Abstract Body

Background / Context:
Math Monster Mystery (M3) is a formative assessment research project consisting of a suite of 23 math games designed to address fourth-grade students’ poor performance on national and international mathematics tests. It also addresses the need for teachers to have easy access to quick and accurate data in order to diagnose student problems, monitor student progress, and improve and differentiate classroom instruction.

Fourth grade mathematics scores reported by the 2007 Trends in International Mathematics and Science Study (TIMSS) and the National Assessment of Educational Progress (NAEP) showed that U.S. fourth grade TIMSS scores ranked 9th of 35 countries. Although the trend is upward, mathematical performance remains a national concern. Similarly, only 39% of fourth graders scored at or above the proficient level on the 2007 NAEP. Although scores for all demographic groups improved on the NAEP, there still remains significant difference between scores and progress of minority and disadvantaged groups compared to white students.

Research shows that one effective means of increasing students’ performance is frequent assessment tasks that help teachers diagnose and monitor student progress so they can adapt their teaching accordingly (Nicol & Milligan, 2006). Bloom’s ideas on mastery learning states that teaching and learning consist of three major components—standards, instructional alignment, and feedback, though he found that most teachers included little variation in their instructional practices (Guskey, 2005). His conclusion that increased variation in teaching is required to improve achievement has been supported by a number of studies (Brimijoin, Marquissee & Tomlinson, 2003; Nicol & Milligan, 2006; Stiggins, 1999). M3 is designed to help teachers achieve this end.

Purpose / Objective / Research Question / Focus of Study:
The overall objective of the project has been to improve the mathematical performance of fourth-grade students by creating a formative assessment in the form of an online game.

Major goals of the project have been to:
• Develop a game environment with multiple mini-games that cover the concepts and align with National Council of Teachers of Mathematics (NCTM) standards and focal points for fourth grade mathematics.
• Construct a back-end database that collects data on each student’s performance and prints out clear, concise, and easily accessed reports for the teacher.
• Create a tool that allows the teacher to customize the assessment for individuals or groups of students by selecting from the menu of mathematical mini-games.
• Complete verification study to collect qualitative data.

Setting:
The project has been field tested in public school classrooms and in home school settings.

Population / Participants / Subjects:
Two phases of testing have been completed—Phase I prototype testing and Phase II verification testing of a completed product. In phase I, the sample consisted of three elementary schools in Guilford County, NC, and students were tested only on an early prototype of the program. Two fourth grade classes participated from School one, a
predominantly white school of 676 students. Three fourth grade classes participated at
School two (420 students) comprised of an ethnically mixed—African American and
Hispanic—student body. Two fourth grade classes participated from School three, an
ethnically mixed school of 400 students.

Phase II verification testing is underway with 1100 public school and home school 4th
and 5th grade students and will be completed by May 15. Data has been collected and
analyzed for 174 of the 1100. 157 of the 174 were 5th graders, and 17 were 4th graders,
ranging in age from 9 to 11 years. The table below describes the 174 student reviewers by
gender and race.

<table>
<thead>
<tr>
<th></th>
<th>White</th>
<th>Latino</th>
<th>Asian</th>
<th>African-American</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>103</td>
<td>81</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Boys</td>
<td>80</td>
<td>56</td>
<td>10</td>
<td>2</td>
</tr>
</tbody>
</table>

**Intervention / Program / Practice:**

Our team has designed and developed a computer adapted, web-based, formative
assessment game—Math Monster Mystery (M3)—that both teaches and measures the
mathematical knowledge and understanding primarily of 4th graders (the program has a
range of third to fifth grade because of its numerous difficulty levels). The heart of M3 is
a series of mini-games that are integrated into an engaging multi-media game narrative.
Twenty-four content knowledge targets are delineated into procedural skills and
conceptual understandings, and the mini-games are designed with 6 difficulty levels. The
M3 system knows how a student is progressing and uses that information to make real-
time adjustments by altering the difficulty level of tasks. The power of the tool is the
combination of feedback, adaptive task sequencing, and assessment of knowledge and
skills.

The procedural skill mini-games are fast moving: a problem appears on the screen,
and the player clicks the correct answer. This approach allows the learner the repetition
requisite for procedural fluency. M3 also incorporates problem-solving or conceptual
mini-games designed to model specific conceptual ideas or logical intersections of the
procedural games.

Computer assessment games allow what paper and pencil tests do not: saving
multiple states and returning to the previous ones; bolstering the strengths and identifying
the weaknesses of students; and keeping an assessment history. M3’s dashboard is an
easy-to-use reporting tool that provides logging and tracking that is rarely if ever found in
formative assessment products. The dashboard allows sorting of performance data by
student, mini-game, item analysis, and mastery level. The teacher can customize the
game to meet the student and program learning needs. Aggregate data is collected and
used to inform instruction and program curriculum decisions.

