Zydeco: Exploring How Mobile Technology Can Scaffold Student Multimedia Data Collection and Use for Science Inquiry

by

Alexander Christopher Kuhn

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Computer Science and Engineering)
in The University of Michigan
2013

Doctoral Committee:

Professor Elliot Soloway, Chair
Professor Mark S. Ackerman
Associate Professor Christopher L. Quintana
Assistant Professor Georg Essl

Dedicated to
the memory of my grandmother,

Ann Surma

who could not see this work through to the end
but would have loved every page,
and to my mother,

Maria Kuhn

for her encouragement and affection
and for continuing to inspire.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the tremendous support and assistance of many people. First of all, I'd like to thank my advisor Elliot Soloway for his support and encouragement throughout the years. In addition, I am very grateful for the monumental aid from Chris Quintana, who has generously given me his advice and support throughout this process.

I am grateful to my other committee members, Georg Essl and Mark Ackerman, who provided suggestions on the research process and study design that were instrumental toward this work.

The Zydeco research group has made all this work possible. I am grateful to Clara Cahill for originating the idea and having the spark to start this project. My thanks also go out to Shannon Schmoll and Alex Pompe for coming along for the ride from the beginning, as well as to Claire Lo, Ibrahim Delen, Alex Migicovsky, and Alissa Ampezzan for their continuing aid on the project. Last (but definitely not least), I thank Brenna McNally for her tremendous assistance and support above and beyond the call of duty to help make this dissertation happen.

Thanks go out to all of my friends and family who have helped me by engaging in various discussions, offering support, and contributing suggestions throughout the years. I am extremely grateful to my mother who has been a constant supporter (and last minute reviewer) for many a paper and chapter of my research. Among these friends I would like to specifically thank Gargi Datta, for her great assistance and support, and Cory Bernieri, who was instrumental in many conversations and editing sessions.

This project would not have been possible without the support from Kira Berman and Amy Harris from the University of Michigan Exhibit Museum of Natural History, as well as the teachers and students who were involved in the research trials. This work was made possible due to the initial support from the University of Michigan Digital Media Commons GROCs Grant and the aid of Linda Kendall Knox, John Merlin Williams, and Dan Fessahazion, as well as the continued funding from the National Science Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

Introduction

This dissertation presents research looking at how to scaffold learners in collecting multimedia data outside of the classroom that will be shared with their peers and then using this collective data set to construct scientific explanations. For this study, we refer to multimedia data as the photographic and audio artifacts that are collected by students. The learner annotates these data and can do so using audio recordings, titles, or textual labels (tags).

Here I describe the motivation for looking at this area of research, what previous research has uncovered and where this research fits in. Next I describe the two scaffolded tools that were designed to mitigate the challenges students face. Following this, the research questions and structure for the research study is discussed, as well as the core contributions from this study.

1.1 Motivation

The importance of the ability of learners to collect and use data during science inquiry is set out in a Framework for K-12 Science Education, developed by the 2012 Next Generation Science Standards [NGSS] (National Research Council, 2012):

> "… all sciences share certain common features at the core of their inquiry-based and problem-solving approaches. Chief among these features is a commitment to data and evidence as the foundation for developing claims."

The NGSS calls for students to engage in practices similar to scientists carrying out investigations. This requires students to collect data from outside the classroom, as the natural world is the "laboratory" for many scientists (e.g. earth scientists, ecologists, ethnologists) (National Research Council, 2012). When conducting research, one aspect of data collection scientists engage in is collecting and using qualitative, multimedia data (such as audio recordings and photographs). Scientists must analyze the multimedia data they and their peers in the scientific community collect in order to construct explanations, answer their research questions, and communicate their findings. Students likewise need experience engaging in these practices and must collect and use not only multimedia data, but also aggregate sets of data collected by peers. These aggregate data sets can grow to be quite large (which could contain several hundred pieces of data or more) at which point students needs to manage and organize data. In order to use the data set, students must find and analyze data for use in constructing an explanation, similar to what scientists routinely engage in.

The NGSS states, "Modern technology makes the collection of large data sets much easier, thus providing many secondary sources for analysis." (National Research Council, 2012). Now scientists and citizen scientists are generating more and more data, with people posting multimedia data and sensor data online for use by scientists and citizens alike to monitor and evaluate. The act of collecting these data and annotating it in a manner that is usable to others is a growing HCI problem, which must be managed for this data to be usable in analyses (Bollier, 2010; Fisher, DeLine, Czerwinski, & Drucker, 2012). This generation of data has resulted in Big Data, a term being used in HCI to denote quantities of data that are so large they are unmanageable by current database techniques. Issues around collecting and using Big Data can be a necessary struggle for scientists, though the task can be overwhelming for novices.

While Big Data problems are beyond the scope of what novice learners are trying to manage they need authentic practices that mimic the struggles of scientists, such as collecting and managing larger data sets that must be explored and filtered to find relevant information in order to be able to address larger problems. The potential of having students collect and use large multimedia data sets has been a recent development with the rise of sensor-rich handheld mobile devices. Being able to scaffold students in sharing and using these larger multimedia data sets enable them to engage in more authentic scientific practices and help illuminate important aspects of collecting, analyzing, and interpreting data (National Research Council, 2012).

The NGSS has defined 8 scientific practices that cover the range of science inquiry activities learners must engage in (National Research Council, 2012). Three of these practices relate directly to collecting and using data:

- Planning and Carrying Out Investigations
- Analyzing and Interpreting Data

- Constructing Explanations

The first, *Planning and Carrying Out Investigations*, calls for scientific investigations to be conducted "in authentic field or laboratory environments" as appropriate to the task, and that learners should "be able to identify what data to collect and make observations on the data to construct and revise explanations" (National Research Council, 2012). *Analyzing and Interpreting Data* is the practice of analyzing data for the investigation in order to "identify significant features and patterns in the data to derive meaning". Lastly, the practice of *Constructing Explanations* has the goal of enabling students to "take into account the data they examined and construct logically coherent explanations of phenomena".

In summary, the NGSS states that students need to have an understanding of how to collect and use data in order to construct science explanations. Students should be able to:

- Collect multimedia data from the field that can be shared with their peers
- Use large (several hundred pieces of data or more) multimedia data sets that consist of their own and their peers' data to construct evidence-based scientific explanations

1.2 Challenges

However, the literature identifies a number of challenges associated with respect to the goals identified by the NGSS around having students collect and use data:

- When collecting data outside of the classroom, students are in stimulating and often unfamiliar environments without the support their teacher normally provides (Balling & Falk, 1980).
- Students need to determine what data is appropriate to collect; the data must help answer their inquiry problem and they must reflect on how to annotate and organize the data in a manner that lends itself to later retrieval and use, which can be difficult tasks for learners (Metz, 2000; Roschelle, 1995; Quintana et al., 2004).
- Students must engage in this data collection without getting caught up in non-germane but difficult tasks, such as the mechanics of data entry (Tossell, Kortum, Shepard, Rahmati, & Zhong, 2010; Songer, 2006).
- Students must search through and make sense of large amounts of student-collected, potentially inaccurate multimedia data that must be used to create explanations to answer open-ended questions with multiple correct answers – a

context which research has found to be one of the most difficult for students (Berland & McNeill, 2009).

- Searching for information is a challenging process for developing learners as learners struggle with how to approach the process of searching and construct search queries (Bilal, 2002; Druin et al., 2009; Hutchinson, Druin, & Bederson, 2007; Jochmann-Mannak, Huibers, & Sanders, 2008).
- Even without the challenge of searching in large data sets, students struggle with using appropriate and sufficient evidence to support their claims (Berland & McNeill, 2009; K. McNeill, Lizotte, Krajcik, & Marx, 2006).

This work designed scaffolds to enable students to overcome these challenges and be able to complete an inquiry activity that culminated with students using the data they and their peers collected to construct an explanation. The next section describes relevant previous works that have designed systems to address similar challenges, and reviews the areas where this work is extends the range of activities learners can be supported in doing.

1.3 Previous Work Toward Answering the Challenges

In seeking to address some of the challenges around collecting and using data, previous research has largely focused on inquiry systems within three areas:

1. Classroom desktop inquiry systems
2. Worksheets for collecting data outside the classroom
3. Mobile systems

Chapter 2 describes these areas of research in depth; here, the goal is to set the context for the work in this dissertation. Previous research systems were not able to address the challenges that arise when students are collecting multimedia data in the field and then using the large, peer-collected multimedia data set -- a context that has been created by the recent affordances of mobile technologies. Accordingly, new forms of scaffolding need to be designed to address the challenges in this context.

Previous learning technology research has focused on classroom learning, designing desktop science inquiry systems that seek to encompass some or all of the scientific practices (Bell, 1997; Quintana, 2001; Sandoval & Reiser, 2004; Suthers, Toth, & Weiner, 1997). These systems encompass the collection and use of data and were focused on a subset of scientific practices, such as using data to construct explanations. In order to collect data for the investigations, these systems had students use curated data sets or Internet resources, or students collected numerical data with probes attached to a computer in the classroom.

However, not all the initial work was looking at data collection in the classroom, as some early systems sought to have students collect data outside of the classroom. These systems sought to have students engage in collecting numerical probe data from the field then sharing the data students collect with peers across the world (Bradsher & Hagan, 1995; Means, 1998). While these systems had students collect the probe data through structured paper worksheets for later input into a desktop in the classroom, the emergence of mobile computers with a variety of sensors and wireless technology provided new affordances for these activities.

More recently, mobile systems have been used to enable students to collect and use data from the field to conduct their investigations. With some of these mobile inquiry systems, students were able to share collected numerical data with their peers (Fraser et al., 2005) (Songer, 2006; Woodgate et al., 2010), improving the ease of sharing data over the early systems that used paper worksheets. In testing these mobile data collection systems where the students share their data with their peers, researchers have found using personal and peer-collected data can make inquiry tasks more authentic and motivating (Griffin, 1998; Martin, Stanton Fraser, Fraser, Woodgate, & Crellin, 2010; Songer, 2006), can enable students to understand new perspectives on data (Fraser et al., 2005), and can allow students to view data they otherwise would not have encountered in the field.

While the previous work with mobile systems has enabled students to collect data, the focus has largely been on numerical data. Some research explored students collecting and using qualitative multimedia data from the field, but these systems had little by way of software scaffolding to help students overcome the challenges of collecting and annotating data beyond reflective prompts and guiding prompts (Laru, Sanna, & Clariana, 2010; Maldonado & Pea, 2010; Vavoula, Sharples, Rudman, Meek, & Lonsdale, 2009). Rather, these systems relied on having an expert (e.g. the teacher or a scientist) directing and scaffolding the students through the process (Laru et al., 2010; Maldonado & Pea, 2010), or limiting students to only using the data they collect (Vavoula et al., 2009).

None of these systems explored how to scaffold students in collecting multimedia data that could be shared with their peers and scaffolding students to be able to use the large, aggregate data set to construct an explanation. In the case of the Myartspace project, it was noted students had issues analyzing and interpreting collected data sets containing up to 40 pieces of datum, showing that additional scaffolds need to be designed to support the collection and usage process (Vavoula et al., 2009).

1.4 The Zydeco Project

The Zydeco Project is a National Science Foundation funded research project that seeks to scaffold science inquiry practices between the classroom and out of classroom contexts using mobile devices and to develop curriculum around these activities. The project's goals investigate five general areas:

- *Curricular*: What kind of science inquiry activities can we consider given these devices?
- *Instructional support*: How do we support teachers in planning and developing activities with Zydeco and then carrying out said activities?
- *Bridging Formal and Informal Environments*: What issues arise in both the formal classroom and informal environments, particularly when students transition between these environments working on a project?
- *Design of the System*: What kinds of designs are appropriate to these activities, to these devices, and to different learner audiences'?
- *Scaffolding the Inquiry Activity*: How can we support learners through their investigation to reflect and perform sensemaking in order to develop scientific explanations?

