



## **Research Memorandum**

ETS RM-12-01

# **Designing and Evaluating an Interactive Score Report for Students**

---

**Margaret Vezzu**

**Waverely VanWinkle**

**Diego Zapata-Rivera**

**January 2012**

## **Designing and Evaluating an Interactive Score Report for Students**

Margaret Vezzu, Waverely VanWinkle, and Diego Zapata-Rivera  
ETS, Princeton, New Jersey

January 2012

As part of its nonprofit mission, ETS conducts and disseminates the results of research to advance quality and equity in education and assessment for the benefit of ETS's constituents and the field.

To obtain a PDF or a print copy of a report, please visit:

<http://www.ets.org/research/contact.html>

**Technical Review Editor:** James E. Carlson

**Technical Reviewers:** Irv Katz and Debbie Pisacreta

Copyright © 2012 by Educational Testing Service. All rights reserved.

ETS, the ETS logo, and LISTENING. LEARNING. LEADING.,  
are registered trademarks of Educational Testing Service (ETS).  
CBAL is a trademark of ETS.



## **Abstract**

In an effort to make students active participants in their learning and to consider students' score reporting needs, we have designed an interactive student score report for middle school students that implements a guided-instructional activity aimed at facilitating student understanding of score report information and improving student engagement. This paper showcases an interactive student score report developed for the Cognitively Based Assessment of, for, and as Learning (*CBAL*<sup>™</sup>) initiative and reports on results from an initial usability study, as well as revisions made to the original version of the score report based upon results from the initial usability study and expert feedback. In addition, results are presented from a follow-up usability study comparing the graphical representations in Version 1 of the score report to Version 2 of the report.

Key words: student score reports, middle school, interactive, usability study, student engagement, score report design principles

## **Acknowledgments**

We would like to thank the *CBAL*<sup>™</sup> staff: Janet Stumper for helping us design and create the student score reports, Jeff Haberstroh and Liz Marquez for providing suggestions and feedback, and Lynn Zaback and Randy Bennett for supporting this work. We would also like to thank Ronald K. Hambleton (University of Massachusetts Amherst), Priti Shah (University of Michigan), Howard Wainer (National Board of Medical Examiners), and Rebecca Zwick (ETS) for participating in a score reporting conference held at ETS and for reviewing and providing feedback about the score report. We would like to extend our gratitude to the students who participated in our studies. Finally, we would like to express appreciation to Randy Bennett, Terry Egan, Irv Katz, Christy Lyon, Debbie Pisacreta, and Cindy Tocci for their comments and suggestions.

## Table of Contents

	Page
Related Research.....	2
Feedback .....	3
Multimedia.....	3
Student Engagement and Motivation.....	5
Interpretation of Score Reports.....	5
Interactive Student Score Report .....	7
Version 1 .....	7
Initial Usability Study.....	12
Version 2.....	13
Follow-Up Usability Study .....	18
Method.....	18
Results.....	19
Expert Review.....	22
Summary and Future Work.....	22
References.....	23
List of Appendices .....	26

A review of commercially available score reports showed that most student score reports are aimed at parents (Underwood, Reshetar, & Leahy, 2006). Parents want to know the student's overall score, the passing score or cut scores for different proficiency levels, how the student's score compares to other scores, whether the student made progress in different areas, and specific recommendations for helping their child. Although this information is important for both parents and students, students have played a passive role in the design of score reports, which results in student score reports that look similar to those provided to teachers, with some minor modifications (e.g., language employed).

Students' lack of sufficient knowledge about tests and statistics, low levels of motivation to examine or use score reports, and passive approach to learning make them a particularly unique audience for score reports. As a result of these factors, score reports for students need to be designed so that they are engaging and motivating.

A review of existing score reports for students shows that current student score reports are usually static PDF documents that include technical terms that students do not understand (Deng & Yoo, 2009). In addition, score reports are usually available at the end of the academic year, which limits their use for guiding student learning. It is not surprising that students, who are accustomed to highly interactive communication and entertainment tools, find these score reports unattractive, disengaging, and somewhat disconnected from their learning process.

Score reports should communicate assessment information to particular audiences effectively, but they should also foster communication among teachers, students, and parents. For example, students and teachers can use score report information to engage in an instructional dialogue aimed at increasing students' knowledge and responsibility for their own learning. At the same time, teachers will gather additional evidence about the students' knowledge in particular domains. Teachers can also use score report information to share student performance information with parents or guardians and help them become more involved in the learning process (Zapata-Rivera, Underwood, & Bauer, 2005).

Interactive computer applications that provide immediate, task-level feedback can positively contribute to student learning (e.g., Hattie & Timperley, 2007; Shute, 2008). Well-designed interactive multimedia can also support student motivation (e.g., McCormack, 2002; Neo & Neo, 2009) and engagement, (e.g., Bottino, Ferlino, Ott, & Tavella, 2007; Cordova & Lepper, 1996; Mandinach & Corno, 1985; McDonald & Hannafin, 2003) as well as enhance

student learning (e.g., Avitabile, 1998; Kennedy & McNaught, 1997; Mayer, 2001; Reeves, 1998; Regan & Sheppard, 1996).

We have designed an interactive student score report for middle school students. This score report was created using a framework for designing and evaluating score reports that includes the following steps: (a) gathering assessment information needs, (b) reconciling these needs with the available assessment information, (c) designing various score report prototypes, and (d) evaluating these report prototypes internally with experts and externally with the target audience (Zapata-Rivera & VanWinkle, 2010). This work is part of the Cognitively Based Assessment of, for, and as Learning (*CBAL*<sup>TM</sup>) initiative (Bennett & Gitomer, 2009). The CBAL initiative includes formative and summative assessments and professional development for the areas of math, reading, and writing.

The interactive student score report implements a guided-instructional activity aimed at facilitating student understanding of score report information and improving student engagement through the incorporation of game-like features. In the current version of this guided-instructional activity, students use a tabbed menu to navigate through the different sections of the score report and collect coins by correctly answering questions about the content of the report. A virtual character guides the students through the score report and provides feedback on their answers. After they have explored all of the sections of the score report, students write about their performance and propose an action plan. This paper summarizes related research, describes two different versions of the interactive student score report that were designed and evaluated following our framework, and reports on results from usability studies with students and a focus group with experts.

### **Related Research**

In order to effectively communicate assessment information to students, score reports need to engage them in an activity that encourages them to understand the contents of the report and use this information to guide their learning process. These new types of score reports should not only communicate assessment information clearly, but also support student motivation and encourage student responsibility for their own learning. Research findings in areas such as providing feedback to students, developing effective multimedia systems, supporting student engagement and motivation, and exploring user interpretation of score reports can inform the



development of these score reports. Next, we present relevant resources and findings in these areas.

## **Feedback**

Hattie and Timperley (2007) provided a conceptual analysis of feedback. They discussed how the type of feedback (e.g., positive and negative) and the way it is provided (e.g., timing of feedback) influences its effectiveness. Shute (2008) reviewed research on formative feedback and identified the characteristics of effective formative feedback for specific situations (e.g., multidimensional, nonevaluative, supportive, timely, specific, and credible), in addition to presenting a list of guidelines for generating feedback.

Black and Wiliam (1998) reviewed literature on classroom formative assessment and showed evidence that supports providing frequent feedback to students in order to improve student learning. In this approach to formative assessment, students are considered active participants, working with teachers to implement instructional strategies that make use of assessment information to inform instruction in the classroom. Research in the area of open student models shows that student knowledge, awareness, and self-reflection can be enhanced by providing students with performance and progress information that is maintained by intelligent systems in the form of student models (e.g., Brna, Self, Bull, & Pain, 1999; Bull & Pain, 1995; Zapata-Rivera & Greer, 2003).

## **Multimedia**

Research related to multimedia applications for learning provides insight into creating and designing score reports for students. This research also provides useful information on motivating students and creating effective learning experiences by implementing specific multimedia design principles.

In a study aimed at exploring the influence of student learning style and presentation mode on student learning, Avitabile (1998) found a significant, positive effect of a multimedia presentation mode on student learning when compared to a control group. Regan and Sheppard (1996) reported on the benefits of multimedia enhanced coursework in motivating students and providing them with opportunities for discussion and use of new vocabulary. They also found that combining multimedia and hands-on exercises can result in powerful learning experiences.

