Investigating the Effects of Varying Labels as Scaffolds for Visitor Learning

Joyce Wang, Susan Yoon, Karen Elinich, and Jacqueline Van Schooneveld

University of Pennsylvania, Philadelphia, PA
The Franklin Institute, Philadelphia, PA
joycew@gse.upenn.edu, yoonsa@gse.upenn.edu, kelinich@fi.edu, jflicker@dolphin.upenn.edu

Abstract: In museum literature, labels have been found to increase visitor learning and contribute to greater cognitive gains. In this study, we seek to understand how various labels support the visitors’ learning experience, and specifically in regards to conceptual and cognitive learning. We investigated the increasing use of three labels (visual digital augmentations and text-based questions and instructions) in four treatment groups of middle school students. Our analyses indicated that when labels were used, there were significant increases in both conceptual and cognitive understanding. Furthermore, observation notes demonstrated that in the highest treatment group, there were increased instances of problem solving behavior as well as association with prior knowledge.

Introduction

Informal environments, such as science museums, have the potential to offer rich learning experiences (NRC, 2009). Though many studies have determined that museum visits increase visitors’ curiosity, motivation, and interest in science, less have presented empirical evidence of visitors’ cognitive or conceptual gains (Allen, 2007; Anderson et al., 2003; Falk et al., 2004). As such, we resonate with this year’s ICLS conference theme, The Future of Learning, to investigate how museums can improve conceptual and cognitive learning through innovative technological advances, as well as through promising research of new approaches to learning. Most commonly, museums support visitors’ learning through scaffolds incorporated into exhibit designs. Thus, we examine how various scaffolds affect museum visitors’ learning experiences (NRC, 2009).

This study follows a series of studies that investigate how designed interactives enhance learning in the museum setting (Yoon et al., 2011a, 2011b). In this study, we seek to understand how different forms of labels support visitors’ learning. We conceive of labels as a type of scaffold that can vary in their representational nature to reveal information visitors would otherwise be unaware of or unable to acquire by themselves. Although labels are traditionally presented as printed text (typically explanatory in nature), we expand this boundary to include any form of a guide that scaffolds a learning experience. We compare the learning outcomes of middle school students placed in several treatment groups that vary in the amount and type of labels as they interact with an exhibit aimed at teaching circuits and electrical conductivity. The research question investigated was “How does the increasing use of various representations of labels affect students’ ability to make sense of the information with which they are being presented?”

Conceptual Framework

In museum literature, labels, originally defined as “written words used alone or with illustrations in museum exhibitions to provide information for visitors, presented as text on exhibit graphic panels or computer screens” (Serrell, 1996, p.239), have been found to impact visitors’ experiences at the exhibits in various ways. When they are written clearly to convey the goals of the exhibit, more visitors will understand, find meaning in, and enjoy museum exhibitions. Some studies have documented how different types of labels change the way visitors interact with exhibits (e.g. Atkins et al., 2009) while other studies have investigated how labels affect the type of conversations that ensue between group members (e.g., Hohenstein & Tran, 2007). Ultimately, labels seek to increase visitors’ learning and to contribute to greater cognitive gains (Falk, 1997; NRC 2009). In this study, we want to build upon these claims and examine different ways in which labels can be constructed to elicit deeper understanding.

We investigated two representations of three labels: visual digital augmentations and text-based questions and directions. We hypothesized that these labels, when used together, would collectively target and enable students to construct their own knowledge, and would reflect an increase in their level of thinking and learning. In reframing the conception of a museum label more generally as a scaffold, we expand upon the conventional consideration of labels as solely offering descriptive information to support the individual visitor, to include ways in which labels purposefully change visitors’ behaviors to support each other’s learning and consequently, their own learning as well.
Digital Augmentation Labels (through Augmented Reality)

At its simplest, Augmented Reality (AR) describes systems that integrate computer-generated virtual elements (objects or environments) or information with the real world environment (Mou et al., 2004). Specifically by superimposing virtual elements and images (which we define as ‘digital augmentations’) onto the real world environments, AR allows users to experience and perceive the newly incorporated information as part of their present world, thereby enhancing their perception and interaction with the real world (Kirkley & Kirkley, 2004).

In this way, AR serves as a scaffold by supporting the user with additional (virtual) information, which might not be directly detected by their senses, to aid in their performance of specific tasks. It is precisely in this scaffolding role that AR offers the unique potential to transform learning at multiple levels. For this reason, we consider these digital augmentations as a type of label in our study.