The project is investigating these areas and supporting six practices set out by the Next Generation of Science Standards (National Research Council, 2012): asking questions and defining problems, planning and carrying out investigations, analyzing and interpreting data, constructing explanations, engaging in argument from evidence, and obtaining, evaluating, and communicating information.

I contributed to the Zydeco Project by focusing on the design of mobile computer-based apps that scaffold students through the collection and use of large amounts of peer-collected, annotated, multimedia data to construct scientific explanations. Zydeco:CollectData, which was designed by the team of researchers (including the author) and entirely developed by the author, and Zydeco:UseData, which was entirely designed and developed by the author. These two tools, shown in Figure 1-1 and described in the use case in Section 1.5, separate the processes of collecting data and using it for explanations:

1. Zydeco:CollectData – running on a mobile device, (e.g. Apple's iPod Touch) this app scaffolds students engaging in multimedia data collection outside of the classroom for later use by themselves and their peers, and
2. Zydeco:UseData – running on a tablet, (e.g. Apple's iPad) this app scaffolds students in using large amounts (several hundred pieces) of personal and peer-collected multimedia data to construct scientific explanations.

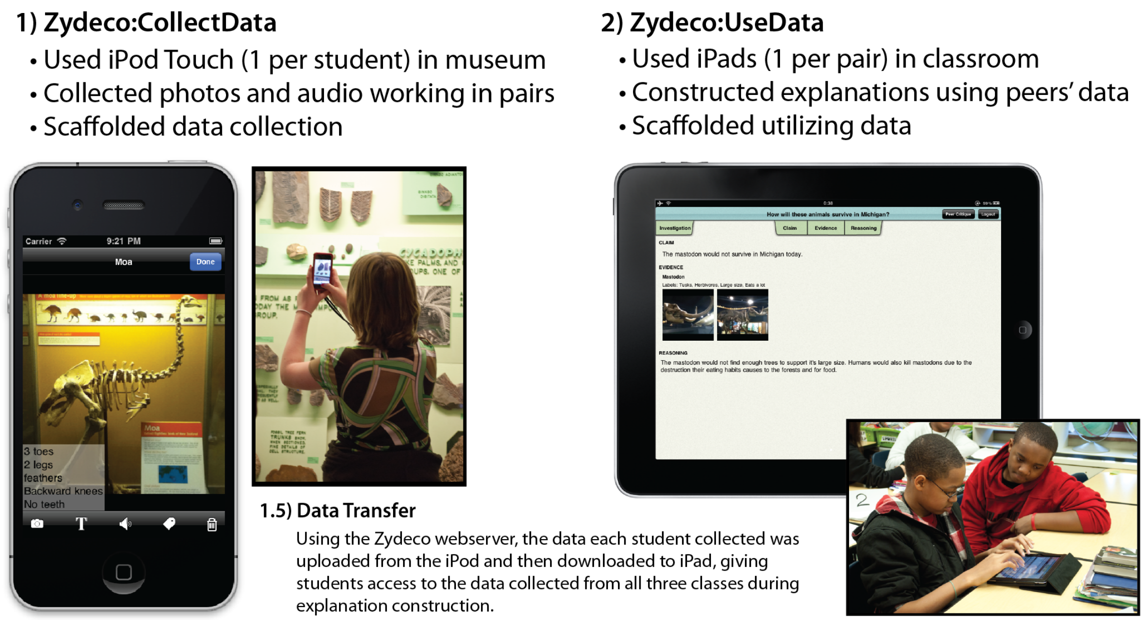

Figure 1-1. Overview of the two integrated tools and how students used the system.

1.5 Use Case

A use case of how students interact with the Zydeco tools, based on a composite of real scenarios, is presented here to anchor the research study and the more abstract discussions on the Zydeco tools that will come in the following chapters. The focus is on how students use the Zydeco tools to collect and annotate data in the field then use the data to construct a scientific explanation. Perspective for the activity is provided by outlining the pre-activities that prepared students to use the Zydeco tools as well as the post-activities students underwent after developing an explanation. During these activities, the teacher acted as a facilitator with the students primarily directing their own learning. This instructional setup followed the practice of Project-Based Science (PBS), a social constructivist based approach where students work in small groups and take on in-depth inquiry projects (Krajcik & Blumenfeld, 2006). The activity structure is detailed in Figure 1-2, with students spending approximately one hour each day on the activity.

1.5.1 Step 1: Preparation Inside the Classroom

Alice and Bob are two of 30 students in a 6th grade classroom. Ms. Smith is their science teacher and has assigned them to be partners for the upcoming activities. Their goal is to learn about animal traits and how animals are similar to each other. To investigate this, the students are presented with the driving question, “How is my animal related to other animals?” To begin, Ms. Smith has each pair choose their favorite animal from a pre-determined list: Grey Wolf, Five-Lined Skink, Sandhill Crane, Brown Bat, White-Tail Deer, or Four-Toed Salamander.

Alice and Bob choose Sandhill Crane and spend the next two days researching their animal. The students are given a list of websites to review for information on their animal and fill out a worksheet (Appendix E) as they proceed through examining various websites.

Schedule of the Activities Students Performed

WEEK ONE

M
Tu

Preparation Inside the Classroom

Day 1-2: The teacher introduces driving question "How is my animal related to other animals?" Students work together in pairs and choose their animal from list of six animals and begin researching their animals (following worksheet in Appendix E).

No School W-F

WEEK TWO

M

Day 3: Students rotate between stations in the classroom to practice grouping animals by similar traits: 1) animal skulls, 2) photos of animals, 3) photos of animals and habitats, and 4) pelts and feathers.

Tu

Day 4: The teacher explains what the students will be doing in the museum and provides an overview of Zydeco:CollectData. Students write down how they might describe their animal’s internal and external traits and discuss them with the class. Students use the tool for a few minutes to get an initial exposure to the application.

W

Day 5: Students practice using Zydeco:CollectData with animal specimens in the classroom. Zydeco:CollectData contains a list of potential tags based on what students wrote on the previous day.

Th

Day 6: The teacher reviews the data students collected and discusses appropriate annotations and inappropriate ones. Students are given an overview of the museum trip, have time to discuss and revise their list of traits (to be used as tags), and take a pre-survey (Appendix C).

F

Collecting Data Outside the Classroom

Day 7: Students spend an hour in the University of Michigan Exhibit Museum of Natural History using Zydeco:CollectData to collect and annotate data on animal artifacts.

WEEK THREE

M

Using Data to Construct Explanations Inside the Classroom

Day 8: The teacher gives an explanation of the Claim, Evidence, and Reasoning Framework by discussing personally relevant examples. At the end of class students go through a tutorial on the Zydeco:UseData system.

Tu
W
Th

Day 9-12: Students construct explanations using Zydeco:UseData to answer, "How is my animal related to (3) other animals?". Students make a claim about how three animals are ranked related to their animal and find data to use as evidence and give reasoning as to why the evidence supports their claim. After making an explanation, pairs swap devices and complete peer critiques. After making revisions, students construct an explanation to answer "Describe how these animals could survive in Michigan", making claims about two of the animals (not necessarily from Michigan) discussed in the first explanation. Students do peer critique on the 2nd investigation when finished.

F

Day 13: First half of class: Students finish their work and prepare to present to their peers. 2nd half: Students get together in groups along with an adult mediator and present their explanations to each other.

Figure 1-2. Overview of the schedule and activities students performed over the course of the study.

The next day Ms. Smith sets up four stations that students rotate between to examine animal skulls, pelts and feathers, and photos to determine how related different animals are and learn about the habitats they live in. Alice and Bob rotate between these stations and learn about different animal internal traits, external traits, and habitats.

Ms. Smith discusses that students will be visiting a natural history museum and collecting data on animals to determine how related they are to their chosen animal. Ms. Smith discusses how animal characteristics around relation can be categorized as an internal or external trait. While collecting data, the students' goal is to determine "What are similar internal traits?" and "What are similar external traits?" to their animal. These are the sub-questions the students will be collecting data under in order to answer their driving question on how their animal is related to other animals.

Ms. Smith provides an overview of the Zydeco:CollectData tool and discusses with the class how they might describe the data they collect on other animals. The teacher asks the class to write down potential labels for the data, which are meant to be around their chosen animal's traits. Alice and Bob write down the internal and external traits the Sandhill Crane has (e.g. two legs, feathers) and Ms. Smith leads a class discussion on the potential labels the groups have chosen. During this discussion Ms. Smith prompts the class to describe what each trait means and corrects students who have created inappropriate traits. At the end of the class, students are able to spend a few minutes getting practice with Zydeco:CollectData, logging into the tool and experimenting with collecting data.

The next day, some animal specimens from the museum are brought into the classroom for students to examine. Alice and Bob are given an opportunity to practice using the Zydeco:CollectData tool to collect data on internal and external animal traits.

Zydeco:CollectData starts by breaking apart the overall investigation into answering sub-questions, and using this as the first step in guiding the data collection process (Figure 1-3). Alice and Bob start by examining a monkey skull and determining it belongs to the "What are similar internal traits?" sub-question.

The software then prompts and requires them to title the photo, which has several preset titles relevant for the class activity (Figure 1-4). They select "Monkey" and are prompted to next provide an (optional) audio note, but decide to dismiss the prompt.

The software guides them to the final step of the annotation process, where they are prompted and required to add at least one tag to the data (Figure 1-5). As they are answering the internal traits sub-question, the tool lists the internal traits created through the class discussion on the previous day to review and they can tap to apply these traits as tags, or create their own. They tag it with *nostril hole on front face* and *few holes in skull*. This completes the data collection process, and they move on to another station to examine other artifacts.

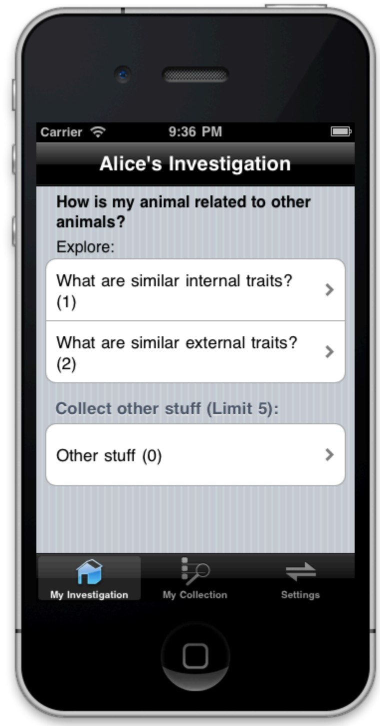

Figure 1-3. Reviewing sub-questions to collect data.

The next day, Mrs. Smith discusses a few pieces of data the class collected, highlighting appropriate and useful annotations that can describe how the animal they were investigated is related to other animals. The class is given the opportunity to revise their list of internal and external traits, which are combined with the lists from all of the other groups in the 6th grade that chose that animal.

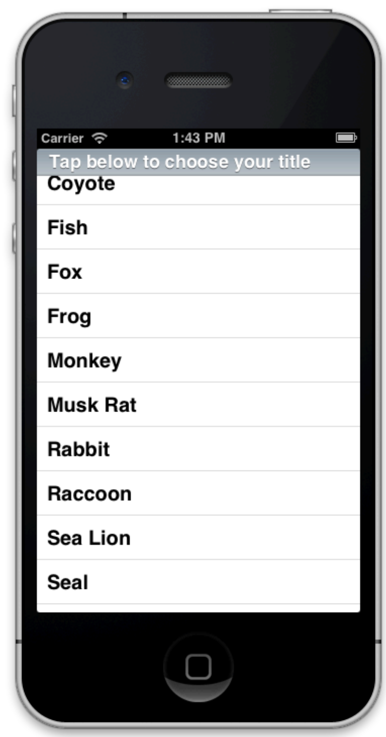

Figure 1-4. Examining a preset list of titles to title the data.

