reasoning abilities targeted by those practices.

Regarding the use of the Jasper activities in different types of classrooms, the present study advanced consideration of the tensions between instructional innovations such as the Jasper materials and conventional "teacher-directed" models of educational practice (e.g., Bransford, Goldman, & Vye, 1991). Prior studies revealed a wide disparity in the way that teachers used the Jasper materials (e.g., CTGV, 1992). This raises the concern that more conventional classroom environments may alter innovative activities to fit within that tradition, potentially limiting their effectiveness, or worse. Early recognition of this conflict led to calls for a more coordinated pursuit of new instructional technologies within more fundamental reforms of educational practice—what Sheingold (1991) called "restructuring for learning with technology" (see CTGV, 1996). Opportunely, the school district in the present study implemented the Jasper materials as part of a broader mathematics curriculum reform that was inspired by the principles advanced by the National Council of Teachers of Mathematics—principles that are entirely consistent with the goals of the Jasper materials. In the course of the study it became apparent that both the Jasper classrooms and the non-Jasper classrooms varied widely in the degree to which overall mathematical practices were consistent with the broader curricular reform goals. Thus, it was possible to conduct post hoc comparisons that examined the independent and combined consequences of both the Jasper activities and the broader reforms.

One other way that the present study advanced prior research concerns the relative impact of constructivist innovations and reforms on disadvantaged students who are presumed to be at greater risk for academic failure (e.g., Herman, 1994; Means & Olson, 1994; PCAST, 1997). Skeptics (e.g., Dick 1991) have expressed concern that constructivist learning environments are too complex for academically disadvantaged students. Such concerns are buttressed by conventional empiricist theories of learning that assume learners must first develop arithmetic skills and basic problem-solving skills before they can learn the more complex problem-solving skills targeted by the Jasper materials and the broader NCTM reforms. If this is true, such innovations and reforms should "make the rich richer," by increasing the gap between low-achieving students and their more able counterparts. However, proponents of constructivist approaches counter that both basic skills and higher-level understanding are best developed within the context of more complex and more meaningful problem-solving activities, and that traditional "drill and practice" remedial approaches place at-risk students
further and further behind (see, for example, Means, Chelmer, & Knapp, 1991). A related concern is that less-able students will find constructivist environments exceedingly frustrating, leading to withdrawal and disengagement (see Dick, 1991). However, McCombs and Pope's (1994) discussion of research on motivating "hard to reach" students concluded that, in addition to supportive teachers and classmates, the learning environment for such students "needs to include instructional practices that give students real experience in how to use their minds and how to take personal control over their thought processes" (p. 16). The Jasper materials and NCTM-inspired math curricula are presumed to afford just such experiences. Furthermore, the use of video in the Jasper materials affords a shared problem-solving context that is expected to support motivation and engagement among students who might otherwise be frustrated by the literacy demands of print-based problem-solving activities. The present study was conducted in two pairs of closely matched schools. One pair served almost exclusively racial-minority students from a relatively low-SES community, whereas the other pair served relatively advantaged students from an affluent suburban community. In terms of standardized measures of math achievement, mean performance of the pairs was separated by more than 40 percentile points. Thus, it was possible to explore the relative consequences of both the Jasper activities and the broader reforms in two remarkably different populations of students within the same school district and ostensibly employing the same curricular framework.

Method

The number of independent and dependent variables ultimately presented a quasi-experimental design that was powerful, but complex. The study examined the main effects and interactive effects of three independent variables (instructional group, SES, and reform orientation) on three sets of dependent variables (subjective motivational experience, motivational beliefs, and mathematical achievement). Following is an explanation of the context in which the study was conducted, a description of each set of variables, and details regarding the sequences in which the instruments were administered.

District and School Context

The study was conducted in a school district in a major metropolitan area in the southeastern United States. The district's roughly 40 K–5 elementary schools serve an extended, ethnically and economically diverse suburban area. Although many of the schools served predominately African American and/or economically disadvantaged communities, few, if any, of the communities could be characterized as "inner city."

Implementation of the Jasper materials. After participating in the earlier multi-state Jasper implementation (described in CTGV, 1992), the district began independently implementing a subset of the Jasper materials in their
fifth-grade classrooms on a school-by-school basis. Two teachers who participated in the original implementation were responsible for training all fifth-grade teachers in participating schools to use the first three Jasper adventures (Journey to Cedar Creek, Rescue at Boone’s Meadow, and The Big Splash). Reportedly, these two teachers worked closely with the district math coordinator and the Instructional Resource Teacher (IRT) at each school. The two teachers conducted daylong workshops at each of the schools that had been scheduled to implement the Jasper materials the coming academic year. These workshops reportedly included demonstrations of the technology, discussions of how the activities were best implemented, and a description of the theory behind the activities. Once these new teachers were trained and the software and hardware made available, teachers then included the three adventures at regular intervals throughout the school year, devoting three or four class periods to each.

Teachers were advised to follow the sample lesson plan that recommended having students spend one period viewing the video and brainstorming solutions as a class, then one or two periods in small collaborative groups generating subgoals and reviewing the video for pertinent data while constructing a plausible solution. Each group was to then present their solution to the class before viewing a possible solution that is included with the video. Reportedly, to varying degrees, classrooms continued to discuss the adventures during other class periods, and some teachers used the activities to illustrate other points throughout the year.

Mathematics curriculum. Starting 6 years before the year that the present study was conducted, the school district began revising its elementary mathematics curricula according to the Curriculum and Evaluation Standards developed by the National Council of Teachers of Mathematics (1989). The scope and sequence for the existing curriculum was rewritten, new curricular materials were obtained or developed, and teachers were provided with extended professional development opportunities. The introduction to the new curriculum guide captured the spirit of the revised curriculum and highlighted its expected impact on achievement and motivation: “Providing opportunities for students to explore concepts, to take risks, and to test a variety of problem-solving strategies will allow for the development of higher order thinking skills and will help students build confidence in their mathematical abilities.” The curriculum guide specified a variety of activities directed at detailed performance goals that were consistent with the NCTM framework. Reflecting a central thrust of the NCTM standards, the revisions curtailed drill and practice activities targeting simple arithmetic operations, while increasing the emphasis on estimation, problem solving, and general mathematical concepts.

The revised curriculum was organized around 18 levels of proficiency. Each level was made up of specific objectives linked to the state curricular standards. Levels 16, 17, and 18 that made up the fifth-grade curricula consisted of 33, 22, and 19 objectives, respectively. For each objective, a sequence of learning activities was specified, and teachers were instructed to
follow the specified sequence for each objective and to complete the objectives in order. The sequence of activities is explicitly developmental, reflecting the assumption that "development of concepts frequently begins with concrete models, continues with the use of pictures, and concludes with more symbolic representations." Various overheads, student worksheets, and manipulables were integrated into each unit. Mathematical discussion was emphasized throughout and explicitly supported with discussion prompts and "cartoon bubble" dialog samples. A standard math text from a major publisher was included, but only to provide additional problem sets and practice within the curricular framework. Classes were organized into groups of 4-5 students of similar ability levels; the curricular guide provided explicit directions and worksheets for activities to be carried out in these groups, as well as other activities that students were expected to complete independently or as a whole class. For example, the 31st objective in Level 16 was uses mental computation strategies such as multiples of ten, hundred, or thousand, with division of whole numbers. The curriculum guide suggests that teachers initially "review the idea of computing mentally." One of the dialog samples provided to facilitate this review featured a character facing a chalkboard with the problem "180 ÷ 3 = ?". The character is shown to be thinking aloud "18 ÷ 3 = 6. So ... 180 ÷ 3 = 60. Or ... 18 tens ÷ 3 = 6 tens = 60". After the initial discussion, teachers are directed to then have students work in small groups to play a competitive game called "Quotient Estimation" where teams within each group use mental estimation to solve problems provided on index cards. Subsequent activities include teacher-led practice where students mentally compute the answers to problems presented on an overhead at a pace designed to require estimation, accompanied by dialog designed to help students understand notions such as the existence of a variety of appropriate estimation strategies.