In classrooms, AR has been employed in the form of handheld games to teach mathematical and scientific concepts (e.g. Dunleavy et al., 2009; Klopfer & Squire, 2008) and to cultivate the development of inquiry skills, argumentation skills, and collaborative learning (e.g. Rosenbaum et al., 2006). For example, Klopfer and Squire (2008) developed a Pocket PC game, enabled with augmented reality, to investigate a chemical spill in a watershed that could potentially contaminate ground and surface water. One of the outcomes of the study indicated that students became active problem solvers who learned to critically analyze strategies and negotiate with each other. They were often observed arguing about different investigative strategies that they believed would give them the most reliable information to formulate a solution (Klopfer & Squire, 2008).

Although AR has been used in school settings to scaffold learning, in museums, it is still a fairly new phenomenon. While a few studies have investigated the impact of AR devices on visitors’ experiences (e.g. Szymanski et al., 2008), more have evaluated the design and creation of an AR system and the feasibility of incorporating AR into the museum environment (e.g., Damala et al., 2008). A significant finding of several of these studies reveals that AR increases visitors’ engagement and interest in museum settings when using these devices (e.g., Hall & Bannon, 2006). While this is an important component of museum experiences, this focuses primarily on engagement and access rather than on learning. In terms of conceptual and cognitive learning, there has been a dearth of research that investigates the impact that AR has on visitors.

Question Labels

Text-based question labels were also incorporated in our investigation as a method to scaffold learning. Questions, as a pedagogical strategy, have the potential to stimulate visitors to think, look, get involved with, and converse with others to ultimately learn about a particular exhibit (Serrell, 1996). In interactive exhibits, question labels can invite visitors to actively participate in an exhibit and thereby integrate their experience with the exhibit’s main message (Serrell). In some cases, they invite and challenge the visitor to use the exhibit to accomplish a certain task (Gutwill, 2006). Regardless of the method of invitation, questions in labels act to encourage physical engagement of visitors, thereby helping them to anticipate and formulate new meanings (Serrell, 1996).

A few studies have evaluated the effectiveness of incorporating questions in exhibit labels to stimulate participation and learning. For example, Hohenstein and Tran (2007) examined the use of guiding questions to stimulate conversations in an object-based science museum. Their findings indicated that in the presence of these questions, visitors were more likely to not only offer explanatory answers to the questions, but to also ask their own open-ended questions. However, it should be noted that the addition of these questions did not have the same effect on all exhibits. Thus, the context of the exhibit must also be considered when creating labels to assist visitors’ experiences. A different study also investigated the impact that questions had on visitor behavior (their learning and their conversations) in their interaction with objects at a science exhibit (Atkins et al., 2008). After including a hybrid label, a label that involved both questions and suggestions for improved visitor engagement, the exhibit was more likely to be perceived as a lesson with an intentional goal and a particular learning agenda. Specifically, in the presence of the labels, parents were observed to explain the phenomena to their children more often than when the label was not present (Atkins et al., 2008). As such, these studies indicate that text-based questions have the potential to encourage visitors to become active learners and thus engage in deeper levels of critical thinking that develop increases in conceptual knowledge.

Instructional Labels to Promote Collaboration

The last label that our study sought to investigate are labels that instruct visitors on collaboration. Instructional labels have been used to guide visitors in using exhibit devices to produce a particular outcome that illustrates a scientific phenomenon (Gutwill, 2006). In instances like these, labels provided directions that signal to the visitor whether the device worked or not depending on how the visitor operated the device. Unfortunately, the disadvantage of these types of instructional labels is that they may communicate an endpoint or finality that would inhibit further exploration by the visitor (Gutwill, 2006). While these studies have acknowledged how museums have commonly used labels to inform visitors of specific ways to interact with exhibits, there have
been a lack of studies that not only investigate the effects of labels that instruct visitors to collaborate with one
another to build knowledge, but that indicate that museums are designing these types of labels.