The teacher leads a discussion on these different traits students provided to ensure all the students understand what each means and discussing inappropriate traits students select. This list of traits (Table 1-1) in effect becomes a co-created language that the students and teachers agree upon and understand. To help Alice, Bob, and their peers focus on finding animals with similar characteristics during the museum field trip, the co-created list of traits is input into the Zydeco:CollectData tool to serve as potential tags for them to use while collecting data:

Sub Question: What are similar external traits?		
2 legs	Feathers	Long neck
Long and skinny legs	No teeth	Long and pointed beak
2 wings	Backward knees	Tail feathers
3 toes	Nose on top of beak	Eyes on side

Sub Question: What are similar internal traits?		
Hollow bones	4 chambered heart	Long jaw
Small brain	Big eye holes	Lungs
Multiple air sacks		

Table 1-1. A list of potential tags students came up with to represent the internal and external traits of a Sandhill Crane.

1.5.2 Step 2: Collecting Data Outside the Classroom

At the museum, Alice and Bob are each given an iPod Touch to use during data collection, loaded with the Zydeco:CollectData tool. As they wander through the exhibits looking for interesting animals that may be related to the Sandhill Crane, they can review the sub-questions they are answering on their handheld (Figure 1-3) to remember what data to look for.

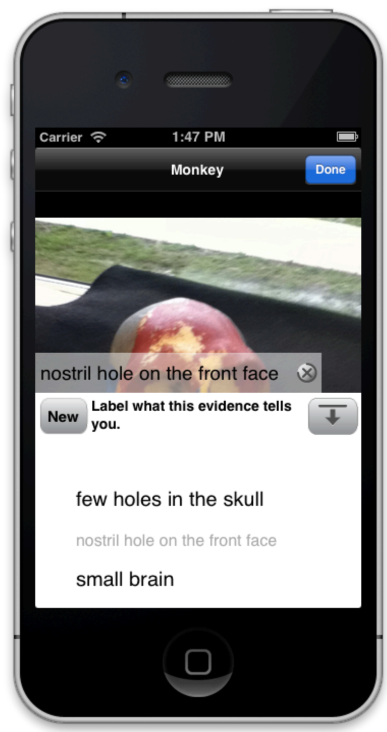

Figure 1-5. Tagging interface displaying tags created and discussed in the classroom.

Bob spots a Moa and points it out to Alice. Alice remarks that it has several similar external traits- such as *feathers*, *two legs*, and *no teeth*- as well as similar internal traits- such as *holes in the skull*- and decides to collect data on external traits to answer the "What are similar external traits?" sub-question, tapping on the question to begin data collection. Alice then chooses to collect a photograph and takes one of the Moa. She is prompted and required to title the photo. Alice can review a list of 165 potential titles (names of many of the animals in the museum), can be panned through or filtered as the she types. She scrolls down to find "Moa" on the list (Figure 1-4). After selecting the title "Moa" from the list, she is prompted to record an optional audio note and decides to read off the information she sees on the exhibit's museum placard. After she finishes recording the audio note, the tool guides her to the final step in the data collection process, which is to annotate the data with tags (requiring at least one tag).

The list of traits the students co-created that reflect teacher discussion, is input into Zydeco and she is able to tag the photo with the traits she noticed by simply tapping on the ones she desires (Figure 1-6); this prevents Alice from having to type the trait in and struggle with data entry and spelling issues. *Small head*, however, was not one of the traits they added to the list, and since she thinks that is a good similarity she types that one into Zydeco herself. Once she finishes tagging the data, she reviews it briefly (Figure 1-7) before moving on to collect more data.

By the end of the field trip, Alice and Bob each collect 8 pieces of data. They have tagged each piece of data with 2-3 tags, and they each have two pieces of data that are additionally annotated with an audio note.

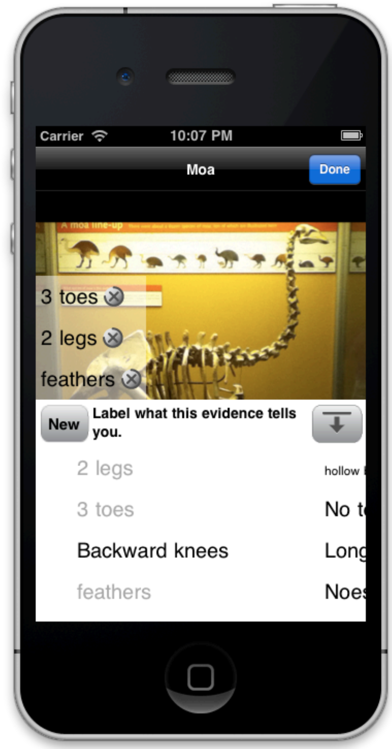

Figure 1-6. Tagging interface displaying tags previously discussed in the classroom (tap to apply).

1.5.3 Step 3: Using Data to Construct Explanations Inside the Classroom

Back in the classroom, Ms. Smith introduces the students to the Claim, Evidence, and Reasoning Framework that was developed to assist students in constructing explanations by breaking the task into the following manageable parts: making a claim, providing evidence to support the claim, and presenting reasoning that links the evidence to the claim. (McNeill & Krajcik, 2011). Each pair is then given an iPad loaded with their peers' data in the Zydeco:UseData system. After being introduced to the system and completing a short tutorial, Alice and Bob begin working on the first investigation, *Determine how three other animals are related to your animal*, which could always be seen at the top of the screen. They are taken to the investigation guide, where they can see the steps needed to complete their investigation. Alice hits "Go Here" next to the prompt that tells them to review the collected data, and is taken to the Evidence page (Figure 1-7).

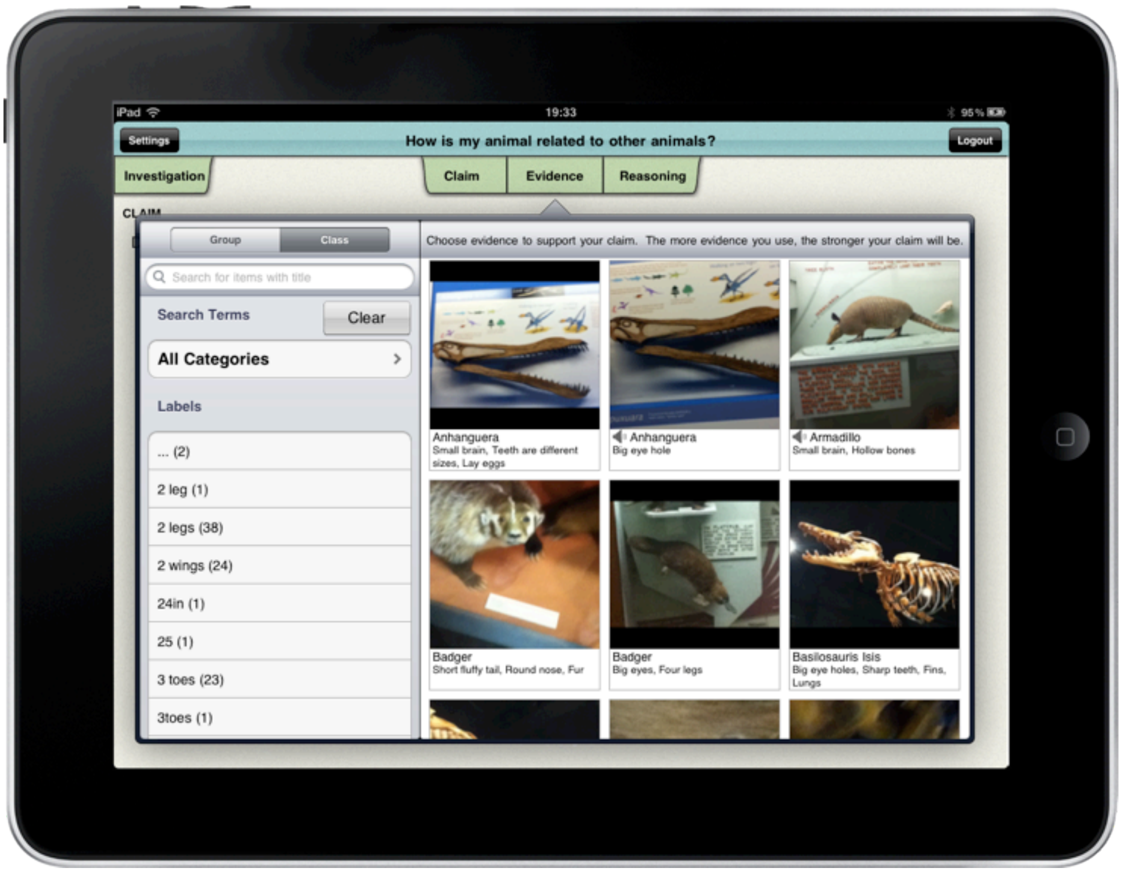

Figure 1-7. Zydeco:UseData evidence page where learners can review the collective data set and search and examine data.

As Alice scrolls down through thumbnails of the data collected by the class (Figure 1-7), Bob suggests that they look at the data they collected. By tapping “group” at the top of the screen, Alice and Bob are able to look at the 16 pieces of data that they cumulatively collected on their field trip. Just like in the class data, they are able to view the titles and tags they applied to each piece of data below each image thumbnail.

Bob wants to take a better look at the data he collected on the Moa. He taps on the thumbnail to enlarge the image and can now listen to the audio note he took in the museum (Figure 1-8). Alice really likes the tag *small brain* that Bob had put on the Moa, and tells him that they should search for that tag in the class data. Scrolling through the list of tags on the left of the screen (above in Figure 1-7), Alice finds and applies the tag “Small brain” as a search term. All the data annotated with that tag appear, including an image titled Deinonychus, which also has the tags *backwards knees* and *feathers*— matching the Sandhill Crane! Bob gets excited and does a title search for more Deinonychus data. There are two data with really good pictures of the animal, but one of them has a tag “fluffy” that Alice thinks is silly, so Bob selects the other picture. The tags on this picture are “backwards

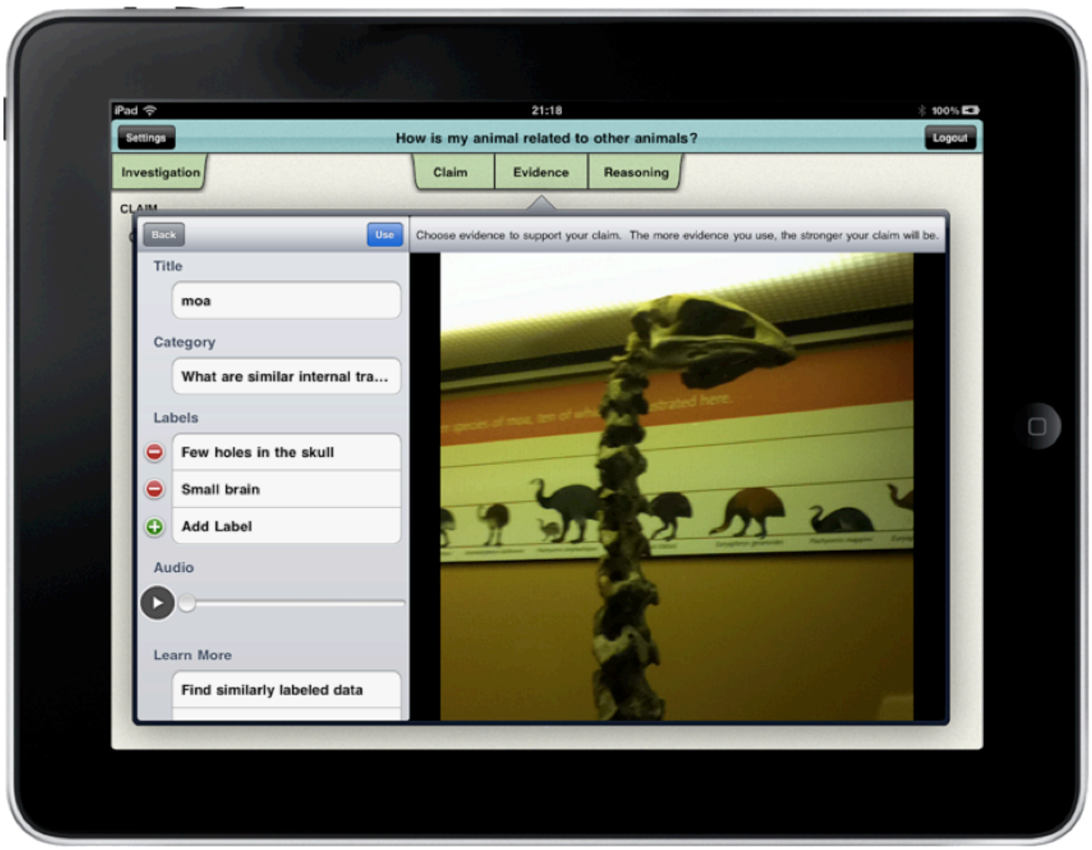

Figure 1-8. Zydeco:UseData full screen evidence viewing to assess the "Moa" datum further.

knees", "hollow bones", and "feathers"; Alice and Bob agree that this is a good piece of data because the tags match both the image and their animal so they select "use". They create an evidence group labeled "Deinonychus" and are taken back to their explanation page.

