Introduction and rationale

Historically math and science education reform efforts have consistently advocated specific classroom practices as substantial influences in instruction and learning. However, there are few research efforts available that have empirically connected specific reform-driven instructional strategies with student learning (NCTM, 1989; NRC, 1996). The political community and general public were awakened to the need for math and science educational reform by the announcement of the Third International Mathematics and Science Study (TIMSS, 2005). The 2002 US ‘No Child Left Behind’ Act has also raised public consciousness of the importance of Standards-based Instruction (SBI) in reform efforts (ESEA, 2002). However, there is a lack of empirical evidence to identify which specific math and science SBI strategies significantly influence student learning; the specific influences of non-reform instructional strategies (non-SBI) on students’ math and science learning are also unclear. The present study seeks to add to this research base.

Specific reform strategies known as Standards-based Instruction consist of classroom activities that encourage participatory student-centred classrooms rather than teacher-directed (lecture-based) classrooms. SBI classroom strategies include student self-assessment, inquiry-based activities, group-based projects, hands-on experiences, use of computer technologies, and the use of calculators. The use of Standards-based Instruction in the retraining and preparation of teachers and in the teaching of math and science (grades 6 to 12) has been the focus of the Oklahoma City Public Schools, Oklahoma, Urban Systemic Program (OK USP) since 1999. The OK USP SBI approaches emphasise hands-on, inquiry, connections, communications, problem solving, real-world applications, and co-operative learning approaches to math and science learning. Examples of SBI strategies that illustrate the student-centred focus of Standards-based Instruction include the following: (1) using manipulatives or hands-on materials such as Styrofoam balls and toothpicks for building molecular models, dominoes, base ten blocks, tangrams, spinners, rulers, fraction bars, algebra tiles, coins,
and geometric solids; (2) incorporating inquiry, discovery, and problem-solving approaches such as making binoculars out of recycled materials, using scenarios from nature and everyday life events for groups of students to research and investigate using math and science concepts; (3) applying math and science concepts to real-world contexts such as banking, energy concerns, environmental issues, and timelines; (4) connecting mathematics and science preparation skills to specific careers and occupations; (5) using calculators and technologies for capturing and analysing original data from original math and science experiments; and (6) communicating math and science concepts through journal writing, small-group discussions, and laboratory/technical reporting of experiments and results. These SBI practices are delivered to teachers in the form of summer professional development academies, informal professional development (individual teacher consultations) and formal professional development (lesson modelling).

Non-SBI practices utilised in math and science classrooms include activities such as teacher lecture, individual student drill and practice worksheets, and computer drill and practice programmes.

Data for this study were collected from a standards-based Preparation, Practice, and Performance (P³ model) data-driven framework. Through quantitative analysis the study seeks empirical connections between reform strategies (Standards-based Instruction: SBI) or non-reform instructional strategies (non-Standards-based Instruction: non-SBI) and student learning (achievement).

**Model**

The framework used for the study is a data-driven decision-making model entitled Preparation, Practice and Performance (P³) model developed for the investigation and presented in Figure 1. The model rigorously aligns the threefold areas of teacher professional development programmes (preparation), classroom implementation (practices), and student outcomes (performance) to empirically assess SBI effectiveness. Data collection methods for

![Figure 1 P3 model: relational database structure](image-url)
each of the three areas of teacher preparation, classroom practice, and student performance created a cohesive relational database for this investigation. A relational database is defined as the data connecting specific teachers’ professional development data to data from specific classroom observations of these teachers’ classrooms and relating these data to their students’ performance data. The students’ performance data are then related back to address the appropriateness of the teachers’ professional development efforts. The focus of the D³ model mandates accountability from all three of the preparation, practice, and performance components. The purpose of the model is to clearly identify gains in student performance (achievement data) as well as pinpointing specific successful classroom practices (classroom observational data) and earmarking specific effective professional development strategies (teacher self-assessment data) by aligning these three components.

It is important to note that preparation does not imply practice. Aligning teachers with instructional strategies that represent reform strategies; educating teachers on SBI activities and materials; and readying teachers to teach with those concepts and pedagogies that are reflective of Standards-based Instruction does not imply that these preparation efforts will be implemented in the classroom. Similarly, reform in teacher professional development does not assume positive change in student performance.

Method

The study methodology consists of the following subsections: (1) the instrumentation used in the study (2) the description of the subjects or participants used in the study; and (3) the procedures used in the study.

Instrumentation

The instrumentation developed and used in the study is summarised in Table 1.