Neo and Neo (2009) reported on the positive results of using a multimedia project in a constructivist-based learning environment. Students experienced high levels of motivation and self-esteem. The authors recommended using these kinds of learning environments as a means of providing teachers with flexible options for authentic environments that they can apply in their classrooms.

Mayer (2001) presented a series of principles for developing effective multimedia applications based on empirical evidence. These principles include the following:

- *Various representations*: Students learn better when both words and pictures are presented than when only words are used; when both words and pictures are used, students have a chance to create verbal and pictorial mental models and to make links between them.
- *Spatial contiguity*: Students learn better when related words and pictures are placed near each other on the page or screen; this way learners do not need to dedicate cognitive resources to scan the page or screen, and the likelihood of keeping the words and pictures in working memory will increase.
- *Coherence*: Students learn better when extra, nonrelevant material is excluded; extra material causes competition for limited cognitive resources and can be distracting.
- *Redundancy*: Students learn better when only given animation and narration than when given animation, narration, and text that is the same as the narration; pictures and written words share the same visual channel, which can cause cognitive overload.
- *Individual differences*: Effects of design are stronger for learners with low levels of knowledge than for learners with high levels of knowledge; students with high levels of knowledge can rely on previous knowledge to make up for a lack of guidance in poorly designed presentations, while students with low levels of knowledge cannot. In addition, high spatial students have the ability to combine visual and verbal content from a multimedia presentation, while low spatial learners may not have the same ability.

## **Student Engagement and Motivation**

Research on student engagement and intrinsic motivation provides valuable information that can be used to inform the design of score reports for students. Malone and Lepper (1987) described four intrinsically motivating factors for individual learning activities:

- *Challenge*: A learning activity should provide the learner with an intermediate level of challenge and should have a goal and an uncertain outcome. The activity should also provide the learner with frequent positive feedback that fosters self-esteem.
- *Curiosity*: An activity must have an “optimal level of informational complexity” in order to stimulate curiosity in the learner. Both sensory and cognitive curiosity play a role in intrinsic motivation. A learning activity can promote sensory curiosity through audio and visual effects. Presenting subjects that the learner already has an interest in can increase cognitive curiosity.
- *Control*: In order to provide the learner with a sense of control over the activity, the learner’s responses and actions should impact the outcomes and create “powerful effects.” Providing learners with choices in an activity can also increase intrinsic motivation.
- *Fantasy*: Fantasy elements, such as those found in computer games, also play an important role in intrinsic motivation. Malone and Lepper propose that fantasies that are integrated with the content being learned are more engaging and educational than fantasies where the content being learned is not dependent on the fantasy. Fantasies should meet the learner’s emotional needs, which can be achieved by presenting imaginary characters that the learner can relate to. Additionally, fantasies should appeal to cognitive factors by associating content to be learned with relevant metaphors or analogies.

## **Interpretation of Score Reports**

Hattie’s (2009) paper discusses 15 principles that can be employed by designers to help users with score report interpretation. The author points out that providing more explanations and descriptions may not help a person interpret a score report; it may in fact cause the user to ignore the information. Some ideas to keep in mind when designing a succinct score report are minimizing scrolling, using color in a purposeful way, and providing only the necessary

information. It is also beneficial to provide users with a report that provides the most important information first, followed by supporting information and details. This will help users follow a path created by the score report developer when interpreting the score report. The author also found that minimizing the amount of numbers and considering the amount of information working memory can handle are both important when designing score reports. Providing too much information on a score report can overwhelm a user. Keeping with a theme for each score report can also help reduce the amount of information on the score report. It is also important that the score report addresses specific questions and that the designer has expected interpretations in mind beforehand. Finally, Hattie states that the meaning and constraints of the score report need to be communicated to the users, as well as the message that score reports are to be utilized not “printed.”

Goodman and Hambleton (2004) presented a review of student test score reports and guides from 11 states, two U.S. commercial testing companies, and two Canadian provinces. They noted promising features of these current reports and made several additional recommendations for designing effective student score reports that can be easily interpreted. Promising features noted by Goodman and Hambleton include features that make the reports more readable, such as headings, organizational devices, a highlights section, and graphs for major findings, as well as features that add meaning for users, such as descriptions of strengths/weaknesses, reporting comparison groups, and reporting in relation to performance levels. Goodman and Hambleton also make the following recommendations:

- Include all information that is essential for proper interpretation of assessment results in score reports, such as purpose, interpretation of results, a description of performance levels, scores, and examples for confidence bands.
- Include detailed information about assessment and score results in a separate interpretive guide.
- Personalize the student score reports and guides by including the student’s name and a letter to parents.
- Include an easy-to-read narrative summary of the student’s results at the beginning of the report. Highlight overall results, diagnostic information, and implications.

## **Interactive Student Score Report**

Our interactive student score report is a game-like activity where students answer questions about different parts of the score report. We hope that these questions will motivate students to pay closer attention to the contents of the score report. An initial version of the interactive report was created following some of the design principles discussed above. This version implemented an opening doors metaphor in which students answered questions to open virtual doors that uncovered different parts of the score report. This version was evaluated and revised to create a second version that implements a coin collecting activity. In the second version, students collect coins by answering questions about each section of the report and use a tabbed menu to navigate through the sections of the score report.

We designed these score reports with specific design and score report interpretation principles in mind. By creating a student score report designed as a game-like activity with fantasy elements, we have taken into account several of Malone and Lepper's (1987) principles for designing intrinsically motivating learning activities. The use of questions about the score report content provides the element of challenge along with the goal of collecting coins by answering the questions correctly.

In keeping with Hattie's (2009) principles for helping with score report interpretation, we aimed to design a student score report that was succinct and clear. Links to definitions, information about question types and tasks in the test, and sample questions eliminate the need for scrolling.

Next, we present the initial version of the interactive score report and the results of a usability study with eight students.

### **Version 1**

The original version of the score report, shown in Figures 1-6, includes doors to different sections of the report that are opened by answering a series of questions that assess students' understanding of the different sections of the score report. (These questions can be modified to support a different goal.) As students answer the questions, a character provides immediate feedback based on their responses (see Figure 1). This is in line with Malone and Lepper's (1987) recommendation to provide frequent positive feedback. The full score report is revealed to students before they answer the last question, as shown in Figure 2. After uncovering the score report, students write about their performance and propose an action plan (to improve their performance).

**CBAL Student Interactive Score Report**

**YourSpace**  
 Student: You  
 Teacher: Brown  
 Subject: Math  
 Grade: 8  
 Test : 3 of 3  
 Test Date: 03/30/10

**What is the Purpose of This Report?**  
 To provide you with your performance on the current as well as past tests that you have taken.

**How Should You Use This Report?**  
 To identify your strengths and weaknesses and use this information to determine where you may need help or more practice.

**How Did You Do On All the Math Tests?**  
 Your mathematics score, 240\*, is at the Proficient level.

\*Your score and performance level are based on how you did on ALL 3 tests.  
 \*\*Tests are not perfectly accurate. The confidence band shows where your score could have fallen (More Information).

**6: Overall Summary**  
 That's right. Your next performance level would be Advanced.

**Question Set 2 - (3 of 6)**  
 Q3. What is the next level that you can progress to?  
 Not yet at Basic  
 Basic  
 Proficient  
 Advanced

Figure 1. Character providing immediate feedback.

**CBAL Student Interactive Score Report**

**YourSpace**  
 Student: You  
 Teacher: Brown  
 Subject: Math  
 Grade: 8  
 Test : 3 of 3  
 Test Date: 03/30/10

**What is the Purpose of This Report?**  
 To provide you with your performance on the current as well as past tests that you have taken.

**How Should You Use This Report?**  
 To identify your strengths and weaknesses and use this information to determine where you may need help or more practice.

**How Did You Do On All the Math Tests?**  
 Your mathematics score, 240\*, is at the Proficient level.

\*Your score and performance level are based on how you did on ALL 3 tests.  
 \*\*Tests are not perfectly accurate. The confidence band shows where your score could have fallen (More Information).

**How Did You Do On All the Tests?**

- You correctly answered questions that require creating, interpreting and using different kinds of data displays.
- You correctly answered questions that require understanding algebraic expressions.
- You provided some well-justified statements to support conclusions.

**To Progress to the Advanced Level the Student Needs To:**

- Solve problems that require the use of functions and relations.
- Evaluate expressions and provide well-justified statements to support conclusions.