Alice and Bob continue searching for new evidence until they have three groups: Deinonychus, Moa and Pigeon. Alice taps on the Claim tab and, following the prompt, writes that, "The mostly closely related animal to the Sandhill Crane we have is the Deinonychus, followed by the Moa then the Pigeon". Bob agrees with her, and for the reasoning he writes, "Because animals that have the most traits in common are more closely related, and the Deinonychus has six traits its in common with the Sandhill Crane— like backwards knees, eyes on side, and feathers— that it is more closely related than the Moa or Pigeon because they have fewer traits in common." They now have finished their first draft of their first explanation (Figure 1-9).

After finishing their first explanation Bob opens up the Investigation Guide to find out what to do next. It instructs them to raise their hand and ask for a peer critique. Their iPad is swapped with another group's, and they have the opportunity to go over the explanation

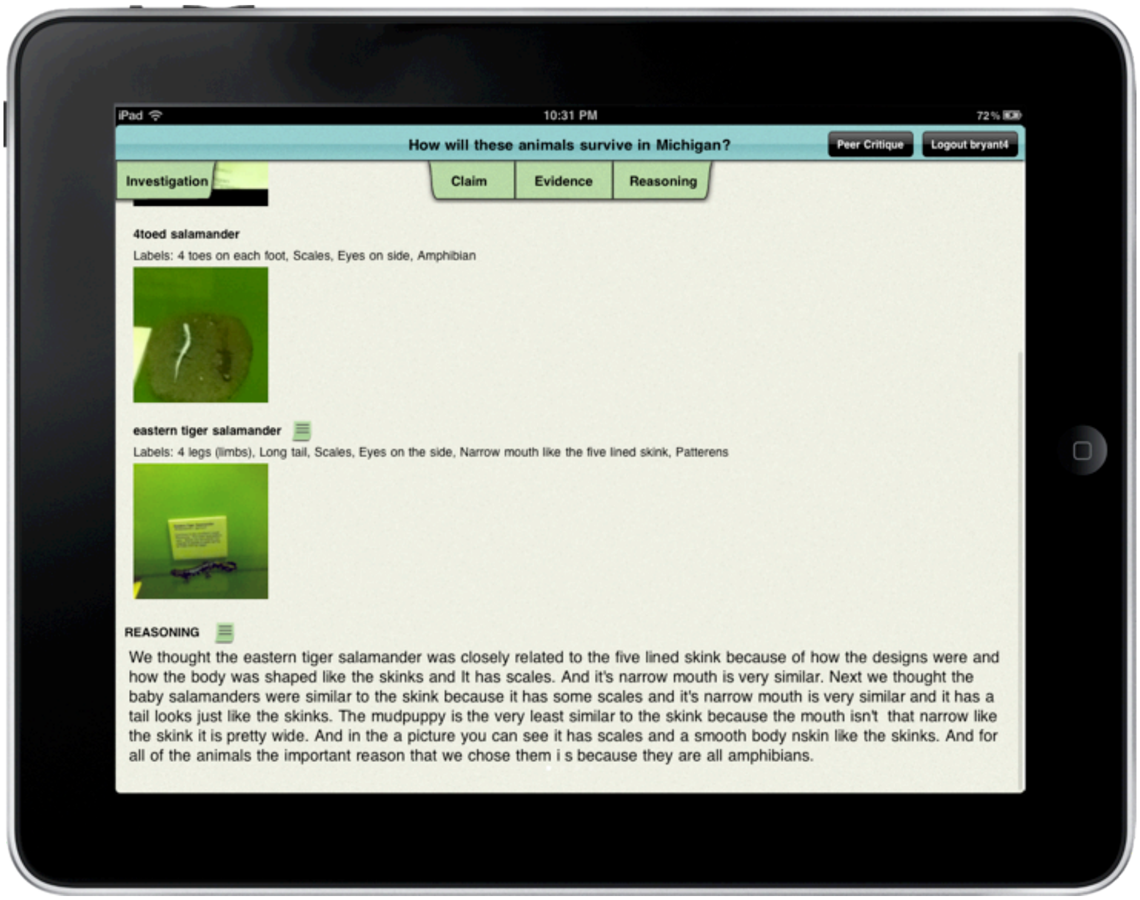

Figure 1-9. Student explanation constructed in Zydeco:UseData with two pieces of evidence and their reasoning visible.

another group made. The other group's animal was the Grey Wolf, and though they had two good explanations with evidence about how Dire Wolves and Coyotes were related, they had not ranked a third animal. Under the peer critique note for Claim, Bob selects the preset critique "Is incomplete" and adds a personal note that, "you need a third animal." In the Evidence peer critique note Alice suggests they just select the "Add more evidence" preset, and choose "Needs more information" for the reasoning.

When they get their investigation back, Alice and Bob can review the notes the other group left for them and make changes to their explanation. Their claim and evidence both had preset critiques selected saying "Is good," but their reasoning had a preset selected that said, "Needs more information," and there is an additional note saying, "You don't say if the Moa or Pigeon is closer." They add that information to their reasoning then use the Investigation Guide to move on to the next assignment, which is to "*Describe how these animals could survive in Michigan*".

After completing this process for the second investigation, Alice and Bob get together with three other groups and a teacher to present their explanations. Using the iPads they are able to point out their evidence and play the audio notes on the evidence they selected as they talk about their claim, evidence, and reasoning for each investigation.

1.5.4 Summary of Use Case

This use case presented a tour through the inquiry processes the Zydeco:CollectData and the Zydeco:UseData tools seek to support. The goal of this study is to focus on the aspects of collecting and using data and to understand how students use the scaffolds to accomplish these tasks (Step 2 and Step 3 in the use case). Students received a variety of scaffolding from their teacher and software as they collected data from the field and used these data:

In class preparation (instructional support)

- Students and teachers discuss how data can be described and annotate and co-create a language they will use to describe the data, instantiated in the form of co-created tags that students will use to label the multimedia data they will collect outside the classroom.
- Students get practice using the system before leaving the classroom.

Zydeco:CollectData System outside the classroom

- The tool integrates the co-created tags, which are a reflection of a discussion with an expert, as a form of scaffolding; students use these tags to annotate the multimedia data they collect.
- The tool contained additional scaffolding to help guide the students through collecting data to answer questions related to their investigation and annotating data in a manner that will be useful for answering their investigation.

Zydeco:UseData System in the classroom

- The tool used data annotated with co-created tags that reflected the expert discussion so students were familiar with the semantics of the tags and how they could be used, which supported students in analyzing their peers' data.
- The tool had scaffolds around filtering and reviewing the data set, supporting exploratory search methods and finding supplemental information in order to use data to construct a scientific explanation.

1.6 Research Study

Previously I described the design and provided a use case of the Zydeco:CollectData and Zydeco:UseData tools. The scaffolds within these tools are being used to evaluate the following questions:

Overarching Research Question: What scaffolds can mitigate the challenges for students inherent in collecting multimedia data as well as using a large, student-collected, multimedia data set to construct explanations?

This overarching question is being investigated by seeing how students perform the tasks of data collection and use given the scaffolds implemented in the Zydeco tools, and is broken down by two corresponding research questions**.** From understanding how students completed the tasks given the scaffolds, the benefits these scaffolds provide and areas for improving the scaffolds can be noted.

Research Question 1: How do students collect and annotate multimedia data given the system's scaffolds?

The goal is to enable students to collect data with accurate annotations that are useful for later analysis and use constructing a scientific explanation. In order to be usable later, the data needs to be annotated in a manner that students can search through and understand when constructing a scientific explanation. Analyzing the characteristics of the data students collect can assist in ensuring scaffolds are built around using the data.

This research question is being analyzed by looking at how students go through collecting and annotating data, focusing on the sort of annotations students apply to the data and whether the annotations are factually accurate and potentially useful towards answering the investigation. From this evaluation, the scaffolds can be revised and areas where additional scaffolding is needed may be discovered.

Research Question 2: How do students use data from the student-collected data set given the system's scaffolds?

Students are carrying out an investigation to answer a question. To answer this question, students must engage in constructing an evidence-based scientific explanation and must analyze data to use in constructing the explanation. The process for using data can entail several steps, which are not necessarily linear:

- Find: Students must find data they want to look at, potentially filtering through the large data set.
- Assess: Once they find potential data, students need to analyze the data, determining if the annotations are factually accurate and if the data is useful for the explanation they have in mind, or to potentially revise their idea about their explanation due to the data they viewed.
- Select: Upon finding useful data, they must make a decision on what data to select for use as evidence in their explanation.

In order to analyze data for use in constructing scientific explanations, the students not only need to collect multimedia data with accurate and useful annotations (using Zydeco:CollectData), but need to have support in searching through their personal and peers' data in order to find and assess data they will use in constructing an explanation (in Zydeco:UseData).

When looking at how students "use" data this research is focusing on how students go about finding, assessing, and selecting data. Though a variety of factors contribute to the quality of the students' final explanations, the final explanations students construct can be an indicator of whether they were successfully scaffolded through the processes of using the tools. As such, this research analyzes the explanations students constructed as a metric to see if they were able to use data in a successful manner.

1.6.1 Overview of Results

Described in depth in Chapter 5, I found that the scaffolds of the two Zydeco tools were able to enable students to collect and use large amounts of annotated, peer collected, multimedia data to construct science explanations.

In evaluating the first research question, "How do students collect and annotate multimedia data given the system's scaffolds?", it was found that the scaffolds were overall successful at enabling students to collect multimedia data with accurate and useful annotations towards answering their investigation (graded by researchers). Twenty-seven pairs of students (aged 11-13) used Zydeco:CollectData for an hour in a natural history museum, collecting 434 annotated multimedia data objects. Specifically, it was found that:

- The average accuracy of tags per group was 88% and all but 1 accurate tag were potentially useful to the investigation.

- The average accuracy of titles per group was 83%, and an average of 85% of titles were correctly spelled per group.
- The average accuracy of audio notes was 70% per group, and 64% of these audio notes per group were potentially useful to the investigation.

In evaluating the second research question, “How do students use data from the student-collected data set given the system’s scaffolds?”, it was found that the students were able to construct scientific explanations with the data they and their peers collected in the field. Students were able to take advantage of the scaffolds around using the data set to find data and then to assess it for use in their explanation. The different data annotations (titles, tags, and audio notes) were used for a variety of useful purposes:

- Tags were most often compared to images as a sense making support to learn about the image and fact check the data.
- Titles were a means to identify the animal and search upon the data set.
- Audio notes were a method to learn more about the animal and remember why the student collected the data later.
- Twenty-two of the 27 pairs selected data primarily from the collective data set and the main means students used to find data they selected was through title searches (accounting for 47% of all the data selected to use as evidence).

1.7 Contributions

This research presents three main contributions:

1. Design of two integrated tools that provide software scaffolding and reflect instructional practices.
2. Evaluation of how students used the scaffolds in the tools.
3. Design guidelines for other designers to follow in developing scaffolds to assist learners in similar tasks.

