What's Inside the Box? An Open Student Modeling Approach in a Museum Context

Diego Zapata-Rivera
Educational Testing Service
660 Rosedale Road MS 16-R
Princeton, NJ 08541
Dzapata@ets.org

ABSTRACT

Adaptive learning environments and technology-rich assessments capture evidence of students' skills, knowledge, and other attributes and use it to adapt their interaction or support assessment claims. Data captured to support assessment claims or implement adaptive behavior can include responses to predefined questions and process data. However, students are not always aware of the type of data being captured and how these data are used by these systems. An open student modeling system, implemented as a museum exhibit called "What's Inside the Box" has been designed to provide students with information about how a technology-rich assessment system makes use of both response and process data to support assessment claims. In this paper we describe the "What's Inside the Box" system and report on the results of a small-scale study aimed at evaluating system usability and perceived value issues.

CCS Concepts

CCS → Human-centered computing → Human computer interaction (HCI) → HCI design and evaluation methods → User models.

Keywords

Open student models; technology-rich, adaptive assessment systems; museum exhibits.

1. INTRODUCTION

Open student models have been used for various purposes including: reducing the complexity of the diagnosis process, supporting student reflection, knowledge awareness, and student learning [1].

Open student models can be used to share response and process data with students, teachers, or parents/guardians in informal learning environments such as museums. In fact, student/user models have been used to generate personalized museum tours, predict users' location, and provide additional information based on user's interests, background information, and path history [2-7]. Stock et al. [8] describe a framework for implementing user modeling applications in museums. This framework includes animated agents that motivate visitors and provide recommendations, adaptive video documentaries and visit summaries. Visitors' user models can be made available to users to make adjustments which may result on better recommendations and an enhanced visitors' experience [6]. Cramer et al. [7] showed that user/student model transparency increases user understanding and acceptance of the system's recommendations.

The open student modeling approach presented in this paper has been implemented as a museum exhibit ("What's Inside the Box?)." Data for the student model were collected using a technology-rich assessment system (the Technology-Rich Environment; TRE) [9]. The "What's inside the Box" system, was designed to show students how their response and process data are used by the system to make assessment claims.

2. OPEN STUDENT MODELING IN MUSEUMS

Informal education contexts such as museums impose particular measurement challenges that can hinder the creation, maintenance and interaction with student/user models. Some of these measurement challenges include [10]: (a) a high degree of freedom and flexibility, which makes it difficult to isolate, keep track and measure individual learning, (b) interactions can vary in duration, type of activity, number of people involved; and (c) interactions may include emerging behavior and unpredictable interactions with other visitors and facilitators.

Several strategies can be used to deal with some of these issues. For example, it is possible to initialize the student model with data from other visitors that share some of the characteristics of the intended audience or borrowing information from existing student models that were created in other contexts [6]. Once a student/user model is available, this model can be used to integrate additional evidence of students' skills, knowledge and other attributes based on their interactions with the exhibits at the museum using a variety of sensors and tracking mechanisms [2-8].

To the extent to which a student/user model is available, different types of recommendations could be implemented. Also, exhibits can use information on the student/user model to adapt their interaction to the particular individual. This can result in an improved user experience. Explaining to individuals why particular recommendations are offered or how exhibits adapt their interaction becomes an interesting challenge, since the adaptation can involve data gathered before or during the visit to the museum.

By keeping track of individual interactions in the museum, it is possible to gather data about how successful particular exhibits are at adapting their interaction and keeping individuals engaged. Also, information in the student/user model can potentially be used to assess student learning at the museum [10]. Students can use this information to plan future visits or follow-up on particular topics. Teachers could receive a report or explore on how their students interacted with the exhibits and use this information to plan instructional and debriefing activities in the classroom.

An interactive museum exhibit featuring an open student model approach serves as a testbed for exploring some of the challenges of implementing student/user models in informal environments. The "What's Inside the Box?" was designed to show students how their response and process data are used by the system to make assessment claims (levels on student model variables and supporting evidence).