In formal education, collaborative learning has been found to have a positive impact upon student
achievement (O’Donnell, 2006). Regardless of how the groupings are formed, whether homogenous or
heterogenous for ability, when students collaborated together, they exhibited increased levels of achievement
and motivation (Saleh et al., 2007). A critical factor that impacts students’ learning in these groups, however, is
the extent to which structures are in place to support effective collaboration skills. Many studies have
investigated the effect of various scaffolds that support collaboration, including the assignment of student roles,
cues, scripts, and rules (Kester & Paas, 2005; Saleh et al., 2007). For example, it was found that when students
were verbally cued to summarize, explain, identify errors, or ask for help from one another, they demonstrated
higher learning gains. These studies indicate the impact and potential that collaboration, when designed for, can
have on learning in the formal environment. Consequently, these results can also offer important considerations
for informal environments.

As labels are an elemental component within the museum space, we suggest modifying them in a way
to encourage visitors to collaborate more. Unfortunately, there has been a dearth of studies that investigate how
labels have been designed to encourage visitor collaboration. Thus, we attempt to address this gap in the
literature by investigating how instructions for collaborative interaction, incorporated in text-based labels, affect
visitors’ learning and behaviors.

**Methods**

**Context & Participants**

This study leveraged a currently funded large scale National Science Foundation informal science
education project in which the central goal is to design, integrate, and increase the use of educational
technologies—and to study their impact—within the science museum learning experience (Yoon et al., 2011a,
2011b). While the larger project focuses on the viability of using augmented reality technologies in informal
science settings, the present study extended this focus by investigating the effects of various labels to scaffold
understanding. In total, 164 students, from nine diverse schools (including public, charter, urban and suburban)
participated. The participants were 6th, 7th, and 8th grade students, chosen because the device addresses
electricity and magnetism, topics which are normally covered in grade 4 of the local school curriculum. Thus,
all students would already have some prior knowledge.

In order to recruit student participants, we enlisted the help of science teachers who were associated
with either the research institution or the museum. Students’ participation in our study was embedded in a
larger school field trip. Prior to the day of implementation, researchers went to schools to administer pre-
surveys to the students. On the day of implementation, students were randomly assigned to visit a particular
exhibit from which researchers would pull students in pairs or groups of 3 and randomly assign them to one of
four conditions (C1–C4) to test the study device. Refer to Table 1 for participant breakdown and a description
of the conditions. (The variance in participant numbers for C4 is due to a larger study, consisting of 6
conditions that required greater participant numbers (Yoon et al., 2012). Whereas the larger study sought to
investigate the effects of certain structured knowledge building scaffolds on visitor learning, this study analyzes
the knowledge building scaffolds as prompts and labels that merely served to complement the device.) The
device “Be The Path” consisted of a table with two metal spheres (located at opposite ends of the table),
connected in a half loop with wires to a battery and a light bulb. As is, the loop is not fully closed (and is
referred to as an “open circuit”) so the bulb does not light. When individuals place their hands on the metal
spheres, the bulb would light up because their body closed the loop to complete the circuit. In the testing room,
students were told to play with the device as if they had come across it on the museum floor while two
researchers recorded observation notes of students’ behavior. After their engagement with the device, the
students were led to a different location to answer some questions regarding their interaction and to reflect upon
what they learned. (These reflection questions were the same questions on the question labels in conditions 3
and 4.) Following the implementation, researchers administered a post-survey (same as the pre-survey) to the
students.

**Table 1: Description of each condition**

<table>
<thead>
<tr>
<th>Cond.</th>
<th>N</th>
<th>Description</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>Control: Device without any additional scaffolds and representative of the current devices out on the exhibit floor.</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>Addition of visual digital augmentation label (DA): When students completed the circuit and made the bulb light up, a projected image (of moving electrons) would appear from overhead and map onto the device and the students’ bodies.</td>
<td>DA</td>
</tr>
</tbody>
</table>
Data Sources & Analyses

1. Pre- and post-surveys were gathered to measure changes in students’ conceptual understanding before and after the implementation. These surveys consisted of 5 multiple choice questions on electricity and magnetism as well as 1 open-ended question (“Think about an electric circuit that supplies electricity to a light bulb. What parts make it work so the bulb lights up? Use words and pictures in your answers.”) Although we included multiple choice questions (compiled from district curricular materials) to collect some baseline data of students’ general understanding of electricity, the concept that our device was demonstrating was most related to the open-ended question. A categorization manual was constructed to evaluate this last question and interrater reliability was obtained on 20% of the responses by two independent researchers ($\alpha = 0.94$). The open-ended responses were coded on a scale from ‘No Understanding’ (Level 1) to ‘Complete Understanding’ (Level 5). Paired samples $t$-tests were conducted for each condition to determine whether there were significant differences in students’ understandings before and after the intervention. An ANOVA was also conducted to determine which particular label had the greatest impact.