Integrated System Design

This research presents the design of two integrated tools for collecting data (Zydeco:CollectData) and using data (Zydeco:UseData) that can be used both inside and outside of the classroom. These tools have scaffolds (described in chapter 3) around using a guided annotation system during data collection and supporting students in using the data as they go about finding, assessing, and selecting data from the collective data set.

Evaluation of How Students Used the Scaffolds

In order to understand how to scaffold students as they engage in multimedia data collection outside the classroom and use their peers' shared data in constructing scientific explanations, this research needed to uncover how students use the scaffolds in the tools and interact with the annotated data in an educational context. Evaluating how students collect data provides a sense of how students approach the data collection process and what sort of annotations they apply to the data. By interviewing students on how they assess the data's annotations as well as looking logs of their data usage in finding and selecting data, information was learned as to which annotations on the data are useful, whether they are able to search through and find data, and their preferences on how they retrieve data and which data they use (discussed in chapter 5). An evaluation of the benefits and areas for improvement on each scaffold is covered in chapter 6.

Design Recommendations

From the evaluation of how students were able to use the scaffolds and areas where students struggled in the task, an emergent set of design guidelines were formulated (discussed in chapter 6). These guidelines can provide support for designers to follow in developing scaffolds to assist learners in collecting and using multimedia data, particularly around activities where learners share and aggregate this data to use a large data set.

1.8 Dissertation Structure

This dissertation explores the background research on existing science inquiry systems, and scaffolding design guidelines that have come out of these systems in chapter 2. Chapter 3 presents the design rationale of the Zydeco:CollectData and Zydeco:UseData tools and describes the scaffolds they each have. Chapter 4 provides a breakdown of the research participants, sites, and the methodology I used in collecting and analyzing the data. Chapter 5 provides the results of students using the Zydeco tools in a field trial, broken down by each research question. Chapter 6 provides a discussion of the implications and design recommendations drawn from this work. Chapter 7 presents a summary of the conclusions and contributions from this study, limitations in the research, and future directions of work.

CHAPTER 2

Background and Related Work

The goal of this work is to understand how to design scaffolding to support students collecting multimedia data from outside the classroom with mobile devices and later use their personal and peer-collected data to construct scientific explanations. This chapter begins by discussing an overview of science inquiry, including instructional practices around Project-Based Science and the tasks and challenges students face while collecting and using data. Next I discuss related science inquiry systems that have sought to address the challenges students face in collecting and using data, and show where this study on Zydeco is expanding the literature. The scaffolding guidelines that have been developed from these related systems are also mentioned, which can inform the design of the Zydeco tools. Lastly, I look into related research outside of the inquiry realm that can be relevant towards designing software to scaffold collecting and using multimedia data.

2.1 Background on Science Inquiry

Science inquiry, the act of posing questions and gathering and evaluating empirical evidence, is a critical component of science education in K-12 settings (American Association for the Advancement of Science, 2009). The Next Generation of Science Standards (NGSS) breaks down science inquiry into 8 scientific practices (National Research Council, 2012):

1. Asking Questions and Defining Problems: Asking a question about phenomenon and determining what questions have to be answered.
2. Developing and Using Models: Establishing models that help develop explanations and can be used to test hypothetical explanations.
3. Planning and Carrying Out Investigations: Requires knowing what needs to be recorded and testing theories with data collected.
4. Analyzing and Interpreting Data: Identify significant features and patterns in the data and identify error.

5. Using Mathematics and Computational Thinking: These tools for representing variables and their relationships enable prediction and assessing patterns.
6. Constructing Explanations: Construction of logically coherent explanations of phenomena that is consistent with the available evidence.
7. Engaging in Argument from Evidence: Defending explanations and examining their understanding in light of evidence offered from others.
8. Obtaining, Evaluating, and Communicating Information: Clearly and persuasively communicate findings and be able to understand scientific texts.

These are practices students need engage in, and they may participate in all or a subset of these practices when conducting a scientific investigation.

2.1.1 Instructional Strategies for Science Inquiry

The tools and curriculum design used in Zydeco research are following a social constructivism approach in teaching science inquiry called Project-Based Science (PBS), which focuses on students learning by working together engaging in science (Krajcik & Blumenfeld, 2006). In PBS students take on in-depth class projects as they try to answer a driving question, such as "Why are the fish dying in the nearby lake?" Students work in pairs or groups and investigate questions that are personally relevant to them in order to situate the problem as more meaningful and engaging to the students (Blumenfeld et al., 1991). During PBS, the teachers act as facilitators to help guide and direct the students as the students are pursuing their investigation (Krajcik & Blumenfeld, 2006). Specifically, a teacher can answer students' questions and help resolve issues as problems occur, but try to avoid leading or directing groups more than necessary. It has been shown that when students are given the ability to investigate the artifacts and phenomena that draw their interest to answer their inquiry, they can be more engaged and have improved educational outcomes compared to when they are led and directed by the teacher on what to do (Bamberger & Tali, 2007).

Another important method used in educational theory and the social constructivist theory of learning is the idea of scaffolding students (Quintana et al., 2004; Suthers, 1998) (Sedig, Klawe, & Westrom, 2001). Scaffolding[1] is the process by which a more

---

[1] There is a distinction between scaffolds and supports in education literature. Supports refer to structures to assist a learner in overcoming the challenging of a task and would never fade as the learner grows more experienced, being useful to helping an expert complete a task as well. Scaffolds are structures that would fade and be removed as the learner becomes more experienced and no longer needs their assistance. However, for the sake of brevity in this dissertation we are referring to these two types as scaffolds as this research does not investigate the fading aspect of scaffolds over the course of the study.

knowledgeable person (be it teacher or peer), instructional support, or software assists the learner and enables them to accomplish problems that otherwise would be too difficult (Quintana, Soloway, & Krajcik, 2003; Wood, Bruner, & Ross, 1976). For example, a teacher can prompt the learner to think through difficult parts of a task, or provide guidance on how to approach the activity and thus provide scaffolding. Software scaffolding is similarly providing support and is discussed in depth in Section 2.2.

The goal of many inquiry investigations is to construct an explanation that answers the question being investigated. However, the task of constructing explanations requires scaffolding for students, being a difficult task for newcomers to inquiry (Berland & McNeill, 2009). There are several simplified frameworks for constructing explanations that provide this support for learners, such as the Evidence, Claims, and Explanations framework (Bell, 1997) or the Claims, Evidence, Reasoning framework (McNeill & Krajcik, 2011). For this research I chose to use McNeill's Claim, Evidence, and Reasoning framework (Figure 2-1) as it has been found to be accessible to middle school students and their teachers (McNeill & Krajcik, 2011). McNeill's model defines a claim as a conclusion to a question or problem, evidence as scientific data that supports the claim, and reasoning as a justification that links the evidence to the claim (McNeill & Krajcik, 2011).

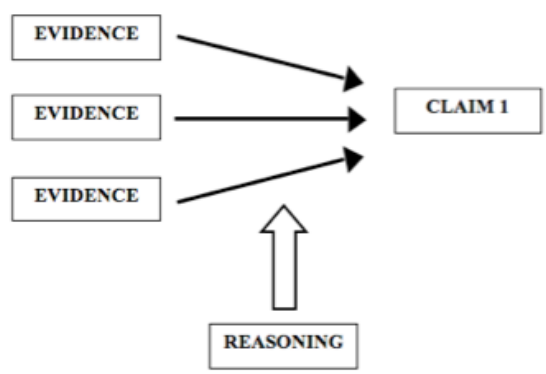

Figure 2-1. McNeill's Claim, Evidence, Reasoning framework.

Even with this scaffolding, the inquiry process that leads up to constructing explanations (which includes collecting and using data) is a difficult task for learners (Berland & McNeill, 2009; Lizotte, Harris, McNeill, Marx, & Krajcik, 2003; McNeill & Krajcik, 2008; McNeill et al., 2003; McNeill & Krajcik, 2011; Sadler, 2004). These processes that lead up to constructing explanations include collecting and using data.

2.1.2 Collecting Data

Collecting data falls under the NGSS practice of *planning and carrying out an investigation*. This work is focused on this data collection aspect; the tasks and challenges students face while collecting data as they carry out investigations outside the classroom are discussed in detail.

When students are collecting data outside the classroom they are in a rich environment that has much sensory stimulation and is often unfamiliar to the students. This has the potential to overwhelm students and add additional cognitive load (Balling & Falk, 1980). Additionally, the students who are outside of the classroom do not have access to the same level of teacher support that is present in the classroom. Research has shown that repeating out-of-classroom visits to the same setting can greatly reduce this cognitive overload on students and increase on-task behaviors (Falk, Moussouri, & Coulson, 1998). However, this study looked at the more common school field trip experience, which involves only one visit to a location, in this case to a natural history museum. While outside the classroom, there are two cognitive tasks the students must perform: 1) determining what data to collect to answer the inquiry, and 2) annotating and organizing the collected data for later retrieval and usage.

Determining what Data to Collect

The task of collecting data is a difficult one for students. Initially, students must reflect on what data they will need to answer their investigation. This is important because students need to collect data that is pertinent to their investigation in order to answer their driving question.

In the context of inquiry outside of the classroom, data collection is happening in a novel and foreign environment where many different artifacts can vie for their attention. This novelty can negatively impact students' ability to reflect on what data to collect due to distraction of the environment and potential cognitive overload (Balling & Falk, 1980). Moreover, reflection and planning are tasks students struggle with in general and need scaffolding to assist them in these tasks (Quintana et al., 2004), much less in stimulating foreign environments. Given that students have limited time to collect data in this environment and may be unable to return, they need to be supported in collecting appropriate artifacts that could be used later or their ability to answer their inquiry question could be compromised.

After deciding on what data to collect, the student must choose an appropriate method to collect the data, such as a text, photograph, audio recording, or numerical probe data (temperature, pH, etc). Following collecting the data, students must describe and annotate the data for later use.

Describing and Annotating Data

Even after determining what data to collect the learner must reflect on the utility and purpose of the data and how it answers their inquiry question. This requires the student to consider not only why they collected the data, but also how to annotate the data in a useful manner. Students must connect the artifacts and phenomena they are observing with their existing scientific knowledge in order to construct factual and useful annotations, a task that is difficult for students (Metz, 2000; Roschelle, 1995).

When learners are unsuccessful at this task they may be unable to find and interpret data they collect later (Vavoula et al., 2009). In the case of Myartspace, students had difficulty managing and inspecting their collection of 40 different multimedia data objects due to an inability to locate the objects they desired. Having issues organizing multimedia data objects can be a problem as photographs and audio notes cannot be searched without accompanying textual annotations or other organizational means of categorizing the data thus requiring either the students or system to manage annotating and organizing the data.

The challenge of describing and annotating data is amplified in a novel environment as they are already over-stimulated (Balling & Falk, 1980) and must then deal with non-germane tasks that further distract them, such as data entry. Research has shown that users have difficulty typing on a mobile device (Tossell, Kortum, Shepard, Rahmati, & Zhong, 2010) and this could frustrate students and cause them to rush, avoid textual data entry, or detract from their ability to focus on the inquiry task. Previous work that had students collecting data from outside the classroom in a group with one handheld device for data entry found that the student in charge of data entry would get so focused on data collection that they would spend little time trying to make sense of the data they were gathering, compared to students who did not have to focus on data entry (Songer, 2006).

2.1.3 Using Data

Using data incorporates the NGSS practice of *analyzing and interpreting data*. This practice involves analyzing the data to derive meaning and identify patterns toward the goal of constructing a scientific explanation. This study focused on how students use data to complete two tasks, *searching through and exploring data* and *assessing data as evidence in* *an explanation*, and discusses the challenges students face in using large amounts of peer collected multimedia data.

Searching Through and Exploring Data

When a student is only interacting with a limited amount of data, they can spend time analyzing each piece of data in turn. However, difficulties with searching and assessing data compound as the data set becomes larger.