**What You Did on Test 3 (Proportional Reasoning & Statistics)**

	Task 1	Task 2	Task 3
	This task includes questions about proportional reasoning as well as creating, interpreting and using different data displays.		
Description of Performance	Number Correct	Content Skills	Process Skills
Question Type 1	4 out of 4 questions	Proportional Reasoning: Estimate and Approximate	Model & Represent
Question Type 2	5 out of 6 questions	Statistics: Use and Interpret Data Displays	Argue & Justify
Question Type 3	3 out of 3 questions	Statistics: Use and Interpret Data Displays	
Question Type 4	2 out of 4 questions		Argue & Justify

**Sample Question**

Click here

Click here

Click here

**Summary of Your Overall Performance on Tasks 1, 2, and 3**

You correctly answered questions that require understanding and applying estimation correctly using pictorial representation.

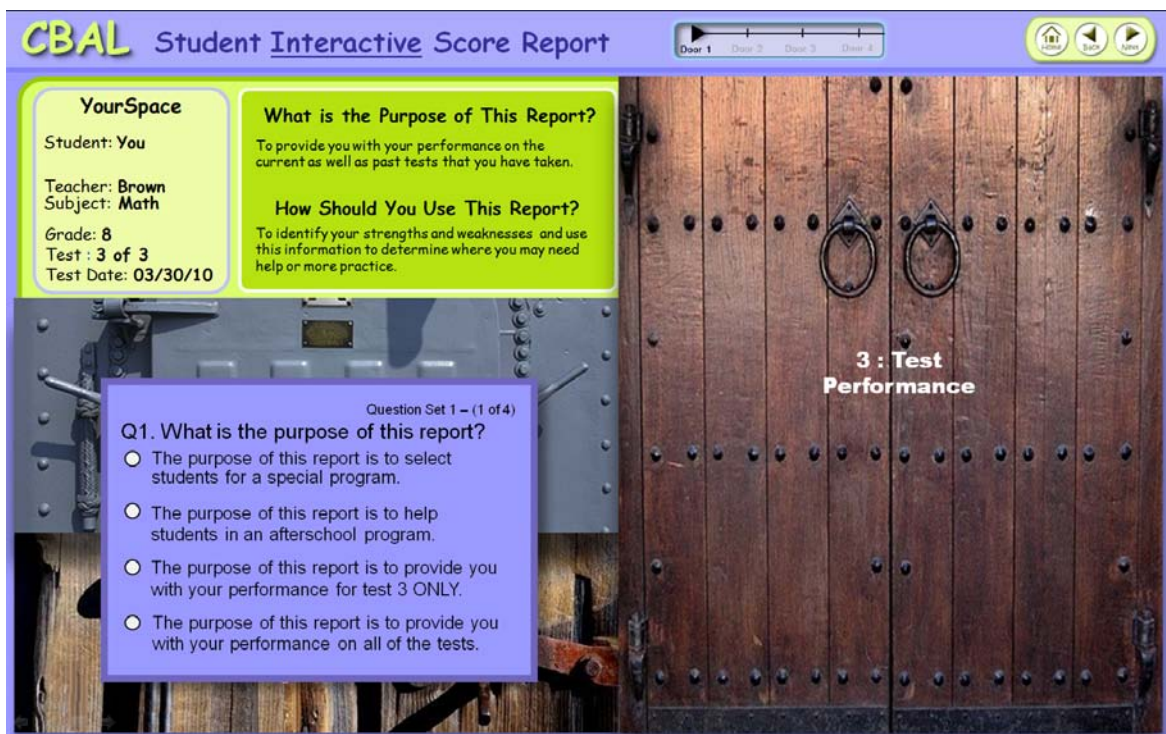
You correctly answered questions that require creating, interpreting and using different kinds of data displays, particularly scatterplots and trend lines.

You seem to require additional work on providing well-justified statements to support conclusions based on data.

Figure 2. Uncovered score report.

This activity is aimed at facilitating student understanding of score report information, improving student engagement, and helping students to identify possible next steps (e.g., read about or ask the teacher for help with a particular topic that requires improvement).

The score report design follows Hattie's (2009) principle of providing the most important information first, followed by details and supporting information. This is seen in the order of the score report sections. Following Goodman and Hambleton's (2004) recommendation to personalize score reports, the first section, shown in Figure 3, includes identifying information about the student, such as the student's name, teacher, and grade, as well as the subject and test that the score report information is from. This section also provides information about the purpose and use of the score report, which as noted by both Hattie and Goodman and Hambleton, is essential to help the student understand the meaning of the report and correctly interpret the score report.



**Figure 3. Purpose and use.**



Section two of the score report, shown in Figure 4, indicates the student’s score and performance level based on all of the tests that the student has taken. In accordance with Goodman and Hambleton’s (2004) recommendation, this section provides information that helps students better understand the meaning of their scores, such as links to the definition of a performance level and information about error and confidence bands.

The third section, shown in Figure 5, provides detailed information about the student’s performance on each task within a test, and includes some of the features that Goodman and Hambleton (2004) described as adding meaning for users (e.g., highlights of overall results and descriptions of strengths and weaknesses). The following information is displayed for each question type within a task: a description of performance, number of correct questions, skills used for the task, and a sample question. This information helps students to understand their performance on particular tasks.

The final section, shown in Figure 6, provides a summary of the student’s overall performance on both the current and past tests and tells the student what he or she needs to be able to do in order to advance to the next performance level.

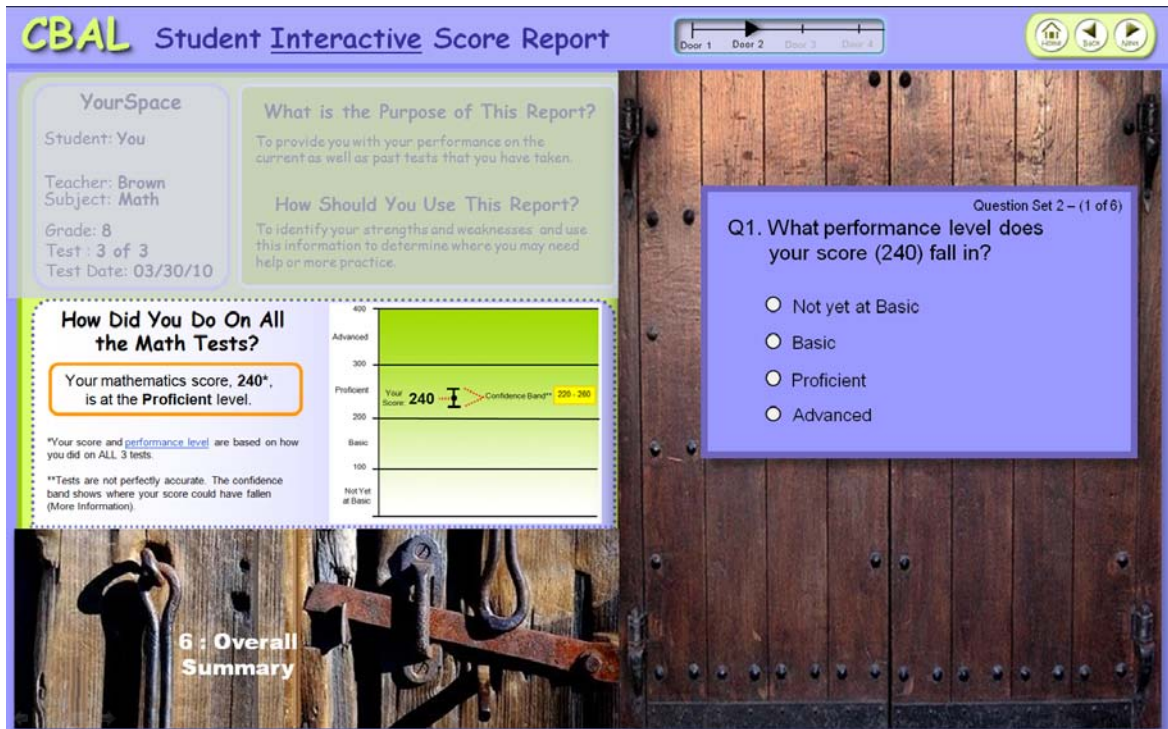


Figure 4. My math performance this year.