3. WHAT'S INSIDE THE BOX?

Recommendations for designing museum exhibits include [11]: (a) design them with a specific learning goal in mind; (b) make them interactive; (c) provide multiple ways of experience concepts, practices, and phenomena; (d) provide support for participants to interpret their learning experience; (e) build on participants' prior learning and interests; and (e) encourage participants to extend their learning outside the museum experience. Following these recommendations, we have implemented "What's Inside the Box?" system.

The contents and data used in the "What's Inside the Box?" system are based on one of three simulation problems used in the TRE project that is intended to elicit evidence for two scientific inquiry skills: Scientific Exploration and Scientific Synthesis [9]. In the TRE assessment system, students are scored based on both responses to particular questions and the process (actions) used to arrive at the answers.

The "What's Inside the Box?" system is intended to be used as a standalone museum exhibit offering the public a view of how student problem-solving in science can be measured using computer-based simulation tasks. Students are asked to solve a scientific problem. Students can witness how the computer (the "Box") takes into account both their interactions with the simulation and their responses to particular questions to make evidence-based claims about what they know and are able to do. Students take from 5 to 10 minutes to complete the activity.

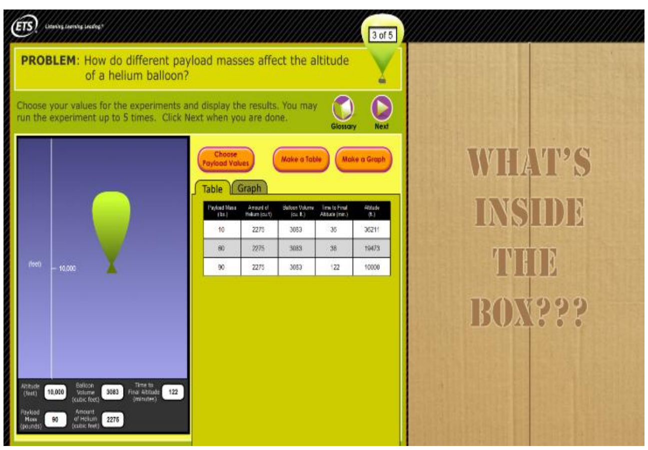

Figure 1. Balloon animation after experiment 3.

Figure 1 shows a screenshot of the system. The screen is divided into two basic areas; the experiment area on the left side of the screen that is used to select payload mass values, set parameters for a table and a graph that will show the data collected, and run the simulation, and the "What's Inside the Box" area on the right that dynamically updates its content based on how students interact with the system. The "Box" starts closed but students can click on it to see its contents at any time. The contents of this area are shown to students once they complete the experiments and after answering data interpretation questions. A glossary of relevant terms is available for students to inspect at any time. Several hidden, sound effects and funny remarks were added to encourage students to explore different areas of the screen.

At the beginning, students view a short introduction describing the different parts of the system. Students are told that they can click on the "What's Inside of the Box?" area at any time to see how the system measures their problem solving skills as they solve the problem. To solve the simulation problem ("How do different payload masses affect the altitude of a helium balloon?"), students can try up to 5 experiments. Students can make use of a table and a graph to record their data. On each experiment, students may choose a payload mass value (one of nine possible payload mass values that go from 10 lbs to 90 lbs in increments of 10 lbs), and select variables to include on the table as columns and or as axes of the graph. After the students selects a payload mass value and clicks on "try it," they see an animation of the balloon moving upwards while values for the variables at the bottom of the balloon area are calculated. The contents of both the table and the graph are updated after each experiment. If, after having run two experiments, the student has not selected variables for the table or the graph, a hint is presented ("Here's a hint. You may want to make a table and a graph").