Research looking at how students search for information found search in general to be a challenging process for developing learners (Bilal, 2002; Druin et al., 2009; Hutchinson, Druin, & Bederson, 2007; Jochmann-Mannak, Huibers, & Sanders, 2008). Students may be unaware of appropriate search term syntax and have difficulty constructing complex queries, particularly when students are searching for information to answer open-ended question (Bilal, 2002; Bilal, 2001). Even if the students overcome the issues with search syntax, they still can struggle with typing and spelling the words they are searching for (Borgman, Hirsh, Walter, & Galagher, 1995; Solomon, 1993).

In dealing with a larger data set that exceeds a size that prevents easy browsing of all the data, these issues related to searching will have to be addressed in some manner for the data to be usable.

Assessing Data as Evidence in an Explanation

Beyond searching through the data, students need to assess the potential data items they encounter to find data to use as evidence in constructing a scientific explanation. The end goal of this activity is for students to use appropriate and sufficient evidence to support their claim (McNeill & Krajcik, 2008). Appropriate data needs to be relevant to the investigation question they are trying to answer and help make and support their claim. Sufficient data involves using enough data to convince others that your claim is valid, often in the form of providing multiple pieces of evidence that support the claim.

However, studies have found students struggle with using appropriate and sufficient evidence to support their claims (Berland & McNeill, 2009; McNeill, Lizotte, Krajcik, & Marx, 2006). They can struggle with what evidence is appropriate to back up their claim and often fail to apply enough data to strongly support their claim, in many cases relying on a single piece of evidence.

An additional challenge arises when dealing with large amounts of data collected by their peers, which may be factually inaccurate or not useful for the investigation (Berland & McNeill, 2009). In a simpler inquiry context students would have a small, curated data set;

the data set is ensured to be factually accurate and all of the data is appropriate for use in their explanation. However, the more inaccuracies and inappropriate data in the set, the more difficult and time consuming filtering, assessing and then disregarding inappropriate data becomes.

2.1.4 Summary of Challenges

A summary of the challenges students must overcome to collect multimedia data outside the classroom for later sharing with their peers, and then use data from this large, aggregate data set to construct scientific explanations is in Table 2-1. These challenges must be addressed with appropriately designed software scaffolds in order for students to complete an inquiry activity in this context.

	Cognitive Task	Challenges facing learners
Collect Data	Students must determine what data is needed to answer their investigation	Students have difficulty planning and monitoring their inquiry plans (Krajcik & Blumenfeld, 2006). They struggle determining what data to collect to answer complex inquiry questions (Berland & McNeill, 2009; Griffin, 1998).
	Students must describe and organize the collected data for later retrieval by themselves and their peers.	Students can be cognitively overwhelmed outside the classroom; non-germane but difficult tasks such as the mechanics of data entry can hinder their ability to collect data (Tossell et al., 2010).
		Students may have difficulty connecting their past knowledge to the artifacts and phenomena they observe (Roschelle, 1995).
		Students may not understand or reflect on how data needs to be annotated to be useful later (Metz, 2000). Students may be unable find data later if not properly organized (Vavoula et al., 2009).
Use Data	Students need to search through and explore the data set	Students can have organizational issues dealing with large amounts of multimedia data (Vavoula et al., 2009). Students have difficulty searching for information (Druin et al., 2009) and have difficult constructing appropriate search syntax (Bilal, 2001).
	Students need to assess data they and their peers collected to judge whether it is accurate and useful for their investigation	Children struggle with using appropriate and sufficient evidence (McNeill & Krajcik, 2011). Students have the greatest difficulty dealing to deal with large, student-collected data sets as this data can be inaccurate and/or not be useful for constructing a scientific explanation (Berland & McNeill, 2009).

Table 2-1. Summary of the challenges students face collecting multimedia data outside the classroom for later sharing with their peers, and then using data from this large, aggregate data set.

2.2 Related Science Inquiry Systems

Many systems have used software scaffolds developed to support students engaging in science inquiry. Software scaffolds are supports built into the software that assist the learner in accomplishing a task they otherwise would be unable to do (Quintana et al., 2003). An example of this could be a virtual tutor prompting the student through a series of questions to decompose a complex task.

The use of software to scaffold science inquiry is an area of ongoing research. Guzdial was one of the first researchers to discuss scaffolding software (Guzdial, 1993). Since then, there has been a variety of work on developing scaffolding for learners, and a sizeable amount of literature focused on scaffolding the various aspects of the science inquiry process (Cho & Jonassen, 2002; Kali & Linn, 2008; Linn, Clark, & Slotta, 2003; Luchini, Quintana, & Soloway, 2004; Quintana, Krajcik, & Soloway, 2002).

This section describes how different systems employ scaffolds to complete inquiry activities, organized by the tasks the systems sought to enable and support. Various systems have supported different aspects of the tasks around collecting and using data. This section reviews scaffolds used in desktop inquiry systems for use in the classroom, in pre-mobile out of the classroom data collection systems, and finishes with a look at more recent research into using mobile devices to collect data from outside the classroom.

2.2.1 Desktop Science Inquiry Systems

There is a volume of research on supporting students who are using data to construct explanations in the classroom with web-based or curated data sets (Bell, 1997; Linn et al., 2003; Quintana, 2001; Quintana et al., 2004; Sandoval & Reiser, 2004; Suthers et al., 1997). Of this research, three systems in particular have relevance to the Zydeco tools as they focused on using web or curated data to construct explanations: Belvedere, SenseMaker, and Explanation Constructor.

Belvedere

Belvedere provides students with shared workspaces for coordinating and recording their collaboration in scientific inquiry as they construct an argument map (Suthers et al., 1997). The system acknowledges learner issues including lack of motivation, limited knowledge of science domains, inability to recognize abstract relationships implicit in theories and arguments about them, and difficulty keeping track of complex debates and lack of scientific argumentation criteria (Suthers & Jones, 1997). Its progressively simplistic scaffolding includes a collaborative workspace where learners construct argument maps relating data and hypotheses, belief level settings for these relations, an AI coach that provides advice, a "chat" facility, facilities for use with Web browsers, and peer critique supported with rubrics (Suthers et al., 1997).

SenseMaker

The SenseMaker system utilizes several similar scaffolds toward its goal, allowing small groups of students to organize and annotate a collection of Web-based evidence that can then be shared with others, including Web browser support, rating the evidence based on usefulness, and providing expert modeling through an AI coach (Bell, 1997). Additionally, SenseMaker uses nesting claim frames on its workspace to allow learners to organize categories for their evidence, visible evidence dots to represent how useful the data is, and an Activity Listing with a checklist of science inquiry processes.

ExplanationConstructor

Taking a different approach from the Belvedere and SenseMaker, ExplanationConstructor, designed to support students' construction and evaluation of explanations through their inquiry (Sandoval & Reiser, 2004), is a highly specialized system fitted to extremely specific content, supporting only three driving questions. Their explanation-driven inquiry approach yielded different scaffolds from what previous systems had utilized, including multiple explanation guides to choose from providing a means to monitor progress, the ability to enter sub-questions to the larger question, content specific evidence and data charts, and self assessments at the end of the project. Even with these differences, ExplanationConstructor does utilize some similar scaffolds; students are able to set a confidence measure on their claims, have an organizer to hold questions and explanations, and provide peer critique based on a rubric.

2.2.2 Worksheets for Collecting Data Outside the Classroom

Before the advances of mobile technologies, researchers scaffolded students engaging in out of classroom inquiry through other means. Kids Network and the Globe Project had particular relevance to the Zydeco system, and are reviewed here.

Kids Network

One of the initial systems for students to share scientific data with their peers was Kids Network, started in 1986 (Bradsher & Hagan, 1995). Kids Network provided several potential trials that schools could participate in, such as determining the pH of rainwater. Each school would go out and capture data from the field and upload the data to each other at specific intervals for analysis and discussion at other sites. Since this research was before the World Wide Web, the students telecommunicated the results to each other and wrote letters to teammates in other schools.

Kids Network enabled students to have a richer science experience at the time, though it required a tech savvy instructor at the school to push for its usage (Bradsher & Hagan, 1995). Because of the high learning cost and since the system was structured around schools conducting experiments at the same time, some schools found it unwieldy and they abandoned usage of the system.

Globe Project

Another initial system in data sharing was the Globe Project, with the goal of students sharing data around the globe (Means, 1998). Started in 1995, the Globe Project was a website where schools could upload data in various formats. The data collection process involves teachers printing out a worksheet with specific data the students need to collect. The students use probes or observe the data in the field, and then enter this data into web forms. The website supported some data lookup; students could examine contour maps that displayed overlays of different numerical data collected around the world and how it changes over time. There was also functionality to enable students to communicate with scientists and other students around the world.

The Globe Project was successful at enhancing students' environmental awareness and scientific understanding of the earth (Finarelli, 1998). This project also identified guidelines for this data collection and sharing to work between schools. These ensure each school is accurately and consistently taking measurements, and persists in taking these measurements over time.

While the Globe Project is still in existence, it has not been updated to utilize mobile devices for any of their data collection and still has a similar worksheet structure for data collection and subsequent web entry.

2.2.3 Mobile Data Collection Systems

With the advent of mobile technologies new ranges of activities became available to learners. Many systems have been made to support students in data collection outside the classroom using mobile devices. They have focused on students with a variety of different activity structures, types of data collected, and mobile supports.

BioKIDS

BioKIDS is a year-long curriculum with an accompanying mobile system that is designed to enable students to learn about biodiversity. One such activity is for students to identify and track populations of organisms for later analysis (Songer, 2006). Using mobile devices, small groups of students (ages 10-12) would record the types of various organisms they find around their schoolyard with the help of a facilitator. The system featured prompts to assist students through the collection process, and also enabled students to review their current and previously collected data *in situ*. Students constructed explanations with this data, though this was structured through standard classroom supports (lectures, class discussion, and workbooks) and the mobile device was primarily used for data collection and was only partially utilized as an analysis and sense making support (Songer, 2006). In some of the curricula units, the number of sightings of organisms by each student group was summed together into a table, enabling them to study the schoolyard ecosystem.

Ambient Wood

Ambient Wood was an exploratory project that had pairs of students (ages 11-12) use mobile devices to guide inquiry in an outdoor environment in order to explore the woodlands (Rogers et al., 2004). The research was studying how an outdoor environment could be digitally augmented, such as by adding digital sounds and having students receive messages and be prompted based on where they are in the environment. This digital augmentation had both student and environment initiated methods of experiencing this digital augmentation. Students could take probe measurements of the moisture readings from the soil and also could utilize a periscope through which they could watch a pre-recorded video clip about the environment. The system utilized location-aware prompts through coming in contact with Bluetooth beacons to provide advice on potential areas to investigate (such as the nearby blackberry bushes). The Bluetooth beacons were also utilized to provide just-in-time information when students approached pertinent areas, either through sending images or voice-overs to the students to investigate areas.

Initial pilot work had found that students were easily overwhelmed by the stimulus coming at them, both from the environment and just-in-time prompts (Rogers et al., 2004); because of this, the software was modified to require students to have to initiate the playback of the just-in-time notification. A facilitator was with the students as they traveled through the outdoors, and students had a walkie-talkie to a remote facilitator to whom they reported the data they found and the remote facilitator would discuss the data and suggest new possible tests (Rogers et al., 2004).

LET's GO

The LET's GO project involved students (ages 16-18) comparing the water quality at a creek versus a pond (Maldonado & Pea, 2010). Students worked in small groups and each of them were given a role: one student acquires water samples, another uses probes to take measurements, while a third records activity with LiveScribe pen and notebook, and the final student photographs the location the sample was taken. The groups go to predetermined GPS coordinates along with a facilitator. The facilitator assists the students throughout the activity and prompts them with leading questions to make hypotheses about the water characteristics.

The students collect the pH, temperature, dissolved oxygen, and conductivity of the water and later record the data to Google Earth with the GPS measurement (and attach the photo of the collection site). Reviewing the data enabled students to improve their understanding of the differences between the two sites. In their initial trials the data was only examined within their groups, though future testing seeks to expand this to multiple classes sharing the data (Maldonado & Pea, 2010).