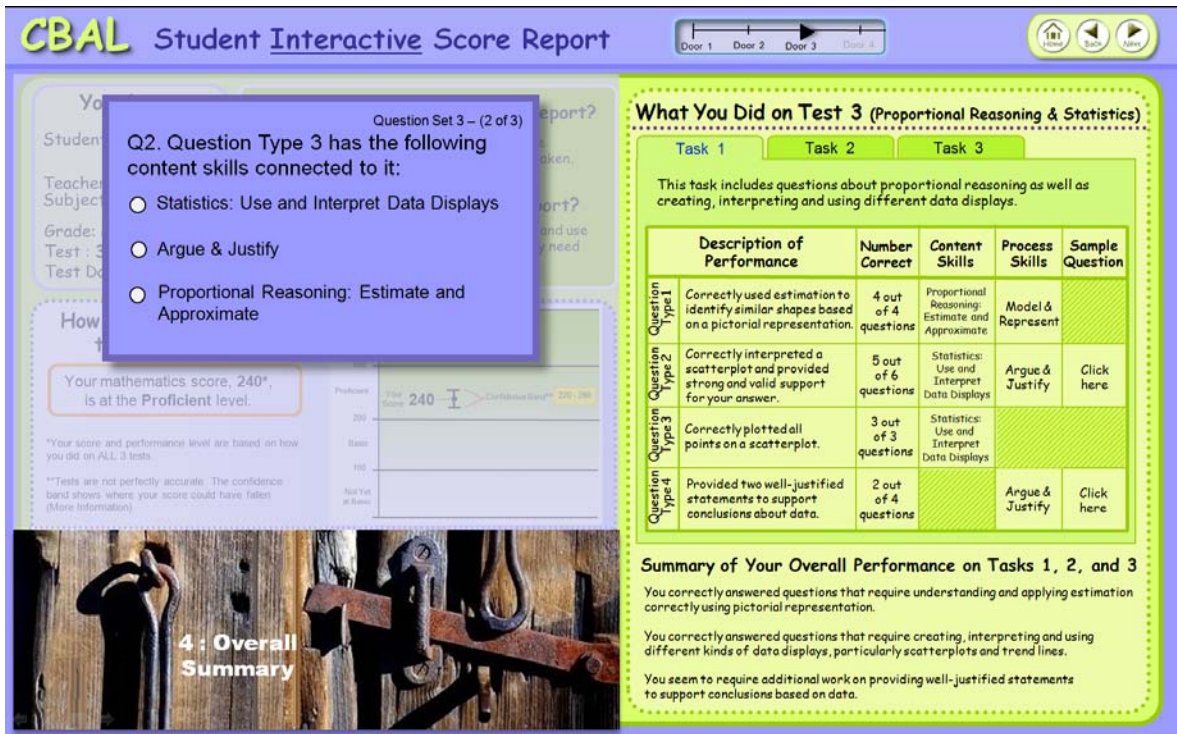


Figure 5. Test performance.

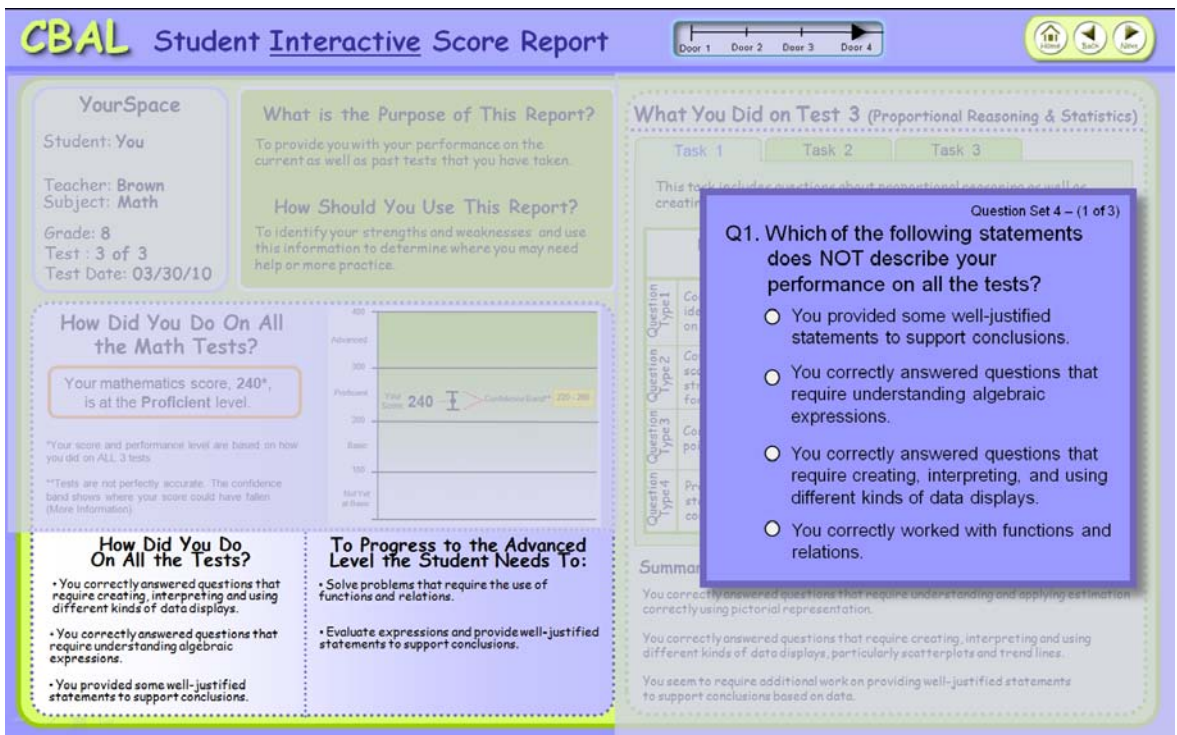


Figure 6. Overall summary.

The figures in this paper show the online version of the score report. When the activity is fully implemented, students will have the option to print their score reports. This printed version will display differently than the online, interactive version (e.g., interactive components could be replaced by static representations or removed from the report).

### **Initial Usability Study**

An initial usability study was conducted with eight local middle school students to identify major issues with content and clarity. These usability studies were conducted in order to obtain feedback from the actual audience that will be using the score reports. Information about the participants' backgrounds is presented in Table 1. Students were asked to think aloud while viewing the score report and answering the embedded questions in the report. They also completed a short background questionnaire about their experiences with computers and score reports, as well as a usability survey about the score report. These instruments are presented in Appendix A and Appendix B, respectively.

Results of this study suggest that students find the activity engaging and the contents of the score report clear. Most of the students were able to explain statistical terms such as *confidence band* with the help of information from the score report. Students made suggestions aimed at improving the look and feel of the score report (e.g., using popular cartoon characters and raising the prominence of the cartoon character). Students also appreciated the opportunity to share their own improvement plan with the teacher or parents/guardians (Zapata-Rivera, 2009).

**Table 1**  
*Participants' Background*

ID	Gender	Race	Grade
001	Male	Asian	8
002	Female	Asian	8
003	Male	Caucasian	7
004	Male	Asian	8
005	Male	African-American	7
006	Male	Asian	8
007	Female	Caucasian/Asian	8
008	Female	Caucasian/Asian	8