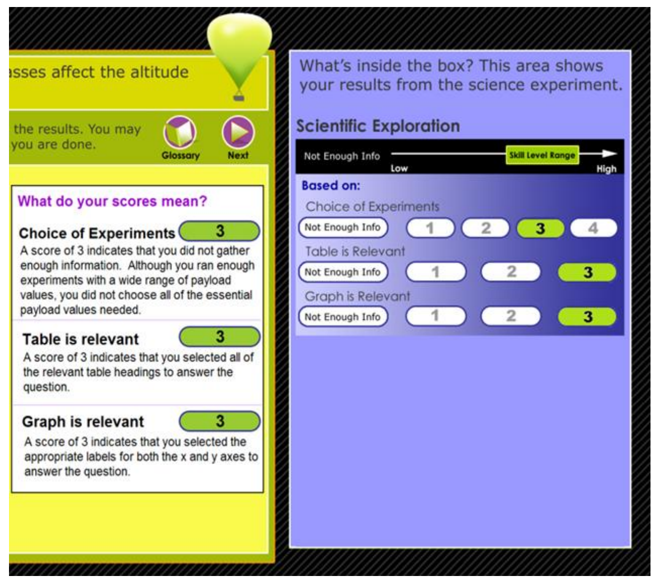

Figure 2. Student model information after running the experiments.

Figures 2 and 3 show student model information in the "What's Inside the Box" area. This information includes skill level ranges for two student model variables: Scientific Exploration and Scientific Synthesis and evidence used to support the skill level ranges. The evidence is represented by levels for relevant observables (i.e., student actions) that are linked to particular student model variables and their corresponding explanations.

Figure 2 shows the status of the student model after the student has run the experiments. The student model depicted in Figure 2 belongs to a student who achieved level 3 on "Choice of Experiments," level 3 on "Table is Relevant," and level 3 on "Graph is Relevant." The explanations provided are intended to inform the student about the actions/selections that were used by the system to the levels.

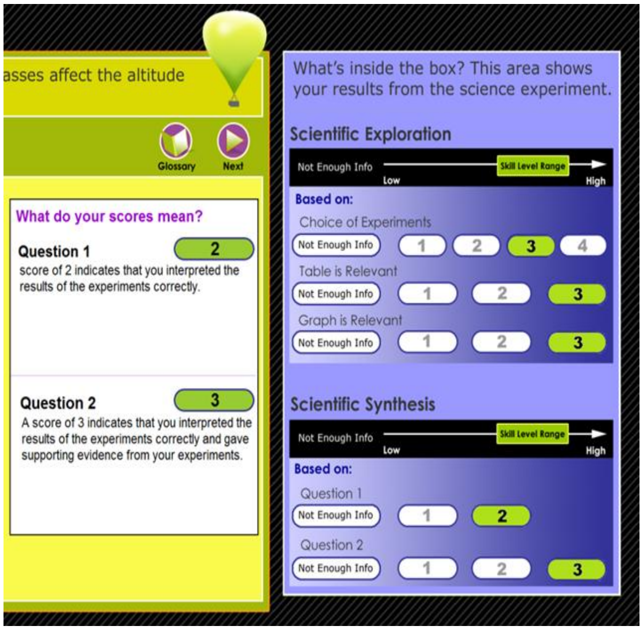

Figure 3. Student model information after answering Scientific Synthesis questions.

Figure 3 shows the status of the student model after having answered Scientific Synthesis questions (i.e., multiple-choice questions interpreting the results of the experiments). The explanations are based on the choices selected by the students.

As mentioned earlier, the system automatically shows the student model at these two particular moments. However, students are free to open the "Box" at any time to see the status of the student model. By showing the contents of the student model (skill level ranges, observable levels and explanations) at these two particular moments, the system provides students with the opportunity to reflect on how their recent actions are used by the system to assess their performance so far.

Relevant observables, their levels and explanations were determined through a study with nationally representative sample involving 2,134 8th grade students [9]. In this study, several process data features were extracted and evaluated for their correctness using scoring criteria called "evaluation rules." Summary scores were created using Bayesian networks.