Reflecting the model of assessment advanced by the NCTM, the curriculum guide specified extensive formative and summative assessment designed to "improve instruction by distinguishing between teaching and learning," "carry forth assessment coincident with instruction," and "communicate the learning activities and outcomes that are valued." Within each level of the curriculum, sets of objectives were interspersed with open-ended performance assessments to be used for "diagnosis of students' strengths and weaknesses, instructional planning, pacing of instruction, and for placement of students." For example, Objective 31 was clustered with four other objectives: divides whole numbers using one digit divisors, divides whole numbers using two digit divisors, uses estimation strategies such as front-end rounding and compatible numbers to predict results of division with whole numbers, and divides whole numbers greater than three-digit by two-digit using a calculator. The group-administered performance assessment for this cluster included a 10-item student answer sheet and detailed administration instructions. For the first problem, students were instructed to mentally calculate the answer to a typical division word problem displayed on the teacher's overhead. The teacher then led a discussion where students were asked
to describe the strategy they used, emphasizing the need to articulate the strategy used and the range of appropriate strategies. Several subsequent problems were presented and students were again asked to write down their answer and the strategy they used to reach that answer. The last four problems to be solved were presented on the actual answer sheet, and students were allowed to use a calculator for the final, more difficult, problems. For this particular performance assessment, students were expected to be able to provide satisfactory answers to 7 of the 10 problems before proceeding. When students failed to demonstrate the specified proficiency, teachers were directed to reteach the content and reassess using problems with changed context or setting, reversed given and wanted information, etc.

Each curriculum level concluded with a 1–2 period cumulative assessment that is “multi-dimensional and allows students to demonstrate their ability to use concepts and skills in the context of real world applications” and “provides a forum for celebrating their accomplishments in mathematics.” In the case of Level 16, this included individually completed problems that were evaluated primarily in terms of students’ ability to explain how they arrived at their answer, as well as group-administered assessments where students are awarded both individual and group scores for the quality of both their response and their collaboration. Problems included more basic computational problems as well as a single multistep “story” problem. The latter problem posed the challenge of collecting pennies to raise $10,000 for a swimming pool. Students collaboratively solved problems such as how long it would take if every student in the school collected 1000 pennies per month, the number of rolls of pennies in $10,000, and the size of the container needed to hold that many rolls.

Relationship between Jasper and the math curricula. The IRT at each school (who had full-time support responsibilities and no class assignment) was responsible for helping teachers incorporate the materials and methods associated with the curricular reforms. As schools adopted the Jasper materials, the IRTs also assisted in their implementation. The implementation of the Jasper materials and the incorporation of the broader reforms were motivated by the same goals. However, the implementation of the Jasper materials was apparently independent of a particular teacher or school’s adoption of the broader reforms. Specifically, the selection of school to implement the Jasper materials in a particular year was reportedly independent of schools’ progress implementing the broader curricular reforms. All fifth-grade teachers at the school implemented the Jasper activities when the particular school was directed to do so; the two comparison schools in the present study were scheduled to implement the Jasper materials the following year, as funds for the necessary hardware became available.

Independent Variables

Figure 1 displays how 19 fifth-grade classrooms from four different schools were configured around the three independent variables employed. 

620
in the study's design. Following is a description of each of the independent variables.

*Instructional group.* Because the school-by-school implementation of the Jasper materials was underway at the time of the study, district administrators were able to identify two closely matched pairs of schools in which only one school was to use the Jasper materials in the coming school year. The “Jasper” school in each pair began using the materials 1 year earlier, whereas teachers in the “non-Jasper” schools were scheduled to begin using them the year after the study was completed. The implementation's progression was directed at the district level, and there was no indication that a particular school's implementation was influenced by factors such as its readiness or desire to use the materials.

*SES.* A consequence of the county's geography was that the schools at one end of the county largely served a very different population of students than schools in the other end of the county, in terms of race, SES, and prior achievement. As shown in Table 1, the students within the two matched pairs of schools were quite similar, but the two pairs of schools were quite different. The population in School A and School B included less than 20% racial minorities, almost no students received free or reduced price lunches, and relatively few students transferred in or out of the school during the year. Mathematical achievement in Schools A and B was similar and quite high. In contrast, nearly every student in School C and School D was a member of a racial minority, more than half qualified for free or reduced-price lunch, and
Table 1
Demographic Data for Students at Participating Schools

<table>
<thead>
<tr>
<th>School (SES/group)</th>
<th>Race&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Other demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black</td>
<td>White</td>
</tr>
<tr>
<td>A (High/Jasper)</td>
<td>5</td>
<td>88</td>
</tr>
<tr>
<td>B (High/non-Jasper)</td>
<td>5</td>
<td>82</td>
</tr>
<tr>
<td>C (Low/Jasper)</td>
<td>80</td>
<td>1</td>
</tr>
<tr>
<td>D (Low/non-Jasper)</td>
<td>97</td>
<td>1</td>
</tr>
</tbody>
</table>

<sup>a</sup>Percentage of student population, according to parental report.

<sup>b</sup>Includes Native American, Alaskan, and Multiracial.

<sup>c</sup>Percentage of students qualified for free or reduced-price lunch.

<sup>d</sup>Percentage of students transferring into or out of the school during the school year.

<sup>e</sup>School-wide mean for fifth graders only on combined mathematical subtests of the ITBS.
more than one-third of the population transferred into or out of the school during the school year. Mean mathematical achievement in both low-SES schools was relatively low, with School C (Jasper) somewhat lower than School D (non-Jasper). As shown in Figure 1, crossing this variable with the instructional group variable placed one school with 4, 5, or 6 classrooms in each cell of the 2 x 2 design.

Reform orientation. Conversations with the district mathematics coordinator and IRTs at each of the four schools while the study was under way suggested a wide range of adherence to the broader curricular reforms in the 19 participating classrooms. The IRTs reported that some teachers had readily incorporated the reforms into their math practices, while others continued in varying degrees to rely on the materials and methods used in previous years. Additional inquiries suggested that the IRTs were quite cognizant of this aspect of the classrooms in their schools, and that a full range of reform orientation was present in the fifth-grade classrooms in all four of the schools. As such, we explored the possibility of incorporating this dimension as described next.

Near the end of the school year, a questionnaire was mailed to the four IRTs asking them to rate each of their school’s fifth-grade classrooms’ consistency with the first four NCTM curricular standards on a 1–10 scale. Two other questions asked each of the IRTs to appraise their familiarity with the NCTM curricular standards and confidence in their knowledge about their rating of the individual classroom environments. Each IRT then participated in a 1- to 2-hr phone interview where they were asked to further justify their survey responses and to answer open-ended questions about the classroom environments. Six of the 10 Jasper teachers participated in a similarly structured survey/interview. A modest honorarium was provided to the IRTs and classroom teachers for their participation.

The IRT’s scores for each classroom on the four classroom environment items were averaged, yielding an overall score that was presumed to represent the degree to which the classroom was consistent with the model of practice advanced by the NCTM. A median split was used to categorize classrooms within each school as either “less-consistent” or “more-consistent” with the NCTM’s model of practice. This within-school categorization was needed to prevent this third independent variable from being confounded by the two other independent variables. This was necessary to maintain a balanced design that would reveal the independent and interactive effects of all three independent variables.

At School A, the ranking procedure yielded two classrooms with the median score; the additional comments from the IRT were used to identify which of the two classrooms appeared more consistent than the other. Table 2 presents the average scores for each of the 19 classrooms and shows how each classroom was ranked. As shown in the bottom row of Table 2, the mean of the average scores for the 9 less-consistent classrooms was 4.8, compared to 8.05 for the 10 more-consistent classrooms, on a 10-point scale. This difference, along with the IRTs’ levels of confidence in their ranking,
Table 2
IRT Ratings of Classroom Environment in Terms of NCTM Standards

<table>
<thead>
<tr>
<th>School (SES/group)</th>
<th>Ranked &quot;less consistent&quot;</th>
<th>Ranked &quot;more consistent&quot;</th>
<th>IRT self-appraisals</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (High/Jasper)</td>
<td>5.0, 5.5, 6.5</td>
<td>6.5, 7.75, 8.0</td>
<td>4</td>
</tr>
<tr>
<td>B (High/non-Jasper)</td>
<td>3.5, 6.25</td>
<td>7.0, 8.5, 9.5</td>
<td>4</td>
</tr>
<tr>
<td>C (Low/Jasper)</td>
<td>3.5, 5.0</td>
<td>9.0, 9.0</td>
<td>4</td>
</tr>
<tr>
<td>D (Low/non-Jasper)</td>
<td>4.0, 4.0</td>
<td>7.0, 8.25</td>
<td>5</td>
</tr>
<tr>
<td>Mean</td>
<td>4.8</td>
<td>8.05</td>
<td>4.25</td>
</tr>
</tbody>
</table>

aThese items asked To what degree did students (a) use problem-solving approaches to investigate and understand mathematical content, (b) communicate with others to discuss mathematical ideas and make conjectures and convincing arguments, (c) investigate mathematical connections in a way that let them see mathematics as an integrated whole and connected to other disciplines, and (d) use inductive and deductive reasoning to support mathematical arguments? Scale anchors were 1 = not at all and 10 = completely.
bHow familiar are you with the NCTM Curriculum and Evaluation Standards? Scale anchors were 1 = not at all and 5 = very familiar.
cHow confident do you feel about your knowledge of the study classrooms in terms of the NCTM Standards? Scale anchors were 1 = not at all and 5 = very confident.

was deemed sufficient to justify including the variable reform orientation as a third independent variable.