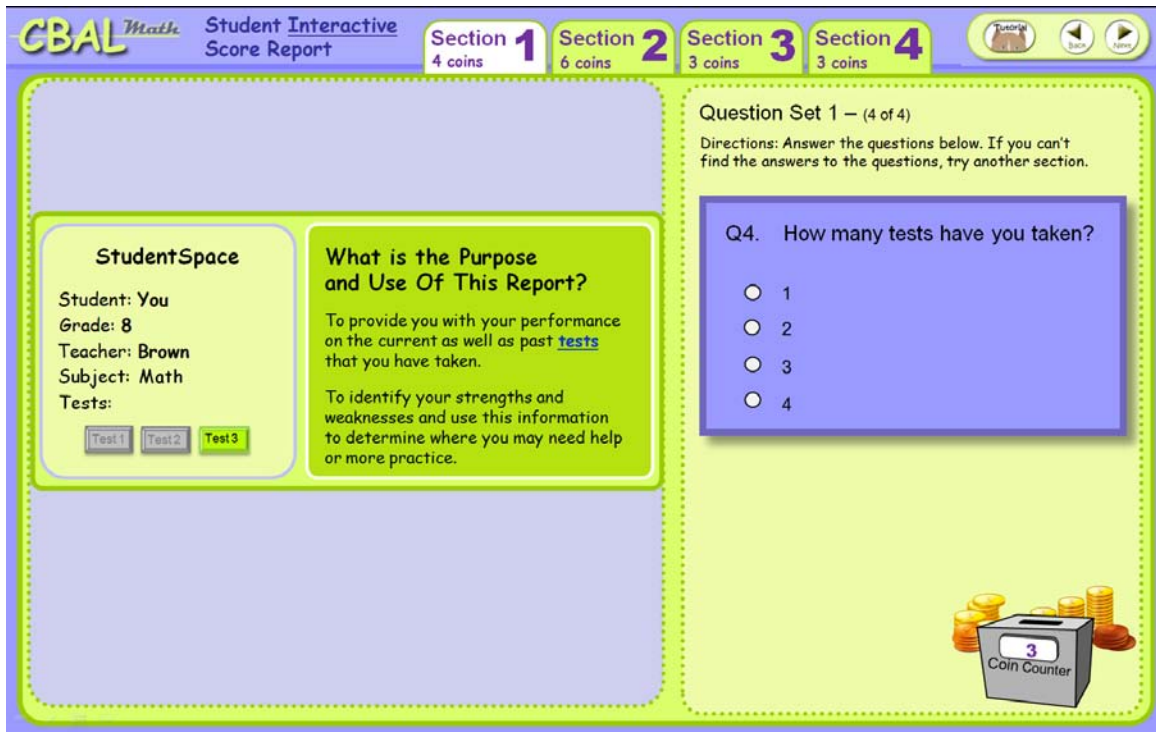
## **Version 2**

Several changes were made to the interactive student score report based on information from the initial usability study, internal feedback, and the aforementioned principles from Mayer (2001), Hattie (2009), and Goodman and Hambleton (2004). One of the major changes made to the score report involved the navigation. In Version 2, there are no doors for students to open, and the order of the score report is no longer sequential. Students navigate through the score report by selecting a section of the score report from a tabbed menu, which then displays that section. Each tab also indicates the number of coins that a student can collect in that particular section. Students can select any section and then answer the corresponding questions, which appear on the right-hand side of the screen.

Figures 7-10 show the new navigation scheme in Version 2 of the score report. Students move through the different sections of the score report, answering questions and collecting coins. The number of coins assigned to a question is dependent on its difficulty. Coins that students collect are put in a safe, which displays the number of coins that they have collected. Students who collect the most coins will earn a spot on a high score list. This coin collecting activity adds a fantasy element to the score report, making it more game-like and intrinsically motivating for the student.

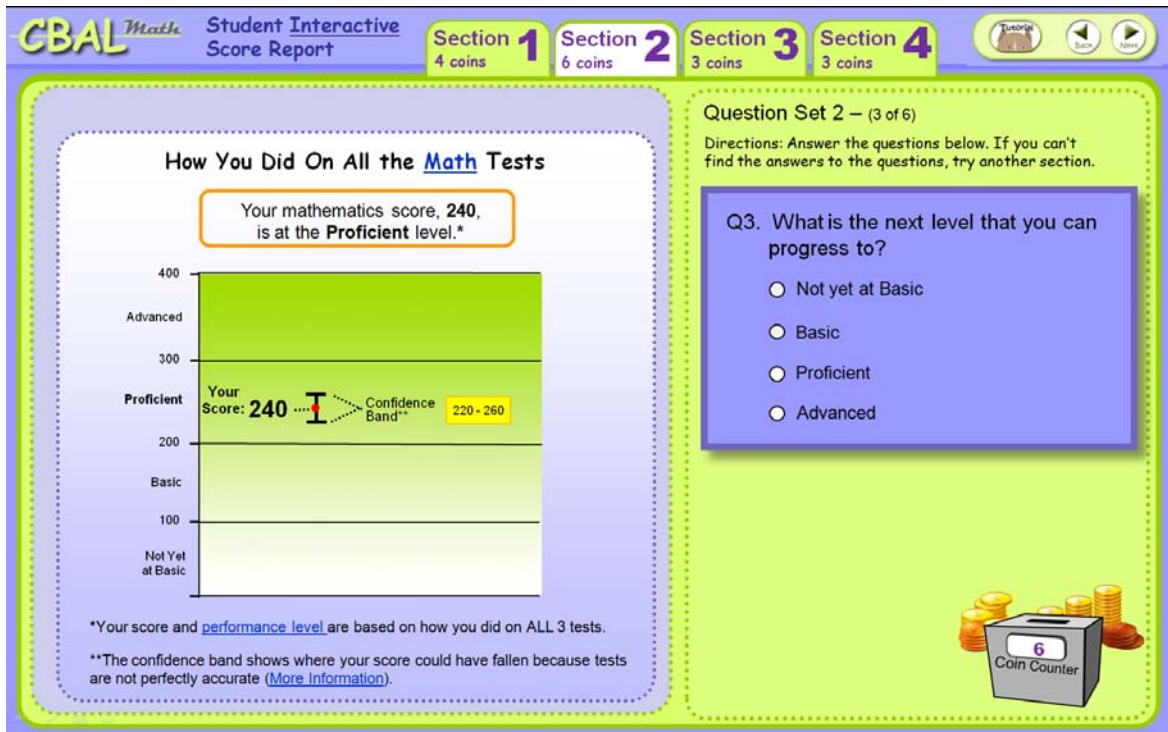
The student uses the tabbed menu to navigate through the report and also uses the navigation buttons at the top right side of the screen to go back to the previous question or advance to the next question within a question set. The student can also click on the tutorial button to review instructions on how to interact with the score report. The tab for the section that the student is currently viewing appears in white.

This new navigation scheme follows both Mayer's (2001) coherence principle and Hattie's (2009) recommendation to provide only necessary information. Breaking the score report up into sections and only showing the section that is relevant for answering questions within a particular set eliminates extraneous and irrelevant material. Additionally, by dividing the score report into four sections and only displaying one section at a time, we are following Mayer's spatial contiguity principle by ensuring that each representation in the score report is close to the questions related to it.



**Figure 7. Student space and purpose and use.**

Figure 8 shows the second section of the score report, which presents the student's test score and performance level. This section utilizes Mayer's (2001) principle of various representations by including both a written description and a graphical representation of the student's score, performance level, and confidence band. This representation also adheres to the spatial contiguity principle, as the written description is placed in close proximity to the score report graph. Following Hattie's (2009) recommendation, we have used color in a purposeful way by using a gradient in the score and performance level graph, in which the color gets darker as the performance level increases.



**Figure 8. How you did on all the math tests.**

Figure 9 shows Section 3 of the Version 2 score report. Keeping in mind Hattie’s (2009) point that too much information can overwhelm the user, we made several revisions to this section in order to decrease the amount of information presented on the screen. In this section, a table that showed task-level performance information for all the tasks has been modified to include tabs and buttons that present information for one task and one question type at a time. Highlights for overall performance on all of the tasks completed are provided at the bottom of the screen. The icons before each statement in this section follow Hattie’s recommendation regarding the purposeful use of color and help indicate whether each highlight is positive (green checkmark), indicates a warning (yellow triangle with exclamation point), or is a problem (red mark).



**CBAL Math Student Interactive Score Report**

Section 1 (4 coins) Section 2 (6 coins) Section 3 (3 coins) Section 4 (3 coins)

**What You Did on Test 3 (Proportional Reasoning & Statistics)**

Task 1 Task 2

This task includes questions about proportional reasoning as well as creating, interpreting and using different data displays.

Question Type	Description of Performance	Number Correct
1	Correctly used estimation to identify similar shapes based on a pictorial representation.	4 out of 4 questions

**Skills**

Content: Statistics: Use and Interpret Data Displays  
Process: Model & Represent

Sample Question: Click here for a Sample Question

Question Type 1 Question Type 2 Question Type 3 Question Type 4

Click on Question Type to get more detail

**Highlights of Your Overall Performance on Tasks 1 and 2**

- You correctly answered questions that require understanding and applying estimation correctly using pictorial representation.
- You correctly answered questions that require creating, interpreting and using different kinds of data displays, particularly scatterplots and trend lines.
- You seem to require additional work on providing justified statements to support conclusions based on data.