Preparation

Teacher assessment form. This self-report instrument is used to assess teachers’ levels of knowledge and attitudes concerning standards-based mathematics and science education. This form was developed using information and adapted items from the NCTM and NSTA standards and the TIMSS survey (NCTM, 1989; NRC, 1996; TIMSS, 2005). A pilot test of the instrument revealed an internal consistency reliability coefficient of 0.93. The form contains five sections:

1. Demographic information. Teachers supply general demographics, including information on math and science formal preparation.
2. Instructional considerations for math and science. Teachers respond to statements that reflect the instructional beliefs or philosophy set forth
by TIMSS (2005) that are directly related to Standards-based Education. A high score on this scale implies that the teacher understands and strongly agrees with the use of the instructional considerations necessary for the standards to be implemented in the classroom. There are twelve items on this subscale, with respondents provided with a 1 to 10 (‘strongly disagree’ to ‘strongly agree’) rating scale for a possible score range of 12 to 120. The reliability of this subscale was found to be 0.87.

3 Preparation considerations for math and science. Teachers respond to specific content and process topics reflected by the standards that are necessary for teachers to have solid preparation and skills for teaching. A high score on this subscale indicates greater perceived preparation of the teacher for teaching standards-based math and science. There are twenty items on this subscale, with respondents provided with a 1 to 10 (‘not at all prepared’ to ‘very prepared’ ) rating scale for a possible score range of 20 to 200. The reliability of this subscale was found to be 0.95.

4 Instructional time considerations for math and science. Teachers review specific kinds of classroom activities that are endorsed by standards-based education and indicate how much time they spend on those types of activities. The higher the score on this subscale the more time the teacher is spending on standards-based activities in the classroom. There are twenty items on this subscale, with respondents provided with a 1 to 10 (‘no time spent’ to ‘virtually all of the time spent’) rating scale for a possible score range of 20 to 200. The reliability of this subscale was found to be 0.92.

5 Assessment considerations for math and science. Teachers indicate their use of specific kinds of assessment procedures in evaluating students’ progress, with half of the suggested types reflecting standards-based

<table>
<thead>
<tr>
<th>Model components</th>
<th>Instrumentation</th>
<th>Description</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation</td>
<td>Teacher assessment form</td>
<td>Overall reliability</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Four subscales:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Instructional considerations</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Preparation considerations</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Time considerations</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 Assessment considerations</td>
<td>0.91</td>
</tr>
<tr>
<td>Practice</td>
<td>Classroom observation form</td>
<td>Overall reliability</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Three subscales:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Math SBI</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Science SBI</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Non-SBI</td>
<td>0.76</td>
</tr>
<tr>
<td>Performance</td>
<td>Iowa Test of Basic Skills (Hoover et al., 1996)</td>
<td>Norm-referenced test</td>
<td>n.a.</td>
</tr>
</tbody>
</table>
assessment procedures and half reflecting non-standards-based assessment procedures to provide teachers with a full selection of the most commonly practised assessments utilised in the classroom. A high score on this scale indicates a high number of assessment procedures (SBI or non-SBI) being used by the teacher. There are twelve items on this subscale, with respondents provided with a 1 to 10 (‘very unfavourable’ to ‘very favourable’) rating scale for a possible score range of 12 to 120. The reliability of this subscale was found to be 0.91.

**Practice**

*Student observation form.* This form is used as an observation instrument within math and science classrooms by an external observer. The form was developed using information and adapted items from math and science education standards and the TIMSS survey (NCTM, 1989; NRC, 1996; TIMSS, 2005). Observers were trained to use the instrument to observe math and science classrooms in grades 6 to 9. The selected observers for the study are former Oklahoma City math and science teachers who had been selected to serve as coaches and were used throughout the project with the math and science classroom teachers. These coaches were utilised as classroom observers because they are not perceived as a threat or intimidating by the classroom teachers being observed. The observation form consists of four sections: (1) demographics, including the school, grade, time, teacher, and class being observed; (2) math standards-based activities/behaviours; (3) science standards-based activities/behaviours; (4) non-standards-based activities/behaviours that are observed in the classroom. Each of the three observation subscales contains eight items representing specific activities/behaviours. Selected examples of SBI math activities/behaviours include: (1) students using manipulatives, (2) students engaged in self-assessments, (3) students working in pairs or small groups. Selected examples of non-SBI activities/behaviours include (1) students listening to a teacher lecture, (2) students working on pencil/paper worksheets, (3) students taking tests and quizzes. Observers mark 1 if the item is observed and 0 if the item is not observed during the classroom observation period (approximately fifty minutes per period). The higher the score on each of the math and science standards-based subscales the higher the degree of SBI implementation in the classroom, whereas the higher the score on the third subscale (non-SBI) the lower the degree of SBI implementation. This form was pilot-tested prior to being used in the study and observers received formal training from the external evaluator in how to observe and use the instrument to establish inter-rater reliability. The internal consistency reliability coefficient for the instrument was found to be 0.78.