Further validation of the classroom categorization was provided by the IRTs’ responses to the survey question “how would you characterize this teacher’s practice.” Although it was not explicitly stated, the question’s placement made it obvious that it was concerned with aspects of practice that are addressed by the NCTM standards. Although these responses were included on the same sheet as the numeric rankings, they were not explicitly included during the categorization process described above. IRTs’ descriptions of teachers in the classrooms that were later placed in the more consistent category included the statements “uses the new curriculum very well,” “makes connections between math and other topics,” and “requires students to give lots of justifications.” In contrast, the IRTs’ descriptions of teachers in the less-consistent category included statements like “very tied to the textbook,” “uncomfortable with small-group activities,” “highly structured,” “teacher-directed,” and “little discussion about mathematics.” It is worthwhile to note that two of the four IRTs volunteered that teachers in the less-consistent classrooms were not necessarily less competent. One suggested that these teachers were reluctant to change because they had been so successful using more conventional methods—particularly because success is partly defined by parents’ efforts to have their students enrolled in particular teachers’ classrooms. Conversely, some of the teachers in the more-consistent classrooms were said to still be having difficulty implementing the new curricula.
Our post hoc, within-school categorization presents potentially important issues, particularly concerning the comparability of classrooms in the two groups across the four schools and the interpretation of resultant effects attributable to this variable. However, these concerns are reduced in light of the distribution of scores on this rating variable shown in Table 2. There was a consistent pattern across schools such that all the classrooms labeled less-consistent had mean scores on the rating scale below those of all the classrooms rated more consistent, with the exception of the one tied score in School A. Even so, we choose the labels more-consistent and less-consistent to highlight the subjective and relative nature of this variable. Ideally, classification would be based on a standardized instrument and protocol to observe classroom activity and interview teachers and students about practices in the classroom. However, we believe that the extensive knowledge of the IRTs regarding classroom practice within their respective schools and the convergence of the various sources of information regarding the range of practices in each school, especially the differences between classrooms within schools, seems adequate justification for taking advantage of this opportunity to extend the scope of the study.

Dependent Variables and Measures

Motivational experiences. Reflecting the increased emphasis on situation-specific conceptualizations of motivation (e.g., Hickey, 1997; Paris & Turner, 1994), the study explored students' motivational/emotional states while they were engaged in the Jasper activities and non-Jasper math class activities. Students' motivational experiences were assessed following the "on-line" motivation assessment method advanced by Boekaerts, where individuals complete short self-report surveys during or immediately following a learning activity (e.g., Boekaerts; 1987; Seegers & Boekaerts, 1993). Students repeatedly completed the 21-item Motivational Experiences Survey (see Table 3) to assess task appraisals and task-specific motivation orientation while engaged in both Jasper activities and comparison math activities. This survey was implemented and revised during two pilot studies, and the items were written for the present investigation. The items were worded retrospectively, and the survey instructions asked students to indicate how they felt during the class activity in which they had just participated.

One aspect of motivational experience examined was task appraisals, (i.e., interestingness, perceived relevance, subjective competence). Given that the Jasper activities are more student-directed and situated in compelling contexts, students were expected to appraise the experience as more interesting and more relevant than their other math activities. Expectations for subjective competence were unclear. On the one hand, the Jasper activities are relatively challenging and less well-defined problems, two activity characteristics shown to decrease subjective competence (Schunk, 1991). On the other hand, these activities build on student interests, give students more control over their learning, and emphasize mastery aspects of the task, activity characteristics that are expected to increase subjective competence.
Another aspect of motivational experience, task-specific motivation orientation, examined the strength of different kinds of goals. Collins, Brown, and Newman (1989) argue that "ideal" learning environments direct learners' cognitive activity toward goals that are concerned with gaining knowledge, what Bereiter and Scardamalia (1989) call "personal knowledge building goals." It was expected that students would report stronger orientation toward goals concerning learning while engaged in the Jasper activities and in classrooms that were more consistent with the broader reforms. We also assessed orientation toward extrinsic "performance"-oriented goals. Conventionally, learning orientation and performance orientation have been considered to be dichotomous, where an orientation toward "looking smart" and impressing classmates and teachers was presumed to be maladaptive (e.g., Nicholls, 1989). We also assessed students' orientation toward work-avoidant goals. It was expected that the interestingness of the Jasper activities would diminish students' orientation toward goals concerned with minimizing effort.

Motivational beliefs. A second set of dependent variables concerned students' motivational beliefs and dispositions—the relatively stable individual-difference variables that have traditionally been the focus of classroom motivation research. These variables address two of the NCTM's (1989, p. 6) five "new goals for students:" becoming confident in one's own ability and learning to value mathematics. Prior studies of students in classrooms

### Table 3

<table>
<thead>
<tr>
<th>Scale</th>
<th>No. of items</th>
<th>α</th>
<th>Example item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task appraisals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjective Competence</td>
<td>2</td>
<td>.66</td>
<td>I was able to do really well</td>
</tr>
<tr>
<td>Relevance</td>
<td>2</td>
<td>.66</td>
<td>What I was learning was very important to me</td>
</tr>
<tr>
<td>Interestingness</td>
<td>2</td>
<td>.66</td>
<td>I was learning something interesting to me</td>
</tr>
<tr>
<td>Motivation orientation</td>
<td>6</td>
<td>.86</td>
<td>My main goal was to learn as much as I could</td>
</tr>
<tr>
<td>Performance/ego</td>
<td>5</td>
<td>.80</td>
<td>I wanted others to notice how smart I was</td>
</tr>
<tr>
<td>Work avoidance 1 (&quot;make it easy&quot;)</td>
<td>2</td>
<td>.63</td>
<td>I wanted to make it easy so I would not have to think hard</td>
</tr>
<tr>
<td>Work avoidance 2 (&quot;get it over with&quot;)</td>
<td>2</td>
<td>.45</td>
<td>I tried to do only what I had to do</td>
</tr>
</tbody>
</table>

*α* Internal consistency (Cronbach's alpha) averaged across all six administrations in the 10 Jasper classrooms.

*All items were preceded with the stem Working on Jasper today . . . or In math class today . . . All items were scored strongly disagree, disagree, both agree and disagree, agree, or strongly agree.*
that used the Jasper materials (CTGV, 1992; Hickey, Pellegrino, Goldman, Vye, & Moore, 1993) revealed significant gains on measures such as math self-concept, personal interest in mathematics, interest in complex math problems, and interest in math class, when compared to students in classrooms that did not use the materials.

Students' motivational beliefs were assessed using the 50-item self-report Motivational Beliefs Survey at the beginning and the end of the school year. Table 4 describes the survey and the relevant measures in more detail. The survey was refined alongside the Motivational Experiences Survey during the pilot studies. The original pool of items was derived from various sources, including published instruments and prior studies of the Jasper materials (CTGV 1992, 1997; Hickey, Pellegrino, Goldman, Vye, & Moore, 1993), and the scales were refined during the pilot studies using factor analytic methods.

Mathematical achievement. In general, we expected that being in a classroom that used the Jasper materials and being in a classroom that was ranked more consistent with the broader reforms would lead to larger gains in mathematical problem-solving ability and conceptual mathematical knowledge. However, it was less clear what to expect for more basic arithmetical fluency. On the one hand, the innovative approaches minimize emphasis on isolated computational skills and associated "drill and practice" activities, and may take away math class time that might otherwise be devoted to such activities. On the other hand, proponents of new approaches insist that even such basic skills are best developed in the context of more meaningful problem-solving activities (e.g., Means, Chelmer, & Knapp, 1991).