Question Set 3 – (3 of 3)

Directions: Answer the questions below. If you can't find the answers to the questions, try another section.

Q3. In which of the following areas do you require additional work?

- Interpreting a scatterplot
- Plotting points on a scatterplot
- Providing well-justified statements to support conclusions about data

12 Coin Counter

Figure 9. Task-level performance and highlights.

**CBAL Math Student Interactive Score Report**

Section 1 (4 coins) Section 2 (6 coins) Section 3 (3 coins) Section 4 (3 coins)

**What You Did:**

**Data Displays:**  
You correctly answered questions that require creating, interpreting and using different kinds of data displays.

**Algebraic Expressions:**  
You correctly answered questions that require understanding algebraic expressions.

**Well Justified Statements:**  
You provided some well-justified statements to support conclusions.

**What You Need To Do To Get To the Advanced Level:**

- Solve problems that require the use of functions and relations.
- Evaluate expressions and provide well-justified statements to support conclusions.

Question Set 4 – (3 of 3)

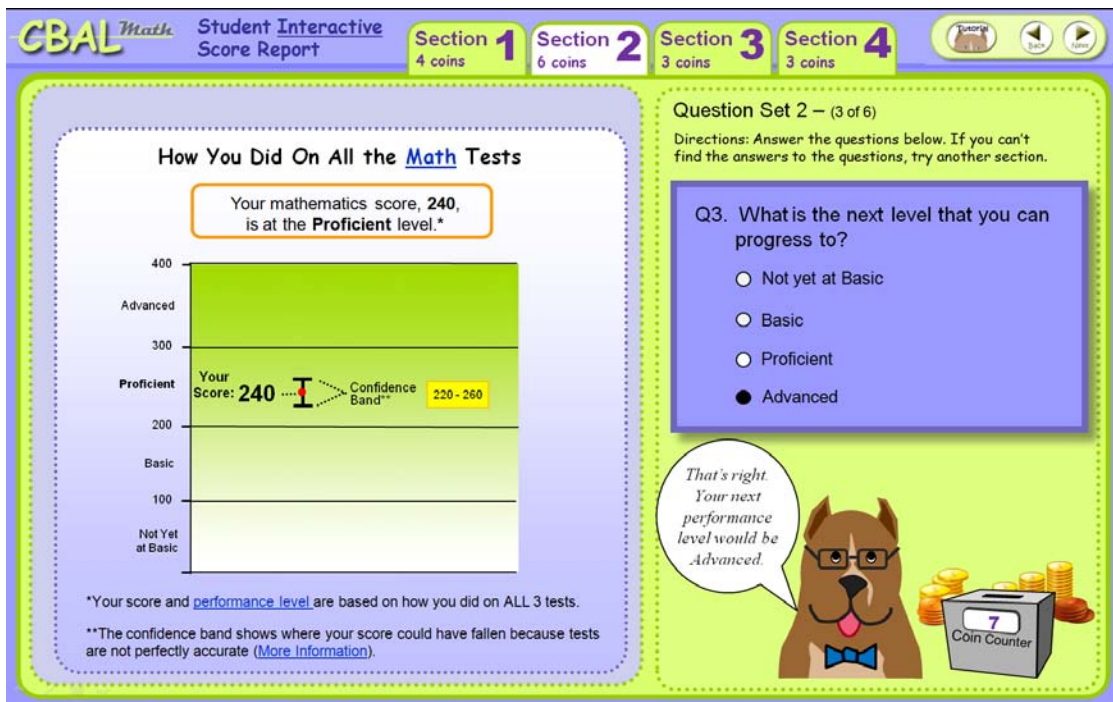
Directions: Answer the questions below. If you can't find the answers to the questions, try another section.

Q3. What would you do to improve your performance?

15 Coin Counter

Figure 10. What you did and what you need to do to get to the advanced level.

Version 2 includes several features that are used to provide the student with guidance and immediate feedback. We have followed Malone and Lepper’s (1987) recommendation of appealing to students’ emotional needs through fantasy by presenting an imaginary character as a guide for the students throughout the score report activity. A virtual dog that always appears in the bottom right area of the screen is used to provide directions and feedback to students about their understanding of the score report (see Figure 11). This feedback includes verification feedback (i.e., correct/incorrect) and hints tailored to help students find the correct answers. When all the options have been tried by the student, the correct answer is provided. The dog is also used to introduce students to the game-like activity and explain the components of the interface as part of an initial tutorial. In addition, the score report includes links to additional information (e.g., definitions, statistical terms, and sample test questions) that students can use to help them answer some of the questions.



**Figure 11.** Virtual dog used to provide immediate feedback, explain how to use the interface, and introduce the activity.

## Follow-Up Usability Study

### Method

A follow-up usability study was conducted with 11 local middle school students in February 2011 to compare the two versions of the score report and to identify major accessibility, readability, and navigation problems. Participants' background information is presented in Table 2. Students interacted with Version 2 of the score report and answered a series of comprehension questions embedded in the report. The Version 2 that students saw in this study did not have the coin collecting activity implemented yet, and the navigation differed slightly from the tabbed menu. In some cases (i.e., when students answered a question incorrectly), they were given the opportunity to answer the question again using Version 1 of the score report. Students were asked to think aloud while answering each question. After viewing Version 2 of the score report, students completed a usability survey (see Appendix C). This survey included five questions in which students were asked to choose among two alternative representations of some of the content of the report and explain their choice. Students also completed a short background questionnaire. This questionnaire was the same as that used in the initial study, except that a question about playing games on a computer and a question about using a game console were combined into one question about video games. Observers provided initial guidance and took notes during the study. Students received a gift certificate as payment for their participation.

**Table 2**

***Participants' Background***

ID	Gender	Race	Grade
001	Female	African-American/Caucasian	8
002	Male	African-American/Caucasian	8
003	Male	African-American	8
004	Male	Caucasian	8
005	Male	Asian	8
006	Female	Caucasian	7
007	Female	Caucasian	8
008	Male	African-American	6
009	Female	Asian	7
010	Female	African-American	6
011	Male	African-American	8




## Results

In general, the students had positive impressions of the score report. Ten out of 11 students agreed that the score report looked nice and said that they would like to use a score report like this one in the future. Most of the students (9 out of 11) liked answering the questions about the different areas of the score report, and 8 out of 11 liked being able to write about what to do to improve their performance. All of the students said that they liked knowing what they had to do to advance to the next performance level.

Overall, students found the score report useful and easy to understand. All but one student said that they felt that this activity helped them to learn about the different sections of a score report, and 8 out of 11 students felt that the embedded questions helped them to understand the score report. In addition, all of the students felt that the pop-up definitions helped them to better understand the score report. Nine out of 11 found the score and performance level graph easy to understand, and 10 out of 11 students said that the summary of what they did was easy to understand.

For the most part, students were able to correctly answer the embedded questions using information from the score report, though some students struggled more than others. In some cases, a few students could not answer a question correctly using Version 2 of the score report, but after being shown Version 1, they were able to find the correct answer. Although it only happened with a small number of students, we saw this occur more often for the questions in Section 1 that dealt with the purpose and use of the score report. Two specific things about the representation in Version 1 struck some students as being clearer than Version 2. First, Version 1 has two distinct sections separating purpose and use, indicated by the titles “What is the Purpose of This Report?” and “How Should You Use This Report?”, whereas Version 2 groups this information all under the heading “What is the Purpose and Use of This Report?”, as shown in Figure 12. Additionally, the representation which indicates the number of tests a student has taken was not as clear in Version 2 to some students as it was in Version 1. Version 2 indicates this through the use of three boxes that say Test 1, Test 2, and Test 3, with the current test box highlighted, while Version 1 does not include these boxes and simply says “Test: 3 of 3” (see Figure 13). This section has been revised by adding the label “Tests:” above the boxes, as shown in Figure 7.