**Performance**

*Iowa Test of Basic Skills, forms K, L, and M* (Hoover et al., 1996). This norm-referenced test is given annually to Oklahoma City public school students in grades 6 to 12 to assess basic skills and cognition levels in reading,
language arts, mathematics, social studies, science, and sources of information. Composite and subscale scores for each of the test battery topics of math and science were utilised in this study as students’ achievement measures. Reliability and development information for the ITBS is available from the publisher.

Subjects
Approximately 10,000 Oklahoma City public-school students and 408 teachers in grades 6 to 9 (204 math and 204 science teachers) during the 2000–01 and 2001–02 school years from randomly selected math and science classrooms participating as subjects in the study. The Oklahoma City Public School District serves approximately 40,000 students in grades pre-K12, with over 85 per cent of the students qualifying for free lunch services and with the following racial/ethnic composition: white (29 per cent), black (36 per cent), Hispanic (27 per cent), Indian (5 per cent), and Asian (3 per cent). The randomly selected math and science classrooms of teachers and students in grades 6 to 9 for participation in this study parallel the demographic composition represented by the school district.

Procedures
The relational database, that is, the empirical connections of teachers’ preparation data with those same teachers’ classroom observation data and their students’ achievement data substantiate the threefold framework of the study: the three areas of teacher preparation, classroom practices, and student performance relative to the common goal of assessing SBI effectiveness. Data collection methods for each of the three areas of teacher preparation, classroom practices, and student performance using the instrumentation appropriate for each component delivered a cohesive database. Teacher (preparation) data were obtained from secondary math and science teachers in the district during summer professional development academies focused on training in SBI classroom implementation. Classroom observational data (practices) were obtained from over 400 randomly selected math and science classroom observation sessions conducted by the OK USP Partners in Excellence members. Separate types of math and science classes were included in the random sampling of classroom observation sessions to include student achievement (performance) data in the form of ITBS math and science scores provided by the school district’s research office.

Analysis of data procedures to examine interrelationships and specific contributions of SBI teacher preparation and practices (as well as non-SBI practices) on student achievement includes key inferential analyses involving stepwise multiple regression procedures. A stepwise regression procedure was selected to first utilise a full regression model, then sequentially identify variables for elimination (Christensen, 1998, p. 427). Specific SBI and non-SBI practices serve as the independent variables for the regression analysis,
with student achievement (math and science ITBS scores) as the dependent variables to determine significant ($p < 0.05$) contributions of individual math and science SBI practices and non-SBI practices to student achievement. Additional analyses of SBI and non-SBI practices using stepwise multiple regression analysis were performed relative to disaggregated data by student gender and ethnicity. Descriptive statistical analyses reflecting the degree of implementation of math and science SBI and non-SBI practices within the observed classrooms provide additional information.

**Results**

Highlights of the descriptive and inferential statistical analyses for the study include the following results. (1) SBI practices that were found to be significant contributors to students’ math achievement include the use of manipulatives, self-assessment, co-operative group projects, and computer technology. (2) SBI practices that were found to be significant contributors to students’ science achievement include the use of inquiry, self-assessment, co-operative group projects, and computer technology. (3) Virtually none of the observed non-SBI practices was found to be a significant contributor to student math or science achievement by gender or ethnic groupings. (4) The 408 math and secondary teachers in the study have each received over 200 hours of informal (consultation and lesson modeling) training in SBI and over 160 hours of formal training in SBI from 2000 to 2002.

The observed frequencies of occurrence of SBI and non-SBI in math and science classrooms ($n = 408$ observations, with 204 math classrooms and 204 science classrooms observed) revealed the following observed SBI and non-SBI frequencies in the combined classroom observations reflective of the data from the student observation form: (1) math SBI 17 per cent, (2) science SBI 36 per cent, and (3) non-SBI 47 per cent of the total observed activities.