Although we were unable to examine scores on the cumulative assessments embedded in the math curriculum, the school district did provide percentile achievement scores on all three mathematical subtests on the Iowa Test of Basic Skills (ITBS). Scores were provided for every student in the study classrooms from the end of the third grade and the end of the fifth grade (the study year)—the district does not test fourth graders. As documented by Romberg & Wilson (1992), 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. As described previously, prior evaluations of Jasper and other such innovations have primarily used standardized achievement measures to document that this particular aspect of achievement is maintained (e.g., CTGV, 1992, 1994; Lamon et. al., 1996; Scardamalia, Bereiter, & Lamon, 1994).

Our initial concerns regarding ITBS scores were modified when we learned that the ITBS scores that we had been provided were based on version of the test (Form K) that had just been revised, ostensibly to reflect the concerns advanced by the NCTM. The publishers of the ITBS state that Form K was developed "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." (Riverside Publishing, 1994, p. 22). On the Math Concepts & Estimation subtest (formerly Math Concepts), the publishers state that Form K increased the emphasis on estimation, required a wider variety of estimation strategies, and included more items (roughly two-thirds) in a contextualized format. The Mathematics Problem Solving and Data Interpretation subtest (formerly Mathematics Problem Solving) requires students to solve unique, complex problems and to interpret and analyze data presented in tables, graphs, and charts. Form K was said to increase the focus on interpretation, minimize the role of computational skills, and include items "that could be thought of as 'problem solving process' or 'strategy' items" that were "based on an adaptation of Polya's four-step problem-solving method" (p. 24). The Mathematics Computation subtest remains "a very direct measure of computation requiring a single operation." Most significantly for this study's purpose, the publishers claim that Form K "does not confound the measurement of concept and estimation skills with the assessment of computational skill" and that the data interpretation measures "require essentially no computation" (p. 25). This purportedly enhanced discrimination promised to help isolate changes in the higher-level skills targeted by the Jasper materials and the broader reforms from the changes in basic computational fluency. Whereas the higher-order skills are expected to increase gradually across the year as students' mathematical reasoning skills develop, computational fluency is expected to increase or decrease more directly in response to massed rehearsal prior to test administration.

The correspondence between the ITBS Form K math subtests and the NCTM standards is certainly debatable. Whereas the previous forms of the ITBS were systematically reviewed in light of the NCTM Standards (Romberg & Wilson, 1992), Form K apparently has not been subjected to such a review (L. D. Wilson, personal communication, June 16, 1997; T. A. Romberg, personal communication, July 20, 1999); nor could the ITBS developers provide additional support for their claims of correspondence (D. A. Frisbie, personal communication, May 26, 1997). Our own examination certainly supported the publisher's claim that Form K is relatively more consistent with the NCTM standards than the previous forms. In one Form K sample booklet, the 30-item Problem Solving and Data Interpretation Subtest consisted of seven sets of items, with each set organized around a single problem context. For example, one cluster of problem-solving items had students use a restaurant menu to figure the total cost of several items, calculate the savings when ordering a meal instead of a la carte, etc. The sets of data interpretation items required students to use an "authentic" symbol chart, bar graph, or a line graph to answer several questions about the information each represented. Per the developer's claims, the computational demands appeared trivial in light of the problem's complexity.

Although not as complex as the problems in the Jasper activities, our examination suggested that some of the problems in the Problem Solving & Data Interpretation subtest on the ITBS Form K capture at least part of the
understanding that the Jasper activities were designed to foster. Relative to the mathematics curriculum, our examination of the instructional activities, performance assessments, and cumulative assessments for Level 16 revealed some problems that were similar in complexity and format to many of the problems in both the Math Concepts and Estimation, and Problem Solving and Data Interpretation subtests.

Assessment and Instruction Sequence

In October, the first two authors visited each of the four schools to meet with principals and IRTs, administer the Motivational Beliefs Survey, and explain the Motivational Experiences Survey administration to the teachers. Teachers were provided with a detailed calendar that specified the administration sequence for the Jasper activities and the various surveys. This sequence is displayed graphically in Figure 2. Specifically, Jasper teachers were asked to schedule the three Jasper activities within a specified range of dates, and to administer a Motivational Experiences Survey at the conclusion of the middle period of each activity. Teachers in both Jasper and non-Jasper classrooms were asked to select three comparison math activities that took place about midway between the periods specified for the Jasper adventures, and to administer a Motivational Experiences Survey at the end of the class period that fell in the middle of the selected activity. The administration instruction reiterated the importance of administering Motivational Experiences Surveys
following a period that was generally representative of the particular objective or unit. Teachers were also asked to complete a brief questionnaire each time a Motivational Experiences Survey was administered. This questionnaire asked teachers to list the specific curricular objective or Jasper activity covered during the period for which the Motivational Experiences Survey was completed, the representativeness of the activity relative to the overall Jasper adventure or overall math curriculum unit, the degree to which the teacher diverged from the Jasper sample lesson plans or the mathematics curriculum guide, and the nature of any grades or rewards provided for students’ participation and success. In March, the same two researchers returned to administer the Motivational Beliefs Survey a second time, gather the assessments, and interview each of the teachers.

Results

Analyses Conducted

Following is a description of the preliminary and substantive analyses carried out using scores from the Motivational Experiences Surveys, Motivational Beliefs Surveys, and ITBS mathematical achievement subtests. Each of the three sets of scores was examined separately using a fully nested design.

Motivational experiences. Analyses of students’ scores on the Motivational Experiences Surveys reveal students’ self-reported motivational states during the third activity period for each of the three Jasper adventures in the 10 Jasper classrooms and a representative class period during each of three comparison mathematics activities in 18 of the 19 classrooms. One high-SES non-Jasper teacher declined to administer the Motivational Experiences Survey to her students.

Information about the representativeness of the comparison (non-Jasper) math activities during the periods for which students completed Motivational Experiences Surveys was gleaned from questionnaires that teachers were asked to complete for each survey administration. Eight Jasper teachers and seven non-Jasper teachers provided specific curricular objectives being covered when students completed the surveys. The 15 teachers identified 35 objectives across the three waves of surveys. In reference to the previous description of the math curricula, the objectives were evenly split between Level 16 and Level 17, and eight came from the cluster of four Level 16 objectives described in detail. Along with objective being covered, 27 responses were provided for the question “How closely were you following the curriculum guide?” Seventeen responses indicated “no difference” or “same” and the remaining responses indicated small divergences such as “used more manipulatives” or “used extra time.” These responses, along with the survey administration instructions, support our expectation that the non-Jasper math activities for which students completed Motivational Experiences Surveys were representative of the particular curricular unit, and that those units were representative of the math curricula for the entire year, as specified in the district’s curriculum guide.
Factor analysis of three major groups of Motivational Experience Survey items was used to verify discriminant validity at each of the six administrations, and several items were dropped. Reliabilities were acceptable (see Table 3) and consistent both within and across the three comparison activities and the three Jasper activities. A total of 397 students in 18 classrooms completed at least one Motivational Experiences Survey. However, many students missed one or more surveys, one high-SES non-Jasper classroom declined to administer any of the Motivational Experiences Surveys, and an administrative conflict led all of the high-SES control classrooms to skip the third Motivational Experiences Survey. Thus, only 46 students in non-Jasper classrooms completed all three Motivational Experiences Surveys, while only 116 students in Jasper classrooms completed all three surveys during the comparison math activities and all three surveys for the Jasper activities. To obtain a data set that maintained a balanced design, scale scores were imputed for the students who completed two of the three surveys by assigning the missing scale score the mean of the other two scores. Imputing scores for the 158 students who missed one of the three surveys for the comparison math activities yielded 328 students, with 10–25 individuals in the 18 classrooms. Imputing scores for the 82 students who missed one of the three surveys during Jasper activities yielded 208 students, with 17–25 students in the 10 classrooms.