<p><b>What is the Purpose of This Report?</b></p> <p>To provide you with your performance on the current as well as past tests that you have taken.</p> <p><b>How Should You Use This Report?</b></p> <p>To identify your strengths and weaknesses and use this information to determine where you may need help or more practice.</p>	<p><b>What is the Purpose and Use Of This Report?</b> </p> <p>To provide you with your performance on the current as well as past tests that you have taken.</p> <p>To identify your strengths and weaknesses and use this information to determine where you may need help or more practice.</p>
Version 1	Version 2

**Figure 12. Comparison of purpose and use sections in Version 1 and Version 2.**

<p><b>YourSpace</b></p> <p>Student: You</p> <p>Teacher: Brown Subject: Math</p> <p>Grade: 8 Test : 3 of 3 Test Date: 03/30/10</p>	<p><b>StudentSpace</b></p> <p>Student: You Teacher: Brown Grade: 8 Subject: Math</p> <p><input type="button" value="Test 1"/> <input type="button" value="Test 2"/> <input type="button" value="Test 3"/></p> <p>Test Date: 03/30/10</p>
Version 1	Version 2

**Figure 13. Representation indicating number of tests in Version 1 and Version 2.**

The buttons for each question type in Version 2’s table representation in Section 3, “What You Did on Test 3,” also proved problematic for some students. Four out of 11 students either did not notice these buttons or did not realize that they needed to click on them to view a different question type. One student suggested that the buttons would stand out more if they were placed vertically on the right-hand side of the question type information, rather than horizontally below the information.

Students struggled with some of the statistical information in the report. While 8 out of 11 students were able to identify what a confidence band was, using information from the report, five students indicated on the usability survey that they did not understand what a confidence

band is. All but one student said that they thought the link to more information about error provided useful information. However, many students had interesting interpretations about what error means. For example, one student used the entire confidence band calculation in the link to describe error, and another thought that error was calculated by multiplying the number of questions a student answered incorrectly by the number of points that each question is worth.

Students had mixed feelings about the dog character that was used to provide feedback in Version 2 of the score report. Seven out of 11 students said that they liked the way the dog looked. Two of the students who did not like the dog said that they felt it was more appropriate for younger students. Based on this finding, future versions of the score report might show students several possible characters and let them choose which one they would like to see in the report. Even though several students did not like the way the dog looked, all of them thought the feedback that it gave was helpful. In general, students liked the representations in Version 2 of the score report, though some of the new representations were better received than others.

In the case of section three: “What You Did on Test 3,” the results were mixed. Five students preferred the representation in Version 1, while five students preferred the representation in Version 2. One student liked both, stating that Version 1 was clearer, but had too much information, while Version 2 was more straightforward, but did not have enough information. Of the students who preferred Version 1, four cited quicker and less complicated navigation as their reason. Two of the students who preferred Version 2 noted that they found it easier to focus on one thing at a time.

In the case of section one: Student Space and Purpose and Use, 4 out of the 8 students preferred Version 2 because the space that displayed the information was larger and more spread out. (Two students did not answer this question, and one student’s answer to this question was not able to be interpreted.) Of the two students who did not express a clear preference for either version, both liked Version 2 for the same reasons noted above. The two students who preferred Version 1 noted that they preferred its use of titles and clear distinction between the statements describing purpose and the statements describing use.

In general, students who answered questions incorrectly did not benefit from viewing Version 1. Examination of the students’ improvement plans revealed that 7 out of 11 students identified and cited relevant information in the score report to create reasonable action plans, while the remaining four students provided plans that were general but reasonable.

### **Expert Review**

A group of experts reviewed the new type of score report and provided feedback and recommendations for future work. The following experts participated in this activity: Ronald K. Hambleton (University of Massachusetts Amherst), Priti Shah (University of Michigan), Howard Wainer (National Board of Medical Examiners), and Rebecca Zwick (ETS). The expert review took the form of a structured-focus group in which the experts viewed a demonstration of the student score report and were given opportunities to provide feedback and to engage in further discussion. The experts appreciated the effort to design score reports for students. Experts' suggestions included simplifying the amount of information presented to students and carrying out small studies to evaluate the graphical representations, definitions, and feedback available to students in the score report.

### **Summary and Future Work**

This paper describes an interactive student score report that has been designed taking into account findings from relevant research areas and our methodology for designing and evaluating score reports. This interactive score report was evaluated and refined based upon results from usability studies with students and suggestions made by experts.

Results from a usability study aimed at comparing two versions of the interactive score report indicated that, overall, students liked the representations in Version 2. Additionally, we learned that, while our new character for giving feedback was perceived as friendlier than our original characters, it needs to be further refined to better appeal to our target age group of middle school students. Future work includes making further revisions to the interactive score report and conducting additional studies to determine its effectiveness to facilitate students' understanding of score reports and increase student motivation.

## References

- Avitabile, J. (1998). Interaction of presentation mode and learning style in computer science. In *NECC '98: Proceedings of the National Educating Computing Conference*. San Diego, CA.
- Bennett, R. E., & Gitomer, D. H. (2009). Transforming K–12 assessment: Integrating accountability testing, formative assessment and professional support. In C. Wyatt-Smith & J. Cumming (Eds.), *Assessment issues of the 21st century* (pp. 43-62). New York, NY: Springer.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-74.
- Bottino, R., Ferlino, L., Ott, M., & Tavella, M. (2007). Developing strategic and reasoning abilities with computer games at primary school level. *Computers & Education*, 49, 1272-1286.
- Brna, P., Self, J., Bull, S., & Pain, H. (1999). Negotiated collaborative assessment through collaborative student modelling. In *Proceedings of the Workshop: Open, Interactive, and other Overt Approaches to Learner Modelling at AIED 1999* (pp. 35–44). Le Mans, France: International AIED Society.
- Bull, S. & Pain, H. (1995). Did I say what I think I said, and do you agree with me? Inspecting and questioning the student model. In *Proceedings of World Conference on Artificial Intelligence and Education* (pp. 501-508). Charlottesville, VA: AACE.
- Cordova, D., & Lepper, M. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, 88, 715-730.
- Deng, N., & Yoo, H. (2009). *Resources for reporting test scores: A bibliography for the assessment community*. Retrieved from the National Council on Measurement in Education website: [https://ncme.org/resources/bliblio1/NCME.Bibliography-5-6-09\\_score\\_reporting.pdf](https://ncme.org/resources/bliblio1/NCME.Bibliography-5-6-09_score_reporting.pdf)
- Goodman, D. P., & Hambleton, R. K. (2004). Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education*, 17, 145-220.

- Hattie, J. (2009). Visibly learning from reports: The validity of score reports. *Online Educational Research Journal*. Retrieved from: <http://www.oerj.org/View?action=viewPDF&paper=6>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81-112.
- Kennedy, D. M., & McNaught, C. (1997). Design elements for interactive multimedia. *Australian Journal of Educational Technology*, 13, 1-22.
- Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction: III. Conative and affective process analyses* (pp. 223-253). Hillsdale, NJ: Erlbaum.
- Mandinach, E. B., & Corno, L. (1985). Cognitive engagement variations among students of different ability level and sex in a computer problem solving game. *Sex Roles*, 13, 241-251.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge, England: Cambridge University Press.
- McCormack, A. (2002). Interactive computer-based technology in EFL. In C. Crawford, D. Willis, R. Carlsen, I. Gibson, K. McFerrin, J. Price, & R. Weber (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2002* (pp. 1151-1152). Chesapeake, VA: AACE.
- McDonald, K., & Hannafin, R. (2003). Using Web-based computer games to meet the demands of today's high-stakes testing: A mixed method inquiry. *Journal of Research on Technology in Education*, 35, 459-472.
- Neo, M., & Neo, T.-K. (2009). Engaging students in multimedia-mediated Constructivist learning – Students' perceptions. *Educational Technology & Society*, 12(2), 254–266.
- Reeves, T. (1998). *The impact of media and technology in schools: A research report*. Retrieved from <http://it.coe.uga.edu/~treeves/edit6900/BertelsmannReeves98.pdf>
- Regan, M., & Sheppard, S. (1996). Interactive multimedia courseware and the hands-on learning experience: An assessment study. *Journal of Engineering Education*, 85, 123-130.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153-189.
- Underwood J. S., Reshetar, R., & Leahy, S. (2006). *Score report design heuristics*. Unpublished manuscript.