Participate Project

The Participate project involved data collected by and shared amongst students (ages 13-15) in order to inform environmental debate (Woodgate et al., 2010). The students collected numerical sensor readings from the field (CO, sound, and temperature data) and had a GPS device and mobile phone working in sync to record the location and to timestamp each sensor measurement. Back in the classroom students used Google Earth and graphing software on a laptop in order to review the sensor readings by location and time.

The researchers found this student generated data was very engaging back in the classroom and provoked a lot of discussion (Woodgate et al., 2010). The study noted that students were similarly engaged when examining the data trails plotted by Google Earth and the low-tech line graphs the students created, suggesting that the students do not have a strong preference towards fancier visualization methods.

SENSE

Fraser's SENSE project had students (ages 10-14) collect environmental data from the field to learn about pollution (Fraser et al., 2005). Students went outdoors and tried to find areas of higher pollution using CO_2 and wind speed sensors attached to a PDA. While a student was taking probe readings, another student video-taped the collection process. This data was synchronized, enabling review of the video of the collection process along with the numerical probe readings on a graph. Students could create textual time-indexed annotations with their data analysis tool, enabling them to mark interesting occurrences (such as a semi-truck driving by).

Two classrooms in separate schools were part of this study and they were able to see each other's data (Fraser et al., 2005). Students found the process of viewing another classes' data helped students reflect and understand new perspectives on the data. The video synchronized with the numbers enabled students not involved with the collection process to understand how it was done, as each class used different techniques for gathering the data.

Adaptation of "Flyer"

Research done by Laru et al. sought to explore how students' inquiry can be supported with multiple scaffolding agents (facilitator and software scaffolding) during a field trip (Laru et al., 2010). Students (age 12) went on a biology field trip with a tutor and nature guide who helped scaffold the students in having an argument discussion. This study adapted a mobile peer-to-peer messaging tool known as "Flyer" to be used as a meta-cognitive and procedure support. The system utilized a framework around claims, grounds, and warrants for the argumentation process and provided sentence starters for each of these aspects.

The research found the combination of the dynamic scaffolding provided by the facilitator on the trip helped engage students in discourse when coupled with the software scaffolds (Laru et al., 2010). The research focused on differences between the high and low-scoring students from their pre-test and found that students who scored high on the pre-test were having twice as much discussion around their scientific arguments.

OOKL (Myartspace)

OOKL (previously called Myartspace) is a handheld program that enabled students to record photos, audio or textual notes in a museum and collect "exhibits" by typing in a two digit code (Vavoula et al., 2009). The students were prompted to textually write the reason why they collected each piece of data in the museum, but had no other form of annotation on the data. After they explored the museum, the students could review the data they and their peers collected on a website. The students then had to create a PowerPoint style presentation on what they learned, and the website enabled them to order pieces of data to construct their presentation.

While the system is now a commercial product, it was initially piloted with a class of students (aged 13-14). The students were able to review not only the data they collected, but also their peers. However, the software did not link the textual, audio, or photographic data together and students had difficulty finding and interpreting the data. While students enjoyed the activity, the lack of identifying metadata caused them to take a great deal of time reviewing data—more than was available for the activity.

2.2.4 Summary of Science Inquiry Systems

This section described a variety of science inquiry systems. These systems have explored a number of directions of science inquiry research, including: early systems designed around sharing numerical data between schools that was collected via worksheets; various styles of data collection outside the classroom, though many of these systems did not allow open-ended data collection that allowed for student-choice, instead relying upon a facilitator directing their inquiry, and; in class inquiry systems that utilized web-based or curated data sets to construct explanations.

However, none of these systems have supported the combination of:

- Scaffolding students engaging in multimedia data collection outside the classroom for later use by themselves and their peers, and
- Scaffolding students in using large amounts (several hundred pieces) of personal and peer-collected multimedia data to construct scientific explanations.

An overview of all the systems can be seen in Table 2-2, including where Zydeco fits into the literature. While none of the systems supported the same context, some of the scaffolding approaches other systems have developed can apply towards the scaffolds used in the Zydeco tools, which are discussed in the design of the Zydeco tools in chapter 3.

General Characteristics				Activities Supported		Processes Within the Activities that Require Supports and Scaffolds					
System	Platform	Project start date*	Participant age range**	Collecting data outside the classroom	Using data to construct scientific explanations	No expert needed to scaffold data colletion	Used collected data for post-activities	Annotate data outside the classroom	Used peers shared data	Primarily multimedia data	Using large (hundreds of pieces) multimedia data sets
Zydeco	Handheld, Tablet	2010	11 to 13	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adaptation of "Flyer" (Laru et al., 2010)	Handheld	2010	12	Yes	Yes	No	No	No	No	Yes	No
BioKids (Songer, 2006)	Handheld	2006	10 to 12	Yes	Yes	No	Yes	Yes	Yes	No	No
Ambient Wood (Rogers et al., 2004)	Handheld	2004	11 to 12	Yes	Yes	No	No	No	No	No	No
Participate (Woodgate et al., 2010)	Handheld, Desktop Website	2010	13 to 15	Yes	Yes	Yes	Yes	No	Yes	No	No
SENSE (Fraser et al., 2005)	Handheld, Desktop Website	2005	10 to 14	Yes	Yes	Yes	Yes	No	Yes	Mix	No
LET'S GO (Maldonado & Pea, 2010)	Handheld, Desktop Website	2010	16 to 18	Yes	Yes	No	Yes	Yes	No	Yes	No
OOKL (previously Myartspace) (Vavoula et al., 2009)	Handheld, Desktop Website	2007	13 to 14	Yes	No	Yes	Yes	No	No	Yes	No
Globe Project (Means, 1998)	Worksheet, Desktop Website	1995	6 to 18	Yes	Yes	Varies	Yes	Yes	Yes	No	No
Kid's Network (Bradsher & Hagan, 1995)	Worksheet, Desktop Website	1986	8 to 16	Yes	Yes	Varies	Yes	Yes	Yes	No	No
Explanation Constructor (Sandoval & Reiser, 2004)	Desktop	2004	13 to 14	No	Yes	-	-	-	-	Mix	No
SenseMaker (Bell, 1997)	Desktop	1997	10 to 14	No	Yes	-	-	-	-	Mix	No
Belvedere (Suthers et al., 1997)	Desktop	1997	12 to 15	No	Yes	-	-	-	-	Mix	No

* This corresponds to first publication if no start date is given, projects may have begun earlier. ** As can best be determined. May be more varied.

Table 2-2. Overview of existing science inquiry systems with a comparison of the activities and processes used within the activity these systems support compared to Zydeco.

2.2.5 Scaffolding Methods and Guidelines from Related Systems

Methods of scaffolding

All of the inquiry systems discussed above implement distributed scaffolding. Distributed scaffolding is a collection of scaffolds to make a task manageable for learners (instead of relying on only one scaffold) (Puntambekar & Kolodner, 2005; Puntambekar & Kolodner, 1998).

However, there are various forms the scaffolds can take and methods they use to support learners through challenges. In looking at distributed scaffolding in more detail, Tabak did a literature review classifying how the distributed scaffolds work to address the needs of learners in overcoming the challenges of the task by placing scaffolds into three categories (Tabak, 2004):

- Differentiated scaffolds: using different scaffolds to address different learner needs.
- Redundant scaffolds: using multiple scaffolds to address the same learner needs.
- Synergistic scaffolds: using multiple scaffolds that work together to address the same learner need in a manner that is more effective than the sum of all the scaffolds effectiveness if used independently of each other (thus having synergy in their effectiveness).

While nearly all systems I described use a combination of differentiated and redundant scaffolding, there has been far less work on synergistic scaffolds. Tabak did some initial work looking at synergistic scaffolding, where she had the scaffolding from the teacher and the software system work together in a manner that was more effective than either independently (Tabak, 2004).

In Tabak's research, she implemented synergistic scaffolding by having the teachers discuss the curriculum and hold class discussions using the same language the software used to create cohesion among the scaffolding systems (Tabak, 2004). It was noted that by having this language reflected in both systems, students were able to become familiar with the scientific vocabulary and process used, which Tabak postulated benefitted their overall ability to perform their inquiry project.

Scaffolding Guidelines

From testing the various systems, researchers have developed scaffolding guidelines for other designers on how to more effectively design systems to scaffold learners (Kali & Linn, 2007; Quintana et al., 2004). An example of one such scaffolding guideline is "*Visibility*: Scaffolds should be very visible because learners will not usually trigger scaffolds" (Quintana et al., 2002).

These guidelines are developed by testing how learners interact with the scaffolds in a system and synthesized into a conceptual set of guidelines By looking at if scaffolding guidelines are able to be used to implement scaffolds across multiple systems researchers can validate their effectiveness (Basu, Sengupta, & Biswas, Under review; Kali & Linn, 2007; Luchini et al., 2004; Quintana et al., 2002; Quintana et al., 2004).

The guidelines can be used to reduce the development time of similar scaffolding systems and help the system be more effective, though the guidelines only provide conceptual advice on scaffolding learners; it is still up to the designer to implement the physical scaffold. However, as these guidelines were derived from previous systems that have tested various scaffolds and described the scaffold implementations, these designs can help by serving as examples for other designers to reuse or get a sense of how the guideline could be implemented.

The scaffolding guidelines researchers create vary greatly in scope. Guidelines may encompass suggestions for all forms of scaffolding (in any field), specific to a general set of tasks such as science inquiry, or for specific contexts (e.g. scaffolding guidelines for supporting inquiry in agent-based simulation environments) (Basu et al., Under review; Kali & Linn, 2007; Quintana et al., 2004).

Some of the guidelines in the literature can provide guidance when designing the Zydeco learning environment. These guidelines are discussed here and related to the design of the Zydeco system in chapter 3.

Some overall scaffolding guidelines that are relevant are the five guidelines Quintana developed, which are (Quintana et al., 2002):

- *Visibility*: Scaffolds should be highly visible because learners will not generally trigger scaffolds.
- *Essentialness*: More scaffolds should be essential than optional, which learners tend to bypass.
- *Coupling*: Scaffolds appearing together should be tightly coupled to focus learner attention. Multiple, loosely coupled scaffolds should be made more essential so learners will use them.

- *Usability*: Scaffolds need to be usable, but they should not make work task too automatic or learners will not mindfully perform the underlying task.
- *Representation*: Neither textual nor graphical scaffolds are inherently better or worse than the other.

The scaffolding guideline relating to usability is of particular note. With all scaffolding, it is necessary to strike a balance between making the software easy to use, but ensuring learners are mindfully performing the task if the task is an essential part of the learning process. "Mindful" means that the scaffold supports the learner in doing the task, while not making the activity too easy to perform or else learners may not learn anything from it (Quintana et al., 2003).

This idea of encouraging mindful behavior can be opposite of more traditional user-centered design, which seek to make the task as efficient as possible for the user (John & Kieras, 1996). The balance and trade-off between the efficiency of interface supports and mindful work is an ongoing area that researchers are exploring, which follows a design strategy called learner-centered design (Luchini et al., 2004; Sedig, Klawe, & Westrom, 2001). Learner-centered design seeks to design software to encourage learning, with usability being a focus but the learning of the student being the highest goal, leading to tradeoffs on encouraging mindfulness versus usability.

In the specific context of science inquiry, there are several researchers who constructed encompassing sets of guidelines to scaffold science inquiry.

Kali and Linn created a design principles database for educational software in the form of a website that anyone can browse and add to (Kali & Linn, 2007). This has been developed since 2001 and has been contributed to by a number of educational designers, initially from conference workshops and courses. The database was structured into a three level hierarchy. At the top were four meta-principles, which are: make thinking visible, make science accessible, help learners learn from each other, and promote autonomous lifelong learning. Under each of these meta-principles are a variety of pragmatic principles, which provide more specific guidance for implementation. Each pragmatic principle then has one or more specific principles or feature rationales, where the interface element that was developed for a specific system or research trial following a pragmatic guideline is described in detail. Their work contributed a multitude of guidelines to scaffold these activities, though many lacked empirical analysis.