2. Student written reflections were solicited regarding their understanding of the science phenomenon being presented. Although there were five questions asked, only one question (“What were you supposed to learn by using this device?”) was coded, as it most clearly illustrated their ability to think critically, theorize, and make sense of the phenomenon. Thus, this question intended to evaluate changes in cognitive learning. (These reflection questions were also the same questions on the question labels in C3 and C4). A separate categorization manual was also constructed to code these responses (on a scale from ‘No Understanding’ (Level 1) to ‘Complete Understanding’ (Level 5)). Paired samples $t$-tests were conducted for each condition to determine whether there were significant differences in students’ understandings before and after the intervention. An ANOVA was also conducted to determine which particular label had the greatest impact.

3. Researcher observation notes were collected and coded to register different types of critical thinking skills that the students were seen engaging in. The categories, adapted from the Critical Thinking Skills Checklist described in Luke et al. (2007) and refined to reflect the typical experiences of a science museum, were: Participating, Interpreting, Evaluating, Associating with Prior Science Knowledge, Problem Solving, Associating with Other Knowledge, and Engaging in Teamwork. These seven categories were further divided into more refined subcategories (for example, Interpreting included behaviors such as ‘asking questions’ and ‘exploring how the device operated’) and were then coded as ‘1’ or ‘0’ depending on whether the students exhibited these components. The frequencies of these exhibited skills were then compared across the 4 conditions and a chi-square test was administered to determine if the results were statistically significant.

Results
Pre-/Post-Surveys (Measuring Conceptual Learning): Aggregated & Disaggregated for Multiple Choice and Open-Ended Questions

A paired-samples $t$-test indicated that significant learning gains were made between the pre-surveys ($M = 4.22$, $SD = 1.59$) and post-surveys ($M = 4.81$, $SD = 1.60$), $t(163) = -5.04$, $p = .01$. Specifically, as our device targeted the conceptual understanding in the open-ended response question, we then disaggregated the data to analyze these scores independent of the entire survey score (which was comprised of the multiple choice and open-ended questions). As shown in Table 2, conceptual knowledge gains were also reflected in students’ open-
ended responses between the pre- and post-surveys. Between the conditions on just the post-survey scores, there was a general increase in understanding on the open-ended questions, which were reflected in increases on the entire test as well (see Table 3 for means). However, an ANOVA conducted between the four conditions revealed that these learning gains as reflected in the open-ended question, \( F(3, 160) = 1.82, p = 0.15 \), and in the test as a whole, \( F(3, 160) = 1.25, p = 0.29 \), were not a direct result of the incorporation of labels.

Table 2: Paired-samples \( t \)-test results of open-ended responses on pre-/post-surveys

<table>
<thead>
<tr>
<th>Condition</th>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>( t ) value (df)</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cond. 1</td>
<td>Pre (OE)</td>
<td>2.17</td>
<td>1.15</td>
<td>0.47(29)</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Post (OE)</td>
<td>2.27</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond. 2</td>
<td>Pre (OE)</td>
<td>2.19</td>
<td>0.95</td>
<td>-2.72(30)</td>
<td>0.01*</td>
</tr>
<tr>
<td></td>
<td>Post (OE)</td>
<td>2.65</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond. 3</td>
<td>Pre (OE)</td>
<td>2.03</td>
<td>1.08</td>
<td>-2.04(30)</td>
<td>0.05*</td>
</tr>
<tr>
<td></td>
<td>Post (OE)</td>
<td>2.48</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cond. 4</td>
<td>Pre (OE)</td>
<td>2.25</td>
<td>0.92</td>
<td>-5.6(71)</td>
<td>0.00*</td>
</tr>
<tr>
<td></td>
<td>Post (OE)</td>
<td>2.72</td>
<td>0.97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: ANOVA results of post-survey scores (aggregated and disaggregated) between conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Post MC Scores (Mean, SD)</th>
<th>Post OE Scores (Mean, SD)</th>
<th>Post-Survey Scores (Mean, SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cond. 1</td>
<td>2.37, 1.16</td>
<td>2.27, 0.78</td>
<td>4.63, 1.50</td>
</tr>
<tr>
<td>Cond. 2</td>
<td>2.16, 0.82</td>
<td>2.65, 0.84</td>
<td>4.74, 1.29</td>
</tr>
<tr>
<td>Cond. 3</td>
<td>2.10, 0.94</td>
<td>2.48, 1.09</td>
<td>4.58, 1.59</td>
</tr>
<tr>
<td>Cond. 4</td>
<td>2.29, 1.20</td>
<td>2.72, 0.97</td>
<td>5.01, 1.60</td>
</tr>
</tbody>
</table>