Based on the student's actions, skill level information and corresponding explanations are determined and presented to the student in the "Box" area. It is worth noticing that each skill and observable has a "Not Enough Info" option. This option was included since it is possible that at some point during the interaction with the system, there may not be enough student data for the system to determine an observable level or skill level range.

Table 1 shows the levels for the observable "Choice of Experiments" which is the first observable depicted in Figure 2. The observable levels and explanations in Table 1 show increasing levels of proficiency at choosing parameters to run experiments. Basically, students should run enough experiments with a wide range of payload values to be able to gather enough data to solve the helium balloon problem.

This is the information used to explain the meaning of scores to the students (see Figures 2 and 3). Students can use this information to improve on their performance during the current interaction with the system or the next time they visit the exhibit.

Table 1: Choice of Experiments

Value	Explanation
Not Enough Info.	A score cannot be determined because you did not run any experiments.
1	A score of 1 indicates that you did not gather enough information. You either did not run enough experiments, or you chose payload values that were too close together.
2	A score of 2 indicates that you did not gather enough information. You either did not run enough experiments, or you did not choose a wide enough range of payload values.
3	A score of 3 indicates that you did not gather enough information. Although you ran enough experiments with a wide range of payload values, you did not choose *all* of the essential payload values needed.
4	A score of 4 indicates that you were successful in gathering enough information to answer the question.

4. USABILITY STUDY

A usability study was carried out to identify major accessibility, readability, and navigation problems as well as to gather feedback on the perceived value of this type of tool.

4.1 Participants

Participants were 11, 6th-10th grade students (7 female and 4 male). Participants received a $15 gift card for their participation in the study. All participants were familiar with museum exhibits and had taken computer-based tests in the past.

4.2 Procedure

Students completed a brief background questionnaire about their experience with museum exhibits and use of computers for learning and testing. Participants interacted with the system on a 40-inch, touch-screen monitor. Participants were asked to "think aloud" while interacting with the system.

The interaction with the system involved the following activities: going through the initial short introduction describing the different parts of the system, working on the helium balloon problem (by choosing payload values, running simulation, selecting variables for the table and graph), exploring the student model on-demand (by clicking on the "Box") or when the system made it available (after the experiments phase and after answering multiple choice questions about the experiments), interacting with the glossary (if needed), listening to sound effects and funny remarks when clicking on some areas of the interface, receiving hints and responding to open questions about the experiments).

A facilitator stayed with the student, took notes, and answered clarifying questions. One additional observer took notes through a two-way mirror. At the end on the interaction with the system, students were given the option to try again.

Finally, students completed a usability survey about their experiences with the system. After the students completed the survey, observers had the opportunity to ask students clarification questions.

5. RESULTS

Participants generally enjoyed the activity and found the system informative and easy to use. Nine or more students agreed or strongly agreed with the following statements: "I liked creating and running the experiments with the balloon," "The demo was entertaining," "The introduction at the beginning helped me understand what I would be doing," "The directions on the screen were easy to understand," "The vocabulary was easy to understand," "The touch-screen was easy to use."

Some issues students thought could be improved or were not useful include the sound effects and voices: "I liked the sounds and voices in the demo" (5 disagreed or strongly disagreed), and the glossary, "The glossary of definitions helped me better understand the demo" (6 disagreed or strongly disagreed").

Students seemed to understand and found the student model information useful: "By using this demonstration, I learned about how a computer-based test measures a student's skills" (9 agreed or strongly agreed), "I understand why I received the scores I did" (10 agreed or strongly agreed), and "I understand how the computer calculated my scientific skill range levels" (9 agreed or strongly agreed).

Students provided some suggestions for improving the system including: reducing the length of the introduction, adding hints to encourage students to open the "Box," adding words to the glossary and making hidden sounds easy to find, and placing them in areas relevant to the task.

Additional observations include: Only one student opened the "Box" before completing the experiments. This student used the information in the student model to help him choose the variables for the table and graph. All the students left it open after the experiments. Although students were informed about the availability of tools such as a table and a graph, they were allowed to proceed with the experiments without using them. Results showed that all students created a table. One student created it before the first experiment, 4 after the first experiment before the hint, and 6 after receiving the hint. All students made a graph. However, most of them made it after they heard the hint.