Two analyses were conducted using scores on the Motivational Experiences Survey. The first was a within-subjects comparison of the effect of activity type, using scores from the three surveys administered during the Jasper activities and the three surveys administered during the comparison math activities in the 10 Jasper classrooms. These scores were analyzed with four-way ANOVAs, with activity type (jasper vs. comparison) and time (first, second, or third Jasper or comparison activity) as within-subjects variables, and reform orientation (more consistent vs. less consistent) and SES (low vs. high) as between-groups variables. The configuration of classrooms within groups and groups within schools presented a hierarchically nested design. This specified that all effects be tested against the mean square error associated with the nested variables. Thus, all between-groups effects were tested against the mean square error for the nested class variable, while within-groups effects were tested against the mean square error for the interaction of the within-subjects variable(s) and class. The presence of two within-subjects variables (activity and time) required a different error term for each of the three combinations of within-group effects. The number of classrooms relative to the number of groups yielded extremely conservative tests, with most effects tested with 6 or 12 degrees of freedom in the denominator. Therefore, the Type I error alpha (the level at which effects were deemed unlikely to have occurred by chance) was set to .10 for all substantive effects. There were large differences in the magnitude of the class effects and their various interactions with the within-subjects variables. A larger class effect indicates that group means varied more within the nested levels of the model, while a larger Class × Time interaction indicates that the
changes across time varied more within the nested levels of the model. The larger effects yield a larger mean square error for tests of substantive effects, thus explaining why some seemingly large substantive effects are not statistically unlikely.

A second analysis of the Motivational Experiences Survey scores yielded the between-groups comparison used to consider the effect of reform orientation on students' motivational experiences. Scores on the three surveys completed during the comparison math activities in the 18 classrooms were analyzed with four-way ANOVAs with time as a within-subjects variable and reform orientation, instructional group, and SES as between-group variables. Again, the hierarchically nested design specified that between-groups effects were tested against the mean square error for class, while time effects were tested against the Class × Time interaction. Once again, the class and Class × Time effects ranged from trivial to statistically significant.12

Motivational beliefs. Analyses of the scores on the Motivational Beliefs Survey reveal students' more general beliefs about school and mathematics at the beginning and the end of the school year. As shown on Table 4, reliabilities of all scales on the Motivational Beliefs Survey were at least .70. A total of 331 students (183 Jasper students and 148 control students) in all 19 classrooms completed both the year-start and year-end administration. Analysis of these scores yielded the Group × Time comparisons used to

\begin{table}
\centering
\caption{Scales on Motivational Beliefs Survey}
\begin{tabular}{llll}
\hline
Scale & No. of items & \(\alpha^a\) & Example item\textsuperscript{b} \\
\hline
Math dispositions & & & \\
Personal interest in math & 4 & .88 & Mathematics is enjoyable \\
Interest in complex problems & 5 & .79 & I like the challenge of solving complex math problems \\
Math self-concept & 5 & .70 & I am able to get good grades in math class \\
Situational interest in math & 4 & .70 & I actually look forward to going to math class now \\
Motivational orientation & & & \\
Learning orientation & 7 & .72 & (I feel most pleased in school . . .) when what I learn really makes sense \\
Performance/ego & 7 & .73 & (I feel most pleased in school . . .) when I am the only one who can answer a question \\
Work-avoidant & 6 & .77 & (I feel most pleased in school . . .) when I get out of some work \\
\hline
\end{tabular}
\end{table}

\textsuperscript{a}Pretest internal consistency (Cronbach's alpha).
\textsuperscript{b}Items were scored strongly disagree, disagree, both agree and disagree, agree, or strongly agree.
consider the effect of instructional group and reform orientation on students' motivational beliefs. The scores were analyzed using four-way ANOVAs, with time (beginning year vs. year end) as a within-subjects variable and reform orientation, instructional group, and SES as between-groups variables. Class was hierarchically nested within classroom environment, SES, and instructional group. Between-subjects effects were tested against the mean square error for the nested class variable, while the effects involving time were tested against the mean square error for the Class x Time interaction.

Math achievement tests. The district administered the ITBS approximately 2 weeks after the year-end Motivational Beliefs Survey was administered. Scores on the three ITBS mathematics subtests were obtained for every student who was present in the study classrooms at the end of the school year; scores from the third-grade administration of the same form of the IBS were obtained for 293 of these students. This included students from all 19 classrooms, with 12–22 students per class. Table 5 shows the means and standard deviations for percentile scores for these students in the two or three classrooms in each cell of the $2 \times 2 \times 2$ design. These scores yielded the Group x Time comparisons used to consider the effect of instructional group and reform orientation on math achievement. Subtests scores were analyzed

### Table 5

<table>
<thead>
<tr>
<th>SES/instructional group</th>
<th>Reform orientation</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traditional, time</td>
<td>Reform oriented, time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Third grade</td>
<td>Fifth grade</td>
<td>Third grade</td>
<td>Fifth grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Problem solving and data interpretation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low/non-Jasper</td>
<td>36.0</td>
<td>23.9</td>
<td>34.0</td>
<td>25.1</td>
<td>47.1</td>
<td>28.0</td>
<td>42.7</td>
<td>25.2</td>
</tr>
<tr>
<td>Low/Jasper</td>
<td>27.8</td>
<td>21.7</td>
<td>31.6</td>
<td>26.1</td>
<td>26.6</td>
<td>23.5</td>
<td>36.1</td>
<td>24.0</td>
</tr>
<tr>
<td>High/non-Jasper</td>
<td>69.7</td>
<td>21.3</td>
<td>63.1</td>
<td>25.1</td>
<td>74.0</td>
<td>23.0</td>
<td>77.8</td>
<td>23.5</td>
</tr>
<tr>
<td>High/Jasper</td>
<td>70.5</td>
<td>24.6</td>
<td>74.7</td>
<td>20.7</td>
<td>73.0</td>
<td>22.1</td>
<td>78.5</td>
<td>18.6</td>
</tr>
<tr>
<td>Concepts and estimation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low/non-Jasper</td>
<td>44.5</td>
<td>26.5</td>
<td>34.0</td>
<td>26.0</td>
<td>50.6</td>
<td>31.5</td>
<td>49.8</td>
<td>28.0</td>
</tr>
<tr>
<td>Low/Jasper</td>
<td>34.3</td>
<td>24.5</td>
<td>24.2</td>
<td>22.4</td>
<td>32.5</td>
<td>25.2</td>
<td>26.0</td>
<td>21.9</td>
</tr>
<tr>
<td>High/non-Jasper</td>
<td>68.8</td>
<td>23.6</td>
<td>71.3</td>
<td>18.6</td>
<td>79.0</td>
<td>24.2</td>
<td>80.3</td>
<td>24.5</td>
</tr>
<tr>
<td>High/Jasper</td>
<td>74.3</td>
<td>23.3</td>
<td>81.1</td>
<td>16.9</td>
<td>76.8</td>
<td>18.1</td>
<td>83.8</td>
<td>18.7</td>
</tr>
<tr>
<td>Math computation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low/non-Jasper</td>
<td>47.3</td>
<td>27.6</td>
<td>39.8</td>
<td>25.3</td>
<td>53.0</td>
<td>26.9</td>
<td>46.3</td>
<td>28.2</td>
</tr>
<tr>
<td>Low/Jasper</td>
<td>34.6</td>
<td>24.4</td>
<td>49.3</td>
<td>26.9</td>
<td>35.8</td>
<td>25.7</td>
<td>22.1</td>
<td>20.4</td>
</tr>
<tr>
<td>High/non-Jasper</td>
<td>66.0</td>
<td>22.3</td>
<td>64.6</td>
<td>19.8</td>
<td>61.6</td>
<td>29.1</td>
<td>67.4</td>
<td>29.8</td>
</tr>
<tr>
<td>High/Jasper</td>
<td>70.8</td>
<td>23.6</td>
<td>75.1</td>
<td>18.4</td>
<td>75.1</td>
<td>20.9</td>
<td>67.0</td>
<td>24.1</td>
</tr>
</tbody>
</table>
using four-way ANOVAs with time (third grade vs. fifth grade) as a within-subjects variable and reform orientation, instructional group, and SES as between-groups variables. Class × Time interactions (the error terms for the substantive effects) were non-significant for Problem Solving and Data Interpretation, and Concepts and Estimation, but there was a highly significant Class × Time interaction for math computation, $F(11,274) = 3.78, p < .0001$. This shows that changes on the Computation subtest from third grade to fifth grade were very unstable within the nested levels of the model. This makes sense, given that changes in Math Computation scores should be more directly influenced by individual teachers' curricular decisions (i.e., time devoted to drill and practice), compared to the other subtests. Large main effects of class for all three subtests reflected ability grouping within schools, while large main effects of SES reflected achievement differences between high and low. An Instructional Group × SES interaction for the Concepts and Estimation subtest reflected the higher achievement scores in the low-SES control school relative to the low-SES Jasper school.

Results Overview and General Issues

The following results reveal the effects of participating in the Jasper activities and/or spending the year in a classroom ranked more consistent with the broader reforms, in terms of motivational experience, motivational beliefs, and math achievement. These results further reveal whether or not these effects were different for students from high-SES communities, compared to students from low-SES communities.