Student Reflections (Measuring Cognitive Learning)
Although there seemed to be a generally increased ability to think and reason at a higher level as more scaffolds were added between C1 (\( M = 2.23, SD = 0.50 \)), C2 (\( M = 2.39, SD = 0.76 \)), C3 (\( M = 2.48, SD = 0.72 \)), and C4 (\( M = 2.43, SD = 0.80 \)), the main effect of the labels was not significant, \( F(3, 160) = 1.93, p = 0.13 \). However, an ANOVA performed between C1 and all other conditions aggregated together, indicated that there was a significant difference \( F(1, 162) = 5.84, p = 0.02 \).

Observation Notes
The observation notes reflected some differences in the students’ interactions between the four conditions. Although there were seven skills being observed, a chi-square test revealed that there were significant differences in the frequencies of only two categories: Problem Solving (\( \chi^2 = 9.166, df = 3, p < .027 \)) and Association with Prior Science Content Knowledge (\( \chi^2 = 12.731, df = 3, p < .005 \)). C4 and C2 students exhibited the greatest instances of demonstrating Problem Solving processes (represented by behaviors that indicated investigation of various configurations in lighting the bulb). Specifically, 96% and 94% of C4 and C2 students respectively reflected these behaviors, as compared to students in C3 (78%) and C1 (84%). Likewise, students in C4 tended to associate their experience with other science knowledge (usually involving recall of similar activities performed in science classes) (26%) more often than students in C3 (9.7%), C2 (13%) and C1 (0%).

Discussion & Implications
Museum studies have consistently acknowledged the importance of exhibit labels on visitors’ learning and behaviors (Atkins et al., 2009; Falk, 1997; Gutwill, 2006; Hohenstein & Tran, 2007). In accordance with this, the goal of our study was to evaluate, not only the effects of various types of labels on students’ learning, but also the effects of increasing the use of these labels on students’ conceptual and cognitive understanding of electricity. Through pre- and post-surveys (to evaluate conceptual understanding), written reflections (to evaluate cognitive skills) and observation notes (to evaluate behaviors that would indicate changes in either conceptual or cognitive understanding), we wanted to understand how students made sense of the phenomenon in both a performative and generative context.

To this end, we found significant increases in students’ conceptual understanding after the museum visit, consistent with prior research (NRC, 2009), for students in C2, C3, and C4. Looking within each condition, we found that Condition 2, 3, and 4 students, collectively, after our intervention, significantly understood better the components and operation of an electric circuit. On average, students’ understanding increased by half a point (which is reflective of half a level of understanding) – a significant result considering
that their interaction with the device was, like most interactions with museum exhibit devices, brief. This positive finding suggests that our device, when scaffolded with various labels, supports students’ learning of electricity and circuits. Additionally, at first glance, we noticed that the increase in students’ conceptual understanding was in proportion to the number of labels they received. For example, it appeared that students in C4, who experienced the most labels, exhibited the most understanding as they had the highest scores. However, an ANOVA conducted across the four conditions indicated that these differences in knowledge growth were not significant. Thus the condition, and subsequently, the labels that comprised each condition, had no significant effect on students’ understanding. A C4 student would increase their knowledge as much as a C2 student. However, as C1 students who didn’t receive any labels also did not exhibit any significant knowledge growth, this would suggest that at least one label is needed to significantly impact learning. Thus, this implies that the amount and types of labels that one can use to scaffold a visitor’s experience does not make a difference.

In regards to the impact of labels on cognitive learning, we evaluated students’ responses on their reflections. Specifically, we hoped that the digital augmentation labels, question labels, and instruction labels on collaborative work, would to varying degrees, affect students’ ability to theorize about how our bodies can complete a circuit and conduct electricity. Our results indicated that initially, it appeared that as we increased the amount of labels to scaffold their learning, they exhibited greater abilities to think and theorize at a higher level. However, similar to their survey results, an ANOVA on their reflections determined that the increases in their ability to theorize were not significant across the four conditions. Interestingly though, when we aggregated Conditions 2, 3, and 4 together to represent ‘Labeled Conditions’ and compared this to Condition 1, the differences in their cognitive scores were significant. Thus, this confirms our earlier speculation that the amount and type of labels are not as significant in learning as the presence of just one of these labels.