Results of inferential analyses report contributions of specific SBI practices on student performance (achievement). Multiple stepwise regression analyses were performed using twenty-four independent variables (each of the eight math SBI practices, the eight science SBI practices, and the eight non-SBI practices observed in math and science classrooms) and math and science ITBS scores as the separate dependent variables to determine the specific contributions of each of the specific SBI practices on student achievement. Table 2 summarises the significant ($p < 0.05$) findings of these analyses. It depicts the multiple effect contributions of specific SBI practices to performance, i.e. secondary students’ math and science achievement ITBS math and science scores. The multiple regression analysis results identify three (use of hands-on materials, self-assessments, and projects-based activities) of the eight SBI math practices and two (self-assessments and use of computers) of the eight SBI science practices as significant ($p < 0.05$) contributors to students’ ITBS math scores. None of the non-SBI practices observed in the classrooms was found to significantly contribute to math or science achievement.
The effects of Standards-based Instruction on students’ math and science achievement disaggregated by gender and by ethnicity were also explored in the analysis, with the results indicating the following outcomes. The multiple effects of two SBI practices (manipulatives and self-assessments) contributed significantly \((p < 0.05)\) to female students’ math achievement, whereas the only significant \((p < 0.05)\) SBI practice found to contribute to male students’ math achievement was the use of calculators in the classroom. Both female and male students’ achievement levels were significantly influenced by the use of computers and self-assessments in science classrooms; however, none of the non-SBI practices contributed significantly to math or science achievement for either gender. The multiple effects of SBI practices (use of manipulatives and self-assessments) contributed 4 per cent of the variance in math achievement for white students and manipulatives contributed 2.5 per cent of the variance in math achievement for minority students (black and Hispanic combined). Significant \((p < 0.05)\) SBI contributors to white students’ science achievement are identified as the multiple effects of self-assessments and inquiry-based activities, whereas the multiple effect SBI contributors to minority students’ science achievement are identified as the use of computers and self-assessments. The only non-SBI practice found to contribute significantly (5 per cent, \(p < 0.05\)) to science achievement is lecture, found only for white students.

The study revealed the following summary findings. (1) Although substantially more non-SBI activities were observed in the secondary math and science classrooms participating in the study than the frequency of occurrence of SBI practices, virtually none of the non-SBI practices was found to significantly contribute individually or in multiple effects to students’ math or science achievement. (2) Virtually none of the non-SBI practices produced significant individual or multiple effects on students’ math or science achievement by gender or ethnicity, with one exception, that is, teacher lecture was found to contribute significantly to white students’ science achievement. (3) The use of manipulatives as a significant key contributor to students’ math achievement was identified for all students, regardless of gender or ethnic affiliation. (4) The use of self-assessments for students in the classroom was found to contribute to science achievement for all students, regardless of gender or ethnic affiliation. (5) The use of computer
technology in science classrooms was identified as a key contributor to both male and female minority students’ science achievement. (6) Co-operative learning-projects-based activities used in mathematics classrooms were identified as a significant contributor (from multiple effects analysis) to students’ math achievement. (7) The use of inquiry-based projects and activities in science classrooms was found to be a significant contributor (from multiple effects analysis) to white students’ science achievement.

Discussion

The results of this study provide empirical evidence in support of specific SBI practices for math and science education. Findings support the views of reform organisations regarding the need for a re-examination of non-standards-based practices currently still dominating many math and science classrooms in American schools. The use of lecture, independent seat work, quizzes, and text homework still preponderates in many math and science classrooms. This study provides evidence in support of standards-based practices such as inquiry, problem solving, co-operative learning, and use of hands-on and technology in math and science classrooms as significant contributors to student achievement. Results add some evidence to the need for females and minorities to utilise self-assessment, hands-on materials, and co-operative projects-based activities as effective standards-based practices in both math and science education. These results corroborate Johnson’s (2002) indicators identifying standards-based reform efforts as key predictors of achievement gains in mathematics and science.

The findings inform the data-driven decision-making needs of schools grappling with strategic considerations of raising student achievement and increasing teacher quality (NCLB, 2000). The Preparation, Practice, and Performance (P³) model utilised by this study also highlights the somewhat neglected practice component in the model. The common belief that increasing teacher quality (preparation) will raise student achievement (performance) even though virtually no effort or focus has been aimed toward assessing the degree of improvement in classroom implementation (practice), specifically SBI practices, cannot be assumed. The notion that highly prepared teachers will propagate high-performing students without assessing and carefully monitoring the degree of implementation of SBI activities that occur in classrooms is an inaccurate assumption (Johnson, 2002).

Results of this study also suggest the P³ model Preparation, Practice, and Performance as a viable framework for empirically validating the effectiveness of Standards-based Instruction as a successful reform effort for systemic change in math and science education. The empirical evidence produced by this study provides rigorous support for specific SBI practices as key contributors to students’ math and science achievement and refutes non-SBI practices as effective contributors to students’ math and science achievement.

In addition, the study findings lend empirical support for examining instructional strategies (practice), student achievement (performance), and
teacher professional development (preparation) as an effective process to improve learning and teaching in math and science education.

References


Oklahoma City Public Schools (2000), *Oklahoma City Public Schools Urban Systemic Program*, Oklahoma City OK: OCPS.


Address for correspondence

Department of Professional and Community Leadership, Building 77, Office 146, College of Professional Studies, University of West Florida, 11000 University Parkway, Pensacola, FL 32514, USA. E-mail cthompson1@uwf.edu