- Zapata-Rivera, D. (2009). *Experimentally evaluating alternative report designs for CBAL*. Unpublished manuscript.
- Zapata-Rivera, D., Underwood, J. S., & Bauer, M. (2005). Advanced reporting systems in assessment environments. In *AIED'05 Workshop on Learner Modelling for Reflection, to Support Learner Control, Metacognition and Improved Communication between Teachers and Learners* (pp. 23-32). Amsterdam, Netherlands: International AIED Society.
- Zapata-Rivera D., & VanWinkle, W. (2010). *A research-based approach to designing and evaluating score reports for teachers* (Research Memorandum No. RM-10-01). Princeton, NJ: ETS.
- Zapata-Rivera, J. D., & Greer, J. (2003). Analysing student reflection in The Learning Game. In *Supplementary Proceedings of the International Conference on Artificial Intelligence in Education: Vol. 5. Workshop on Learner Modelling for Reflection* (pp. 288-298). Sydney, Australia: International AIED Society.

## **List of Appendices**

	Page
A. Student Background Questionnaire for Initial Usability Study .....	27
B. Usability Survey for Initial Usability Study.....	28
C. Usability Survey for Follow-Up Usability Study .....	31



## Appendix A

### Student Background Questionnaire for Initial Usability Study

How comfortable are you with **computers**?

- Not at all Comfortable
- Somewhat Comfortable
- Comfortable
- Very Comfortable

How often do you use the **Internet**?

- Not at all
- Few times a month
- Once a week
- Daily

How often do you **play games on a computer**?

- Not at all
- Few times a month
- Once a week
- Daily

How often do you play **video games on a game console**?

- Not at all
- Few times a month
- Once a week
- Daily

Have you ever seen a **student score report** before?

- Yes
- No

If so, what do you usually do with it?

## Appendix B

### Usability Survey for Initial Usability Study



*Student Report*

*We would like to know about your experience with CBAL Student Report.*

*Total confidentiality for students, teachers, and schools will be ensured.*

*All responses are anonymous and will only be used to improve the program.*

**The following questions refer to the student score report you just saw. Please, read each question carefully and mark the answer that you think is best.**

<b>How much do you agree with each statement?</b>	<u>Strongly</u> <u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Strongly</u> <u>Disagree</u>
1. I feel that I learned about the different pieces of a score report by using this software.	①	②	③	④
2. I liked uncovering the different areas of the interactive report by answering questions.	①	②	③	④
3. The directions on the screen were easy to understand.	①	②	③	④
4. The vocabulary was easy to understand.	①	②	③	④
5. I would like to use an interactive report like this one in the future.	①	②	③	④
	<u>Strongly</u>			<u>Strongly</u>
<b>How much do you agree with each statement?</b>	<u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Disagree</u>
6. The interactive report looked nice.	①	②	③	④
7. The definition pop-ups (links) helped me understand the interactive report.	①	②	③	④
8. The pop-ups that described the contents of each door helped me understand what was coming next.	①	②	③	④
9. The sample questions helped me understand the types of questions on the test.	①	②	③	④
10. I found the content and process skill pop-ups useful.	①	②	③	④

<b>How much do you agree with each statement?</b>	<u>Strongly</u> <u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Strongly</u> <u>Disagree</u>
11. The questions on the screen were easy to answer.	①	②	③	④
12. The questions helped me understand the interactive report.	①	②	③	④
13. I liked being able to write in my response to some of the questions.	①	②	③	④
14. I liked being able to write in a plan for improving my performance.	①	②	③	④
15. I liked being able to share my improvement plan with the teacher.	①	②	③	④
16. I liked being able to share my improvement plan with my parents or guardian.	①	②	③	④

<b>How much do you agree with each statement?</b>	<u>Strongly</u> <u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Strongly</u> <u>Disagree</u>
17. The graph with the score in it was easy to understand.	①	②	③	④
18. The <i>more information</i> pop-up provided useful information.	①	②	③	④
19. I understand the purpose of this score report.	①	②	③	④
20. I understand what a PAA is.	①	②	③	④
21. I understand what a content skill is.	①	②	③	④
22. I understand what a process skill is	①	②	③	④

<b>How much do you agree with each statement?</b>	<u>Strongly</u> <u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Strongly</u> <u>Disagree</u>
23. The table that describes the types of questions used in PAA3 was easy to understand.	①	②	③	④
24. The PAA3 summary below the table was easy to understand.	①	②	③	④
25. The “How You Did on All the PAAs” section was easy to understand.	①	②	③	④
26. The “To Progress to the Advanced Level a Student Needs to” section was useful to help me understand where to improve.	①	②	③	④
27. The characters provided useful information.	①	②	③	④
28. I liked how the characters looked.	①	②	③	④

29. What does the term *standard error of measurement (SEM)* mean?

30. What does the term *confidence band* mean?

31. Please provide any additional comments or suggestions you may have in the space below:

## Appendix C

### Usability Survey for Follow-Up Usability Study



*We would like to know about your experience with CBAL Student Report.*

*Total confidentiality for students, teachers, and schools will be ensured.*

*All responses are anonymous and will only be used to improve the program.*

**The following questions refer to the student score report you just saw. Please, read each question carefully and mark the answer that you think is best.**

<b>How much do you agree with each statement?</b>	<u>Strongly</u> <u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Strongly</u> <u>Disagree</u>
1. I feel that I learned about the different pieces of a score report by using this tool.	①	②	③	④
2. I liked answering questions about different areas of the score report.	①	②	③	④
3. The directions on the screen were easy to understand.	①	②	③	④
4. The vocabulary was easy to understand.	①	②	③	④
5. I would like to use a score report like this in the future.	①	②	③	④
<b>How much do you agree with each statement?</b>	<u>Strongly</u> <u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Strongly</u> <u>Disagree</u>
6. The score report looked nice.	①	②	③	④
7. The pop-ups of definitions (e.g., Tests, Performance Levels etc.) helped me better understand the score report.	①	②	③	④
8. The pop-up (in the green box) at the beginning helped me understand what I would be doing.	①	②	③	④
9. The map was easy to use.	①	②	③	④

How much do you agree with each statement?	<u>Strongly</u>			<u>Strongly</u>
	<u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Disagree</u>
10. The questions helped me understand the score report.	①	②	③	④
11. The questions on the screen were easy to answer.	①	②	③	④
12. I liked being able to write about what to do to improve my performance.	①	②	③	④
13. I liked knowing what I needed to do in order to progress to the advanced level.	①	②	③	④
14. The summary about what I did was easy to understand.	①	②	③	④

How much do you agree with each statement?	<u>Strongly</u>			<u>Strongly</u>
	<u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Disagree</u>
15. The graph with the score in it was easy to understand.	①	②	③	④
16. The <i>more information</i> pop-up provided useful information.	①	②	③	④
17. I understand the purpose of this score report.	①	②	③	④
18. I understand what a confidence band is.	①	②	③	④

How much do you agree with each statement?	<u>Strongly</u>			<u>Strongly</u>
	<u>Agree</u>	<u>Agree</u>	<u>Disagree</u>	<u>Disagree</u>
19. I liked how the dog looked.	①	②	③	④
20. The feedback provided by the dog was helpful.	①	②	③	④

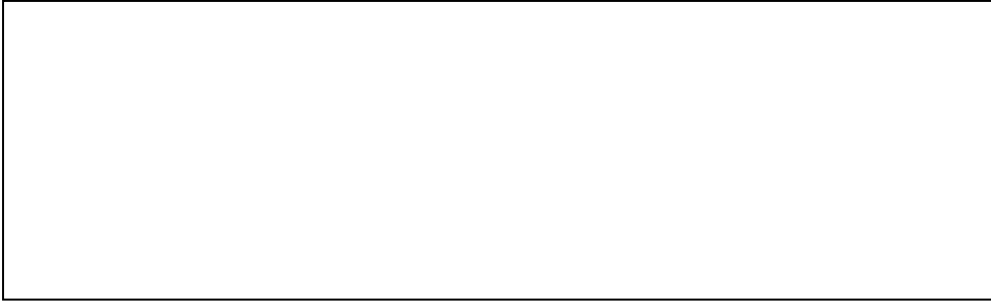
**Please answer the following questions about which version of the score report you prefer.  
Ask the researcher to show you the part of the score report mentioned in the questions.**

21. Which character do you prefer and why?

22. Which navigation do you prefer (the doors or the map) and why?

23. Which table representation for “What You Did on Test 3” do you prefer and why? Were the parts of this section (Description of Performance, Highlights of Your Overall Performance) easy to understand? Were the icons easy to understand?

24. Which version of the section describing what you did on the test and what you need to do to get to the advanced level do you prefer and why?



25. Which version of the “Purpose and Use” section do you prefer and why?