Several general conclusions gleaned from the IRT and classroom teacher interviews and the questionnaires that some teachers completed when they administered the Motivational Experiences Surveys are relevant for interpreting these results. First, the Jasper materials were not entirely implemented in the manner that the developers intended. It appeared that all 10 teachers followed the general sequence outlined in the sample lesson plan. Although all of the high-SES classrooms followed the recommended practice of completing the adventures during math period across at least 4 consecutive days, all of the low-SES classrooms completed each adventure during three or four class periods on a single day. All six of these Jasper teachers listed their first (or only) goal for the activities as something like “showing students how math problem solving is useful in the real world.” Meanwhile, none alluded to the broader goal of supporting extended collaborative investigation around complex problems. Most teachers reported or were reported to have used “fact sheets” to structure the problem-solving activity, and in one of the classrooms, the Jasper activity was largely reduced to having students compete with each other in answering questions on the fact sheets (an activity appropriately called “Jasper Trivia”). Apparently, many teachers felt that these locally generated worksheets were useful for keeping students focused on the task and preventing them from getting frustrated—particularly during the first adventure. However, these aids may
Motivational and Academic Consequences

have constrained student problem generation in a manner that compromised the larger activity.  

A second finding from the interviews involved students' concern with their progress in math class. Students' progression through the 18 levels of the math curricula was determined by their proficiency according to the performance assessments embedded in each level and on the cumulative assessment at the end of each level. In the high-SES schools, IRTs and teachers reported that students and parents were quite concerned with progressing through as many levels as possible—particularly so during fifth grade because students' level at the end of that year dictated starting level (and so, class assignments) for middle school the following year. The significance of this finding was not apparent until we considered how the high-SES students appraised the relevance of the Jasper activities. This is considered subsequently.

Consequences of Participating in the Jasper Activities

These results reveal whether (a) students in the Jasper classrooms reported different motivational experiences during the Jasper activities than during their other non-Jasper math activities; (b) whether changes in motivational beliefs and math achievement for students in the 10 Jasper classroom were different than for the students in the 9 non-Jasper classrooms; and (c) whether the effects of the Jasper activities on motivational experience, motivational beliefs, and math achievement were different in the high-SES schools than in the low-SES schools.

Motivational experience. These results contrast students' self-reported motivational states during the third period for each of the three Jasper adventures and during a class period that fell midway in the unit for each of three comparison math activities. Most, if not all of the comparison activities assessed were conducted in accordance with the district's mathematics curriculum guide. Table 6 contrasts scale scores on the Motivational Experiences Surveys completed during the Jasper and comparison activities in the 10 Jasper classrooms. Students reported lower subjective competence during the Jasper activities than those same students reported during the comparison math activities \( F(1,6) = 9.15, p < .025 \). This was not unexpected, given that the Jasper activities were more complex, challenging, and novel than the comparison activities. However, the main effect of activity type was subject to an important interaction with time and reform orientation \( F(2,12) = 2.99, p < .10 \). As shown in Figure 3, subjective competence during the Jasper activities (circles) started out lower than the comparison activities (triangles) in both more-consistent classrooms (bold lines) and the less-consistent classrooms (thin lines). However, in the less-consistent classrooms, by the third activity, subjective competence during the Jasper activities (thin/circles) was higher than it was during the comparison math activities (thin/triangles). Thus, for students in the five classrooms ranked less consistent with the broader reforms, the lower levels of subjective competence for the Jasper activities were specific to their initial encounters with the activities, and
repeated exposure to these activities led to increasingly higher confidence in ability to succeed. In the Jasper classrooms ranked more consistent with the broader reforms, students reported lower subjective competence during all three Jasper activities (bold/circles) relative to the comparison activities (bold/triangles), although subjective competence did increase across the Jasper activities.

Another finding shown in Table 6 ran counter to expectation and prior research—students in the 10 Jasper classrooms appraised the Jasper activities to be less “relevant” than the comparison activities, $F(1,6) = 7.46, p < .05$. However, this main effect was subject to an interaction with SES that led us to a plausible explanation for this finding. Before considering this interaction, the large main effect of SES on motivational experience should be noted. Compared to the suburban students in the high-SES schools, the relatively disadvantaged low-SES students reported dramatically more positive task appraisals and motivation orientation. For some measures, the difference was as large as a full standard deviation. The existence of such differences and their implications has been discussed extensively in the context of self-concept research (e.g., Graham, 1994). The present concern is whether SES interacted with the other substantive effects—indicating that the differences attributed to the Jasper activities were themselves different in the high-SES and low-SES students. There was only one such interaction, between SES and activity type for appraisals of relevance. As shown in Figure 4, the low-SES students appraised the Jasper activities and comparison activities as equally relevant, while the high-SES students appraised the Jasper activities as less relevant than the comparison activities, $F(1,6) = 3.77, p < .10$. This finding is noteworthy, given that the interviews with the high-SES IRTs and teachers revealed student and parental concerns with advanc-
Figure 3. Change in subjective competence, by activity type and reform orientation, across three activities.

ing through as many levels of the math curriculum as possible. Follow-up interviews with high-SES Jasper teachers later revealed that many of these students (correctly) perceived that devoting class time to the Jasper activities effectively delayed their progress through the levels of the curriculum. The low-SES IRTs and teachers initially volunteered no such parental or student concerns, and later confirmed that students did not express concern that the Jasper activities were delaying their progress through the levels of the broader math curriculum. This suggests that the high-SES students appraised the relevance of the Jasper activities in terms of the extrinsic reward structure of the larger academic context, rather than the intrinsic relevance of the actual learning activities.

Another finding presented in Table 6 concerns motivation orientation, with students reporting lower levels of performance orientation during the Jasper activities, \( F(1,6) = 20.3, p < .005 \). Apparently the Jasper environment diminished students' focus on overtly demonstrating one's competence and "looking smart." This likely reflects the lack of explicit individual-level performance appraisals in the Jasper activities, relative to the comparison activities. There were no explicit assessments described in the Jasper sample lesson plans; of the 18 responses recorded on the Motivational Experiences Survey questionnaire regarding reward/grading practices, 17 reported providing only "verbal praise" and one reported providing points to groups for
the quality of their participation. Meanwhile the comparison math activities featured extensive public and private assessment with fairly high stakes.

Motivational beliefs. In contrast to earlier studies of the Jasper environment, there were no significant Instructional Group × Time interactions for any of the motivational belief measures. Nor were there further interactions of Instructional Group × Time with either SES or reform orientation. Thus, motivational beliefs appeared unchanged for students in the 10 classrooms that participated in the Jasper activities, relative to students in 9 classrooms whose math curricula did not include the Jasper activities.

Mathematical achievement. Regarding the consequences of spending a year in a classroom that participated in the Jasper activities, Figure 5 shows the Instructional Group × Time interactions for the three ITBS math subtests. A highly significant interaction was present for Problem Solving and Data Interpretation, $F(1,11) = 21.3$, $p < .001$. As shown in the top panel of Figure 5, students in the Jasper classrooms increased substantially, while students in the non-Jasper classrooms decreased slightly. The large proportion of variance accounted for reflects the fact that the interaction was represented throughout the nested levels of the design. In other words, the scores in every Jasper classroom increased while the scores in every non-Jasper classroom stayed the same or went down slightly. A similar pattern was present for Concepts and Estimation (middle panel), but the interaction accounted for a small proportion of variance ($F < 1$); mathematics computation (bottom panel) declined slightly in both groups. These results strongly suggest that including the Jasper activities in the fifth graders' math curricula had a substantial, positive impact on their ability to solve moderately complex math problems.
Figure 5. Changes in ITBS math subtest scores, by instructional group.
Interactions with SES. As stated previously, the only interaction of SES with motivational experience measures was for relevance appraisals (shown in Figure 4) and the motivational belief measures revealed no interactions of SES with the Instructional Group × Time interaction. However, examining the math achievement scores revealed that SES interacted with the Instructional Group × Time interaction for concepts & estimation, $F(1,11) = 6.09, p < .05$. As shown in Figure 6, positive consequences of participating in the Jasper activity on this measure are present for high-SES classrooms, but not the low-SES classrooms. Thus, participating in the Jasper activities appeared to enhance problem-solving ability in both the high-SES and low-SES classrooms, but the positive consequences for conceptual knowledge and estimation skills were limited to the high-SES classrooms. This may have been the result of the relatively compressed implementation of the Jasper activities in the low-SES classrooms across consecutive class periods on a single day. Although we lack sufficient comparative information on actual instructional practices, it seems that the combination of lower achievement and a reduced number of hours devoted to the Jasper activities in the low-SES classrooms
might have minimized the focus on the conceptual aspects of mathematics and estimation skills during the Jasper activities. To the extent that there were no significant relative declines in any scores for the Jasper students, there is no evidence that participating in the Jasper activities had negative consequences for students from either SES stratum.