When asked "Do you understand the relationship between what you did in the experiment and how that was reflected in your skill ranges on the right side of the screen?" most of the students responded affirmatively. Some of the explanations provided include: "I understand that every answer I got wrong or right was recorded and deciphered and matched to form my skill range," "Yes, the results were explained clearly, although I'm not sure that a younger child might know that the independent variable goes on the x-axis and dependent on the y-axis," "Yes, I understand how my scores and skill levels were determined based on the choices of my experiments and my answers to the final questions," "It has a scale from high to low and it shows how you do and how you could have done, and you see explanations after you see score ranges," and "I was measured based upon relevancy of the topic of the tables and graphs I made and also on the accuracy of my answers when questions were asked."

Three of the students spontaneously decided to try again citing the following reasons: "I hope to improve my score," "I would try again because I didn't get the highest score," and "this is kind of fun. I would continue until I get a perfect score."

Some students' reactions to the information found in the "What's Inside the Box?" area include: "OK, that makes me feel better" (this student received perfect scores), "lots of room for improvement" (this student tried twice), "4 is good for the experimental choices. [Table] Oh, I didn't choose 2 headings. [Graph] only one of the two" (this student used this feedback to do a better job during the second round).

Finally, although the student model was produced based on data collected from 8th graders using the TRE system, we decided to open the usability study to 6th-10th grade students since the content of simulation problem implemented in the exhibit was accessible to all of them. We did not observe any major differences in the way these students used the system.

6. DISCUSSION AND FUTURE WORK

The results of the study provide initial evidence on how students interact with an open student model museum exhibit based on a technology-rich assessment system. Students seemed to understand and value the features included in the "What's Inside the Box?" system. We believe that open student models and embedded assessments have potential to support student learning and reflection in informal educational contexts [10].

Explanations provided by students about how their actions were used to update the skill level ranges and observable levels indicate that the information in the student model was understood as intended. We argue that these students may be better prepared to interact with adaptive learning and assessment systems that make use of response and process data.

Results showed the desire of some students to try again and use the student model information to improve their scores. This is interesting since museum exhibits compete with other exhibits for visitors' attention. Also, students may have more opportunities to practice their science inquiry skills and appreciate the open student model.

Even though only one student opened the "Box" before the system made it available after completing the experiments, all of them left it open after that. This seems to indicate that once students are aware of the availability of this information, they want to keep it on the screen.

Suggestions provided by students indicate that they would welcome more hints encouraging the use of the student model. These hints could be implemented as system alerts or implemented as an artificial agent that will accompany the student during the interaction with the system and alert the student about important changes to the student model.

The approach described here demonstrates how open student models could be used to create museum exhibits to help students become aware of how some technology-rich assessment systems use their response and process data to support assessment claims.

By gathering student model information before students interact with the exhibit (in this case by building upon the results of the TRE study), it is possible to create user/student models that can be used to create adaptive museum exhibits. Open student modeling applications in this contexts can provide information about why particular recommendations are made, what data are used to support these recommendations, and how museum exhibits adapt their interaction based on the information in the user/student model. By keeping this information in a generalized/life-long student model [12], the benefits of the museum visit can be transferred to other contexts such as the classroom and vice versa.

Teachers can also benefit from understanding how these types of systems use response and process data. This information can be useful in understanding how their students solve problems and the types of responses they provide to particular questions. Making teachers and students participants can also contribute to the acceptance and adoption of adaptive, technology-rich learning and assessment systems in the classroom.

This work may inform related research in areas such as exploring approaches for helping people understand computer science ideas in museum contexts and designing and evaluating interactive reporting tools and other materials to explain assessment concepts to various audiences.