In addition to conceptual and cognitive learning, we were interested in understanding how these three labels affected students’ behaviors and interactions, that would help us to understand their changes in learning. A chi-square test revealed that there were significant differences in students’ behaviors in two categories: Problem Solving and Association with Prior Knowledge. Specifically, we saw a marked difference in C4 students’ interactions; as this label pointedly instructed students on ways to collaborate together, it was encouraging to see more students investigating and attempting additional configurations to light the bulb. For example, a few groups discovered that when the students all linked hands, the bulb would light up. Other students would experiment using different parts of their body including their elbows, fingertips, and the backs of their hands. The discovery of these phenomena often led to expressions of excitement and wonder. Other studies have confirmed that museums can stimulate joy, delight, awe, wonder, appreciation, surprise, etc. (NRC, 2009) and that these affective responses are linked to other forms of learning (e.g., Steidl et al., 2006). It was interesting to note that C2 students also demonstrated very high instances of Problem Solving behaviors. This would suggest to us that the novelty of the digital augmentation label by itself, was just as impactful as the instructional label in encouraging this kind of behavior. When the digital augmentation was combined with another label, specifically the question label, however, this exploratory behavior decreased. Likewise, C4 and C2 students were observed making associations with other knowledge in more instances than C3 or C1 students (though the frequency for this behavior is not as similar as that of Problem Solving). The labels in these conditions encouraged students to connect their interaction between our device and other knowledge gained from prior experiences. Some recalled similarities to other school science class activities while others remembered that they had learned about this content previously. In C4, because our instructional labels strongly influenced students to collaborate together, we observed greater amounts of dialogue relating to electricity (though this was not formally measured). Thus, it is not surprising to us that as students are talking about these questions and sharing their observations and hypotheses, their conversations would include more instances of connections to prior knowledge.

Though we did not find as many significant results as we had anticipated, we would not rush to conclude that these labels are ineffective. Instead, we suggest revisiting the context and conditions in which our intervention was administered as well as the design of the instruments themselves. First, we posit that the survey question may not have been sensitive enough to measure the smaller differences in students’ thinking and learning. We worded the question broadly so that students would not feel constrained to answer in a particular manner. However, as a number of responses were too general (i.e., “We learned about electricity.”) to offer any significant information, we were not able to accurately assess what students learned and thus, had to code their responses as “No Understanding”. Second, the variation in learning results could be explained by the different forms of assessments that were used to measure thinking (hands-on versus paper-pencil) and to the different environments in which the students were assessed. We suggest there may have been a failure of transfer between what was observed and written. Educational researchers have long since theorized that a learning experience may look differently depending on the testing context, as the manner in which the information is learned affects transfer (Bransford & Schwartz, 1999). Thus, while it appeared that students were able to think at a higher level when they were engaging with the device (especially in C4), this aptitude did not transfer once students were removed from the original context and asked to reflect upon it in writing. Lastly,
although we tried to control for this, it was apparent that many students rushed through both the reflections and the post-surveys in an attempt to hurry on to the next activity. For example, some students remarked impatiently that they wanted to finish the reflections quickly so they could return to the main exhibits and continue their exploration. Similarly, during the administration of the post-surveys, many students objected that this was exactly the same thing that they answered previously. Although the researchers explained the rationale for requiring both a pre- and post-survey to the students, informal comparisons (acquired during the collection of the post-surveys) reflected that students had not put as much effort and thought into the post-surveys as they did in the pre-surveys. In the next iteration of this study, we suggest collecting this information in a different format, perhaps embedded in interviews.

This study sought to investigate the effects of layering three types of scaffolds, in the form of labels, on students’ conceptual and cognitive learning. While the results didn’t support all of our hypotheses, we did learn that the addition of at least one of our labels significantly increases conceptual learning. Furthermore, we were encouraged by changes in students’ interactions to reflect more scientific inquiry behaviors. As we begin to plan for future phases of this study, the findings we’ve discovered here will guide our decisions as we seek to revisit the context, methods and instruments used to evaluate museum learning.

References


**Acknowledgements**

This material is based upon work supported by the National Science Foundation under Grant No. 0741659. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.