**Consequences of Reform Orientation**

These results reveal the effects of spending the year in one of the classrooms ranked more consistent versus less consistent with the broader NCTM-oriented reforms in the three areas of motivational experience, motivational beliefs, and math achievement.

*Motivational experiences.* The only statistically significant main effect or interaction for motivational experience was that students in the more-consistent classrooms reported higher subjective competence than students in the less-consistent classrooms [4.12 versus 3.80, \( F(1,10) = 5.29, p < .05 \)]. Given the emphasis on performance assessment in the new curriculum, there is reason to believe that students were exposed to *more* assessment in the classrooms ranked more consistent with the reforms. As such, the enhanced subjective competence in the more-consistent classrooms suggests a meaningful positive change in the motivational environment of these classrooms (rather than less exposure to assessment).

*Motivational beliefs.* As shown in Figure 7, math self-concept increased in the more-consistent classrooms while decreasing in the less-consistent classrooms, \( F(1,11) = 3.45, p < .10 \). A similar pattern of change was observed for interest in complex problems, but the statistical significance did not reach
the cutoff for Type II error, $F(1,11) = 3.10, p \approx .11$. As shown in Figure 8, personal interest in math was maintained in the more-consistent classrooms, but declined in the less-consistent classrooms, $F(1,11) = 3.21, p < .10$. Thus, spending a year in a classroom that was ranked more consistent with the broader reforms helped students maintain desirable beliefs about mathematics, whereas spending a year in the other classrooms led to decreases in those beliefs. Thus, we conclude that classroom environments that were more consistent with the NCTM-oriented reforms did, in fact, help achieve two of the NCTM's "five new goals for mathematics students": becoming confident in one's own ability and learning to value mathematics.

**Mathematical achievement.** Figure 9 shows the Reform Orientation x Time interactions for the three ITBS subtests. Significant interactions were present for Problem Solving and Data Interpretation, $F(1,11) = 6.73, p < .025$, and Concepts and Estimation, $F(1,11) = 4.64, p < .10$. In both cases, the achievement of students in the more-consistent classrooms increased relative to students in the less-consistent classrooms. The opposite pattern was observed for math computation; while this effect appeared quite large, it did not reach significance, $F(1,11) = 2.58$, because of the relatively large proportion of variance accounted for within the groups. Given that the district's revised curriculum shifted the focus away from the kinds of problems in the math computation subtest and towards the kinds of problems on the other two subtests, these results make perfect sense. We conclude that classrooms that were reported to have more faithfully enacted the revised curricula were more successful in increasing scores on the two higher-level mathematical achievement subtests, compared to the other classrooms.
Figure 9. Change in ITBS math subtest scores, by reform orientation.

Interaction with SES. No significant SES × Reform orientation interactions were found for the motivational experience measures and no signifi-
significant interactions of SES with the Time \times Reform Orientation interaction were found for the motivational belief measures (all $F < 1$). Thus, the motivational consequences of reform orientation (or lack thereof) appear to be the same for high-SES and low-SES schools. On the math achievement measures, a significant interaction of SES with the Time $\times$ Reform Orientation interaction was found for the Math Concepts and Estimation subtest, $F(1,11) = 5.31, p < .05$. As shown in Figure 10, the increase on this subtest in the high-SES classrooms (squares) was the same in both the more-consistent and less-consistent classrooms. However, the decrease on this subtest in the low-SES classrooms (triangles) was significantly greater in the less-consistent classrooms, compared to the more-consistent classrooms. Thus, with one exception favoring the low-SES students, the consequences of reform orientation were similar for the high-SES and low-SES students. With the exception of decreased performance on the Mathematics Computation subtest, there was no evidence of negative consequences of being in a classroom ranked more consistent with the reforms for students from either community.

Figure 10. Change in ITBS math concepts and estimation scores, by reform orientation and SES.
Motivational and Academic Consequences

Consequences of Jasper Activities and Reform Orientation

Of particular interest in the present analysis are the consequences of spending the year in one of the five classrooms that used the Jasper activities and was ranked more-consistent with the NCTM curricular reforms. Regarding motivational experience, there was no indication that students' reported experiences were particularly noteworthy in these classrooms; for motivational beliefs, the largest increases on some of the measures were observed in these five classrooms, but the effects were inconsistent and quite small. Regarding mathematical achievement, Figure 11 presents the results from the Reform Orientation x Instructional Group x Time interactions for the three subtests. As shown in the top panel, students in the five more-consistent/Jasper classrooms showed the largest gains in Problem Solving and Data Interpretation (a gain of 6.8 percentile points), while students in the four less-consistent/non-Jasper classrooms showed the only decline (4.1 percentile points). Even though this interaction effect ultimately accounted for a small proportion of the variance, $F(1,11) = 0.15$, it is noteworthy because of the evidence it provides for the consequences of combining both instructional innovation and the broader curricular reform. As shown in the middle panel, a similar, less-pronounced pattern was observed for the Math Concepts and Estimation subtest. However, the interaction for mathematics computation went in the opposite direction and accounted for a statistically significant proportion of variance, $F(1,11) = 4.84, p < .05$. As shown in the bottom panel of Figure 11, mathematics computation scores declined in the five more-consistent/Jasper classrooms, but increased in the less-consistent/Jasper classrooms. This interaction likely reflects systematic differences in massed rehearsal of computation fluency prior to testing in the two groups of Jasper classrooms.

Discussion and Implications

The purpose of this research study was to determine whether constructivist reforms and instructional programs in the domain of elementary mathematics make a difference in students' academic achievement and motivational responses. The specific program under consideration was the Adventures of Jasper Woodbury (CTGV, 1997), which has been shown previously to positively impact student attitudes and achievement in mathematics (CTGV, 1992, Hickey, Pellegrino, Goldman, Vye, & Moore, 1993; Pellegrino, Hickey, Heath, Rewey, & Vye, 1992). The latter results have typically been obtained with teachers who were more likely to have attempted to establish classroom environments that promote the student-centered pedagogy and learning goals advocated by the NCTM (1989, 1991). The present study attempted to assess whether similar outcomes would be obtained in a wider range of classrooms, under conditions where the decision to implement the instructional program in a particular classroom was unrelated to that teacher's interests or orientation, and where implementation was more likely to reflect typical school and system adoption characteristics. A variety of interesting
Figure 11. Change in ITBS math subtest scores, by instructional group and reform orientation.
results were thus obtained regarding student outcomes relative to the instructional program and reform agenda. Consideration will be given first to outcomes related to the general reform orientation of specific classrooms before then considering effects related to the presence of the Jasper program in these diverse classrooms.

The district in which this study was conducted had previously adopted the NCTM reform recommendations which relate very directly to both the nature and content of mathematics instruction. The district had taken numerous steps in support of implementing such a district-wide policy for the elementary mathematics curriculum. Not surprisingly, however, teachers within and across schools varied widely in the degree to which their enactment of the "official" curricula reflected the reorientation of mathematics instruction, instructional strategies, and curricular emphases associated with the NCTM recommendations. This natural variation made it possible to examine, on a within-school basis, the possible effects of the degree of reform orientation on student achievement and motivational outcomes, whether or not the Jasper program was being used on a regular basis. Several of the study's findings can be interpreted as supportive of the positive impact of NCTM-oriented curricular reforms.

The IRT and classroom teacher interviews, together with observations made while administering assessments at the beginning and end of the school year, verified that the less-consistent classrooms appeared relatively representative of more conventional teacher-directed models of educational practice. The classrooms that were ranked as more consistent with the broader reforms appeared more representative of what NCTM had recommended as desirable for mathematics instruction and learning. With respect to student outcomes, the classrooms ranked more consistent with the NCTM reforms yielded substantially larger gains on the two higher-level mathematics achievement subtests—Problem Solving and Data Interpretation, and Math Concepts and Estimation. These are highly desired results given prior assessments of specific weaknesses and deficiencies in the mathematical skills of American children (e.g., United States Department of Education, 1996). Furthermore, these results were observed in the context of mixed and inconsistent results for basic computational skills.