Additional work in this area involves: exploring the types of graphical representations/guidance mechanisms that should be used to externalize the student model information in museums and other contexts (e.g., after school programs and school events such as school assemblies and science fairs); investigating whether the benefits of open student models in general, and feedback provided by this museum exhibit in particular, transfer to other technology-rich assessment adaptive systems; and exploring how other audiences (e.g., teachers and parents) interact with this types of systems.

7. ACKNOWLEDGMENTS

We thank Margaret Vezzu, Margaret Redman, Debbie Pisacreta and Tom Florek for their work in this project, and Randy Bennett, Eric Hansen, Irv R. Katz, Lei Liu, Hilary Persky, James Carlson, Lucia Sanin and two anonymous reviewers for their comments and suggestions.

8. REFERENCES

[1] Bull, S., and Kay, J. 2007. Student models that invite the learner in: the SMILI open learner modelling framework. *International Journal of Artificial Intelligence in Education*, 17(2), 89–120.

[2] Bohnert, F., Zukerman, I., Berkovsky, S., Baldwin, T., and Sonenberg, L. 2008. Using interest and transition models to predict visitor locations in museums. *AI Communications*, 21(2-3), 195-202.

[3] Bright, A., Kay, J., Ler, D., Ngo, K., Niu, W., and Nuguid, A. 2005. Adaptively recommending museum tours. In *Proceedings of the Ubicomp Workshop Smart Environments and Their Applications to Cultural Heritage*. http://smart.arces.unibo.it/pdf/04-Adaptively-Recommending_Bright.pdf.

[4] Wang, Y., Stash, N., Sambeek, R., Schuurmans, Y., Aroyo, L., Schreiber, G., and Gorgels, P. 2009. Cultivating personalized museum tours online and on-site. *Interdisciplinary Science Reviews*, 34(2-3), 139-15.

[5] Lane, H.C., Noren, D., Auerbach, D., Birch, M., and Swartout, W. 2011. Intelligent Tutoring Goes to the Museum in the Big City: A Pedagogical Agent for Informal Science Education. In G. Biswas, S. Bull, J. Kay & A. Mitrovic (Eds.), *Artificial Intelligence in Education: 15th International Conference* (Vol. 6738, pp. 155-162): Springer Berlin / Heidelberg.

[6] Kay, J., Lum, A., and Niu, W. 2005. A scrutable museum tour guide system. In *Proceedings of the 2nd Workshop on Multi-User and Ubiquitous User Interfaces* (pp. 19-20).

[7] Cramer, H., Evers, V., Ramlal, S., van Someren, M., Rutledge, L., Stash, N., Aroyo, L., and Wielinga, B. 2008. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Model. User-Adapt. Interact*. 18(5):455-496

[8] Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Krüger, A., Kruppa, M., Kuflik, T., Not, E., and Rocchi, C. 2007. Adaptive, intelligent presentation of information for the museum visitor in PEACH. *User-Model. User-Adapt. Interact*. 17(3), 257–304

[9] Bennett, R. E., Persky, H., Weiss, A., and Jenkins, F. 2007. *Problem-Solving in technology rich environments: A report from the NAEP technology-based assessment project*. NCES 2007-466, U.S. Department of Education, National Center for Educational Statistics, U.S. Government Printing Office, Washington, DC.

[10] Zapata Rivera, D. 2012. Embedded Assessment of Informal and Afterschool Science Learning. *Summit on Assessment of Informal and After-School Science Learning*. Retrieved from http://www7.nationalacademies.org/bose/1Informal_Ed_ZapataRivera_2012_Paper.pdf

[11] National Research Council. 2009. *Learning Science in Informal Environments: People, Places, and Pursuits. Committee on Learning Science in Informal Environments*. Philip Bell, Bruce Lewenstein, Andrew W. Shouse, and Michale A. Feder, Editors. Board on Science Education, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

[12] Kay, J. (2008) Lifelong Learner Modeling for Lifelong Personalized Pervasive Learning. *IEEE Transactions on Learning Technologies*, Vol. 1, No. 4, pp. 215-227.