M3 is aligned with NCTM standards, is reliable and valid, and provides teachers with
immediate information about student performance. The M3 reports are easily accessible
from the web immediately following the student’s attempt. Data may be sorted by
concept for an individual student or clustered by groups of student who are performing at
similar levels. Using this information and a simple backend tool, the teacher can easily
customize future M3 game play to maximize learning. For example, if 25% of the class is
experiencing difficulty with multiplication fluency while studying in the algebra section
of the curriculum, the teacher can create a customized M3 collection of games for those
students to include mini-games to practice multiplication and algebra concepts. Since
75% of the class is not experiencing the difficulty with multiplication, there is no reason to include it in their formative assessment experience.

This program was developed over a 30-month period with SBIR funding from the U.S. Department of Education.

**Research Design for Phase III Testing:**

Phase III testing for M3 is being planned. In this phase, our team will implement a substantive quantitative and qualitative testing protocol that is designed to span the academic year with a significant sample size. This is the next step in establishing the validity and reliability of the tool for improving student mathematical performance. The full-scale testing will also seek to determine M3’s effectiveness and usability as a teacher tool for differentiated classroom instruction. A preliminary description of the desired field study is below.

The full-scale testing will implement a randomized crossover field trial design and comparison to a control group. Item analyses (item difficulty and p-values) will be conducted for each of the mini-game assessments. Comprehensive pre- and post-tests using computer-adaptive testing formats will be developed so that students’ achievement levels may be determined without causing students to sit through overly long assessments. Teachers will participate in focus groups regarding their use of the M3 mini-game modules and the program’s effectiveness as formative assessment. Other features of M3 to be assessed will include: the feasibility of incorporating it into their classroom practice, the use of the game as a practice tool for students, suggestions for improving the dashboard or the game, and what they liked most about the game. The analysis will also determine if there was a systematic relationship between level and type of teacher use of the M3 assessment program and students’ mathematic achievement levels. Similar analyses will be conducted to examine student practice factors and achievement. A sample of students will be interviewed about their use of M3, their attitude toward it, and how it might be improved. Teachers will be asked to select a representative sample of students for these interviews, based on their ability and ethnicity. At the end of the program, all students will be surveyed about their use of the games.

**Data Collection and Analysis:**

In the full-scale testing with all 24 mini-games students’ performance will be tracked by software embedded in the system. Surveys and interviews will complete the data collection for the study.

**Findings / Results:**

There are three phases of testing, but only two preliminary field tests have been implemented, Phase I Prototype Testing and Phase II Verification Testing. Phase I occurred at the end of the prototype development, as part of a Phase I SBIR contract with the U.S. Department of Education; only two mini-games had been completed for testing at that point. As noted previously, three elementary schools participated, and teachers as well as students responded to survey questions. Results to Phase I research questions are below:

- Data from the teacher survey showed that 100% of the teachers felt that the dashboard reports were helpful in adapting their instruction to fit students’ performance although the time constraints of the study did not allow for extensive use of the tool. All of the teachers implemented the game during regularly scheduled computer lab times and
50% of them also used the game in the classroom for extra individual and group practice.

- Students indicated that the design of the two mini-games, which were on multiplication and fractions, was appropriate for their grade level and was easy to use. All 23 students interviewed reported that they particularly liked the narrative, graphics, and game play.
- Students indicated they thought that the program helped them learn multiplication and fractions, and several students mentioned that they found the game a fun way to practice math.
- Data from the pre- and post-assessment that were submitted to a t-test for independent samples indicated no significant differences between the two groups for either school.
- Teacher indicated that the game was self-motivating and self-explanatory, so students could work at their own pace. The teachers felt the repetition was important for improved learning.

**Phase II Verification Testing:**
Verification testing is underway with approximately 1100 students and will be completed by May 15. Preliminary data collected in the verification phase of testing from both teachers and students has been very positive. Overall they have rated the game high in terms of content and ease of use. Data has been collected and analyzed for 183 of the 1100. Of the 183 fourth- and fifth-grade students initially tested, 174 filled out review sheets, and the large majority of those indicated that M3 assisted them in learning math concepts, and most reported they found the game enjoyable. Teachers have been enthusiastic about using it as a formative assessment tool and reported they thought it would integrate well into their curriculum and classroom.

- Of the 174 students completing M3 review sheets, 54% rated it as Excellent, with another 28% rating it as Good; 13% rated it as Fair, and 5% rated it as Poor.
- When asked about how they thought the game fit with what they were learning or had learned in 5th grade math, 99 (57%) said that it “fits really well,” 56 (32%) said it “fits somewhat,” and 19 (11%) responded that it “really doesn’t fit.”
- Students were asked if M3 helped them learn math concepts. Of those responding, 142 responded, “Yes” and 32 responded, “No.”
- When asked about who they thought would really like to play the game, common student answers included “4th or 5th graders,” “students,” and “elementary students.”
- Of the 174, only 14 reported any technical difficulties.

This qualitative data from public school students, home schoolers, expert teachers, internal testers, and math professionals will be used to inform updates and changes to the mini games as we move to the full-scale validation testing.

**Conclusions:**
Qualitative data has been very positive in favor of M3 as an effective addition to the 4th grade mathematics classroom; we feel confident that the phase II verification study will provide evidence for continued study of the intervention. Conclusions and recommendations with regard to its effectiveness for improving mathematics performance will require the implementation of full-scale testing as described above.
Appendix A. References