As noted in the introduction, a major issue with respect to introducing curricular and instructional reforms that emphasize the teaching and learning of more complex and abstract knowledge and skills is whether there will be a differential impact for students more versus less academically advantaged. Thus, it is especially noteworthy that the desirable consequences for math achievement were most pronounced in the less advantaged, low-SES schools, refuting the argument that academically disadvantaged students cannot benefit from reforms that focus on more complex mathematics competencies and knowledge.

In both economic strata, students in the more-consistent classrooms also displayed enhanced math self-concept and interest in math, and students in these classrooms also displayed greater subjective competence in mathemat-
ics throughout the school year. The enhanced mathematical achievement results, coupled with the likely increased focus on assessment in the more-consistent classroom, appears to rule out an interpretation of the motivational findings as simply reflecting student positive reactions to a "watered-down" and less demanding curriculum and instructional environment.

Given the design of the study, it was possible to examine for effects of the Jasper instructional program independent of classroom reform orientation. The mathematical achievement results in the area of problem solving and data interpretation clearly showed that the Jasper instructional implementation had very desirable consequences, with no evidence of negative consequences. Positive outcomes in the area of problem solving were shown for the Jasper program for both the advantaged high-SES schools and the relatively disadvantaged low-SES schools. This pattern of findings for the Jasper materials supports the argument that academically disadvantaged students can profit from the complex problem-solving activities associated with the Jasper materials and that such students do not suffer negative academic or motivational consequences.

It is very interesting, and perhaps not surprising, that the Jasper activities, in and of themselves, did not enhance students' motivational beliefs. This result is particularly noteworthy given the concurrent finding that reform orientation yielded precisely the effects on motivational beliefs that have been attributed to the Jasper activities in prior studies (i.e., Hickey, Pellegrino, Goldman, Vye, & Moore, 1993, CTGV, 1992). The present data suggest that the apparent positive motivational effects of the Jasper activities obtained in prior studies may have been due to teacher attitudes and instructional practices as they attempted to faithfully implement the Jasper program. In other words, the prior motivational gains likely reflect the more general instructional orientation of the teachers who used the Jasper materials, rather than the materials themselves.

Of special concern in the present study were the consequences of employing the Jasper activities in classrooms representing different models of educational practice. The present study documents the positive academic consequences of the Jasper activities when used in a range of classrooms, and helps rule out any negative consequences when used in more conventional classrooms. Most importantly, the findings reveal that the positive consequences of the Jasper activities were the largest in the classrooms that were more consistent with the student-centered model of practice advocated by the NCTM. This is what might be expected given the design principles behind the Jasper Adventures and instructional program and their consistency with the NCTM standards. Furthermore, it supports the common-sense notion that coordinating both instructional innovation and curricular reforms around a common constructivist perspective is the most effective way of enhancing students' conceptual understanding of mathematics and their ability to solve complex mathematical problems.

The results of the present study also have implications for professional development activities associated with instructional innovations such as the
Jasper materials. At the most general level, the results suggest the need for ongoing training and feedback for teachers implementing student-centered curricular and instructional reform principles such as those advocated by the NCTM and actually instantiated in the Jasper materials and program. The simple fact that teachers were observed to utilize the Jasper materials in ways that are antithetical to the original design purposes and/or NCTM recommendations makes it clear that the "affordances" built into materials and programs designed to support constructivist learning and teaching will not necessarily be realized. Rather, considerable scaffolding may be required for teachers to be able to maximize their effectiveness with these materials, especially when they are relative novices in their use (see CTGV, 1997, Chapter 6). The concept of a "zone of proximal development" may be especially useful and applicable in the case of designing professional development programs that support teachers in understanding and adopting various curricular and instructional reform programs, especially those with a constructivist learning and teaching philosophy.

It is worthwhile to note that this study was conducted during the initial phases of one of the first large-scale, district-initiated implementations of the Jasper materials. After the study was conducted, a range of training and curricular support materials (including training videos) became available, as well as additional Jasper adventures. This particular school district also continued to refine its support for the activities and their integration into the broader mathematics curriculum, and the teachers have had more experience using and integrating them. The district has also implemented three of the subsequent Jasper adventures in its sixth-grade classrooms (there are 12 all together). Thus, there is every reason to expect that the positive consequences of the Jasper activities would be even greater than those observed here, especially if data were obtained longitudinally across multiple grades.

Notes

This article is based on research conducted in partial fulfillment of the doctoral degree by Daniel T. Hickey at Vanderbilt University. The study was carried out with the support of a grant from the American Educational Research Association, which received its funds for its AERA Grants Program from the National Science Foundation and the National Center for Educational Statistics (U.S. Department of Education) under NSF Grant No. RED-9255347. We gratefully acknowledge the support and participation of the administrators, teachers, and students in the participating school system, particularly the mathematics coordinator, director of testing, the instructional resource teachers at each of the four schools, and the 2 classroom teachers who coordinated training. Additional preparation of this article was supported by a Postdoctoral Fellowship from Educational Testing Service. The authors acknowledge the input and assistance of numerous individuals at Vanderbilt University and the Vanderbilt Learning Technology Center, particularly Susan Goldman, John Bransford, Craig Smith, and David Cordray. Additional input was provided by Ana Maria Brannan, Ernest Valente, Caroline Brayer Ebby, Shane Bland, Bobby Gaensler, and several anonymous reviewers.

1For more information, visit http://peabody.vanderbilt.edu/divs/ltc/Research/.
2This and subsequent quotes are taken directly from the district's curriculum guide. The actual reference is omitted to preserve anonymity.
3With scale anchors 1 = not at all and 10 = completely, the questions asked the degree to which students: (a) used problem-solving approaches to investigate and understand
Hickey, Moore, and Pellegrino

mathematical content, (b) communicated with others to discuss mathematical ideas and make conjectures and convincing arguments, (c) investigated mathematical connections in a way that let them see mathematics as an integrated whole and connected to other disciplines, and (d) used inductive and deductive reasoning to support mathematical arguments.

The first question asked "How familiar are you with the NCTM Curriculum and Evaluation Standards?" with scale anchors 1 = not at all and 5 = very familiar; the second question asked "How confident do you feel about you knowledge of the study classrooms in terms of the NCTM Standards," with scale anchors 1 = not at all and 5 = very confident.

The interviews were conducted during the summer and we encountered difficulty arranging and conducting the interviews. As such, the findings from the interviews are used primarily to buttress conclusions derived from other data.

Only 8 Jasper teachers and 7 non-Jasper teachers completed questionnaires; just 6 Jasper teachers and 2 non-Jasper teachers completed a questionnaire for each of the Motivational Experiences Surveys they administered to their students. As such, conclusions from the questionnaire will be presented in the context of the other results and then only for the questions that can be answered with this limited set of responses.

It should be noted that this imputation attenuates changes across time; time effects were examined using the restricted data set when feasible.

For technical reasons, all analyses were conducted using SAS PROC-MIXED (SAS Institute, 1992), which has been shown to yield nearly identical model estimates as those yielded by hierarchical linear modeling methods, such as the HLM software (Bryk, Raudenbush, Seltzer, & Congdon, 1988), according to Bland & Appelbaum (1992).

Specifically, all between-groups effects were tested against the MSE for Class (Instructional Group [SES]).

Effects involving time were tested against the Time × Class MSE; effects involving activity were tested against the Activity × Class MSE; effects involving the Time × Activity interactions were tested against the Time × Activity × Class MSE.

The class effect was statistically significant for interestingness (p < .001), learning orientation (p < .01), and work avoidance orientation 2 (p < .05), with 6,198 df, while the Class × Time effect was statistically significant for Interestingness (p < .02) and work avoidance 2 (p < .05) with 12,396 df.

The class effect was statistically significant for work avoidance (p < .02) with 10,310 df, while the Class × Time effect was statistically significant for relevance (p < .03) and learning orientation (p < .06) with 20,620 df.

Such concerns were one focus of an array of teacher training videos and classroom support materials under development at the time of the study.

The overall Activity × Time interaction was not significant, F(2,12) = 1.46. Note that this time effect is based on the imputed sample, which may have attenuated the change across time. It is difficult to interpret between-group effects in the restricted sample because of the loss of two classrooms.

References


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. Author.


Manuscript received March 2, 1998 Revision received November 22, 1999 Accepted May 25, 2000