Several researchers worked together to create a set of scaffolding guidelines for addressing the high-level challenges students struggle with during science inquiry (Quintana et al., 2004). These guidelines are broken down by the guidelines to address the challenges of sense making, process management, and reflection and articulation. These guidelines were based on a review of desktop science inquiry tools and literature on scaffolding strategies. Their guidelines are organized into high-level scaffolding guidelines that address a challenge of inquiry and have more specific guidelines to implement the scaffolds that are referred to as scaffolding strategies (similar to Kael and Linn's meta-principles and pragmatic principles). A breakdown of the scaffolding guidelines and strategies developed by Quintana et al. are:

To address the challenges of sense making:

1. Use representations and languages that bridge learners' understanding
    a. Provide visual conceptual organizers to give access to functionality
    b. Use descriptions of complex concepts that build on learners' intuitive ideas
    c. Embed expert guidance to help learners use and apply science content
2. Organize tools and artifacts around the semantics of the discipline
    a. Make disciplinary strategies explicit in learners' interactions with the tool
    b. Make disciplinary strategies explicit in the artifacts learners create
3. Use representations that learners can inspect in different ways to reveal important properties of underlying data
    a. Provide representations that can be expected to reveal underlying properties of data
    b. Enable learners to inspect multiple views of the same object or data
    c. Give learners "malleable representations" that allow them to directly manipulate representations

To address the challenges of process management:

4. Provide structure for complex tasks and functionality
    a. Restrict a complex task by setting useful boundaries for learners
    b. Describe complex tasks by using ordered and unordered task decompositions
    c. Constrain the space of activities by using functional modes
5. Embed expert guidance about scientific practices
    a. Embed expert guidance to clarify characteristics of scientific practices

    b. Embed expert guidance to indicate the rationales for scientific practices
6. Automatically handle nonsalient, routine tasks
    a. Automate nonsalient portions of tasks to reduce cognitive demands
    b. Facilitate the organization of work products
    c. Facilitate navigation among tools and activities

To address the challenges of reflection and annotation:

7. Facilitate ongoing articulation and reflection during the investigation
    a. Provide reminders and guidance to facilitate productive planning
    b. Provide reminders and guidance to facilitate productive monitoring
    c. Provide reminders and guidance to facilitate articulation during sense making
    d. Highlight epistemic features of scientific practices and products

These guidelines are relevant to this research as they are a comprehensive set of guidelines for scaffolding the standard challenges that arise in science inquiry. However, these guidelines were all developed for desktop software used in the classroom. This research seeks to support students in engaging in inquiry using mobile systems outside of the classroom and continuing the inquiry back in the classroom, which can present additional challenges due to the device and context and may require developing new guidelines or adapting existing guidelines to address the challenge.

When designing mobile software, the different context and form factor of the device presents new opportunities and limitations. These opportunities arise in part out of the ease in which the mobile device can be carried and used in a variety of contexts due to its portability. However, the small size of the device limits the amount of information that can be displayed on screen and can change the methods for data entry than would be employed on a desktop machine. Issues and affordances that arise from mobile devices are:

- *Usable in any context*: Mobile devices can be carried and utilized in a variety of contexts. In an educational context, mobile software may be used outside of the classroom and have reduced or no support from the teacher, requiring additional software scaffolds to compensate.
- *Small displays*: Mobile displays are significantly smaller than their laptop or desktop counterparts. These smaller displays enable less information to be displayed and require designers to reduce clutter and only display essential interface elements on screen (Mohageg & Wagner, 2000).

- *Data entry*: Most touch-screen smartphones are equipped with virtual keyboards. Typing on virtual keyboards can be more difficult and prone to errors than a desktop keyboard, both due to the smaller spacing between keys and the lack of tactile feedback. However, mobile devices are equipped with a variety of sensors that can be utilized to collect data and be utilized as an input mechanism, such as audio recorders, cameras, GPS, accelerometers, magnetometers, and gyroscopes.

User-centered design research has developed several guidelines to overcome some of the challenges in developing software for mobile devices, which include:

- Minimize number of elements onscreen (Bergman, 2000; Weiss, 2002).
- Automatically complete work for the user whenever possible (Passani, 2002) (Weiss, 2002).
- Provide word selection instead of requiring text input (Gong & Tarasewich, 2004).

Educational researchers have also begun to explore how scaffolding guidelines can change in a mobile context. Though her work was focused on students in a classroom, Luchini developed three scaffolding guidelines for designing educational software on mobile devices, (Luchini et al., 2004). These are:

- Evaluate trade-offs to select scaffolds: Scaffolds need to be evaluated based on how usable they are versus their effectiveness at supporting students; a scaffold that provides substantial cognitive support but is very difficult to use may not be worth including, while a scaffold that has provides less support but does not make the system harder to use would be better to include.
- Design “double-duty” interface elements: Due to the limited space on mobile screens, develop interface elements that provide two or more forms of support for students (such as a button providing a visual cue with its text and an additional support when activated).
- Streamline scaffold implementation: Make scaffolds as usable as possible to encourage students’ use.

Though these guidelines discussed provide conceptual recommendations and advice for designing software on mobile devices, they need to be applied towards scaffolding the challenges that students will face in the context. As researchers have not explored how to support students engaging in inquiry as they collect multimedia data with handhelds outside the classroom and then use their large, personal and peer-collected data to construct scientific explanations, this research must determine how to apply these guidelines in this context and potentially extend or develop new scaffolding strategies to expand the literature.

2.3 Methods to Scaffold Multimedia Data Collection and Use

The main challenges this activity seeks to support is annotating multimedia data outside of the classroom and organizing this data for later usage. A traditional method of organizing data that has been used in computer file systems is to give objects a file name or otherwise title them. However, this can only convey a limited amount of information so a title alone may not be enough to enable students to understand the purpose of multimedia data in order to use it, as well as make it difficult to organize large quantities of information.

Data has often been organized into folders (by a title), though folders have the drawback of having a rigidly defined hierarchy. With a folder, an object can only reside in one place at a time, enabling objects to only be classified into one category, which can hinder later retrieval if not well placed.

One method that has been used to support the collection and later usage of multimedia data is tags (MacGregor & McCulloch, 2006). Tags are short textual labels used to annotate, organize, and describe data. The benefit of tagging is that it removes the rigid nature of having a few defined categories or hierarchies that data must be classified under and enables users to create their own categories and descriptions on the data.

However, little research has studied how young learners (ages 11-13) use tagging systems in general, much less in an inquiry context. To inform the design of the searching mechanisms, relevant literature on how students search for information on the web or within collections is brought to inform the strategies this study uses.

Beyond tagging, another form of annotation that can be used to describe data is a voice note. A voice note enables a user to record their thoughts on a piece of data and enumerate reasons to support using the data or ask themselves questions to think about concerning the data.

The literature is explored on these different research strands to inform the design of appropriate scaffolding around applying and using annotated data.

2.3.1 Tagging Systems

Tags have been used on a number of social media sharing sites, such as Flickr, Del.ic.ious, and Last.fm as a means for users to describe and organize content (Bischoff, Firan, Nejdl, & Paiu, 2008). Tags are short textual labels used to annotate, organize, and describe data (Ames & Naaman, 2007). These systems enable users to upload their own content and provide textual annotations by the creator and / or other users on how to describe data, which users can search upon.

Flickr is an example of a tagging and annotation system for photographic data (Yahoo inc.). Users upload photos and tag the photo to describe it, while other users can search upon the tags to find similarly tagged photos. In Flickr, tags are limited to a single word, though users get around this by omitting essential syntax or punctuation (e.g. squareformat or sanfrancisco). Additionally, users can add a textual description and title to their photograph to further label it, and organize it within groups.

Tagging systems also use different ways in which the data in the system is tagged. Thomas Vander Wal is credited with coining the term folksonomy to describe the tag language of collaboratively creating tags, and introducing the conceptual distinction of broad and narrow folksonomy (Wal, 2005, 2007). Folksonomy is defined as the act of tagging by the person consuming the information (Wal, 2007). In a broad folksonomy many users tag one object (e.g. del.icio.us), while in narrow folksonomy only one or a few people provide tags, and often it is the content creator adding the tags (e.g. Flickr) (Wal, 2005)

Flickr is an example of a narrow folksonomy, as the user uploading the photo is typically the only person to tag it (Yahoo inc.). Research on the motivation behind tagging photos for use in social media sharing sites such as Flickr and ZoneTags, two sites geared around posting and sharing photographs, found that tags are applied for a cross of personal/ social and organizational/ communicative functions (Table 2-3) (Ames & Naaman, 2007).

		Function	
		Organization	**Communication**
Sociality	**Self**	* Retrieval, Directory * Search	* Context for self * Memory
	Social	* Contribution, attention * Ad hoc photo pooling	* Content descriptors * Social Signaling

Table 2-3. Motivations for users tagging photos in Flickr/ZoneTags (Ames & Naaman, 2007).

Using Tags for Information Retrieval

Tags can provide a means to retrieve information by pulling up all the data that is annotated with the same tag. This can provide a mechanism for performing textual searches upon data, which may not otherwise have identifying characteristics that can be used in textual searches, such as images.

Tagging systems have been found to be usable both for search and knowledge discovery, though research has been divergent on their effectiveness (Bischoff et al., 2008; MacGregor & McCulloch, 2006). Previous research has compared the effectiveness that novices and experts had using keyword searches and exploratory tag searching to find relevant information when given a task (Kang, Fu, & Kannampallil, 2010). Novices (which would be representative of students) were more effective at using exploratory tag searches, while experts were better off with keyword searches.

Some issues with using tags for information retrieval arise due to the fact that the users of the system add tags, which can mean that a number of tags are ambiguous, overly personalized to the creator, and inexact in their description. Research has found that how tags are utilized to describe objects and their usage in searching for objects is highly context dependent (Bischoff et al., 2008). For example, a study found that users on Flickr who use tagging for their personal photographs prefer tags related to the time an event occurred, while the topic of the photograph is more important to tag in a public photo (Bischoff et al., 2008). However, these varied tag vocabularies between users hinders the tags usefulness for search (MacGregor & McCulloch, 2006).

For example, in a small-scale study to assess tag literacy of users on del.icio.us and Flickr, it was discovered that 28% and 40% respectively of the tags were deemed erroneous (Guy & Tonkin, 2006). Erroneous was deemed as either misspelled, from a language not included in multilingual dictionary software they used, in a form the dictionary could not understand, or a compound word consistent of more than two words or having a mixture of languages. Beyond this, an additional 11% of the del.icio.us tags and 8% of the Flickr tags were in plural form, which is not usable for searching upon items tagged in the singular form, reducing the overall tag effectiveness. This study suggests educating users to add "better" tags and improving the system to allow "better" tags are two ways to improve searching upon this data.

Issues with tagging can stem partially from there being little or no mechanics to control issues relating to synonyms (different word, same meaning) or homonyms (same word, different meaning). Clay Shirky has argued that there is no need to worry about synonyms, as users are choosing their particular tag names for the underlying meanings of the name (Shirky, 2005a, 2005b). An example of this is tagging information as “cinema” or “movie”, which Shirky argues the tagger has a unique meaning intended and information loss would occur if these tags were aggregated under one name. However, Shirky does not discuss an approach that would scale or the impact on general resource discovery that happens when there are a variety of synonyms of tags.

One interface element that is commonly seen in tagging systems to search the data is a tag cloud (Figure 2-2). A tag cloud displays a list of different tags found on the data set and may vary the font size, boldness, and text color of the tags to enable the user to distinguish between how frequently the tags occur (typically having bolder and larger text for frequently used tags) (Hearst & Rosner, 2008; Rivadeneira, Gruen, Muller, & Millen, 2007). However, Heart and Rosner argued that these tag clouds are more of a social signaling method than a data analysis tool (Hearst & Rosner, 2008). They claim that tag clouds can be an effective signaler over what the content of a website is (either for personal or social interaction), but suggests that the tag cloud is inferior to a more standard alphabetical listing of tags.

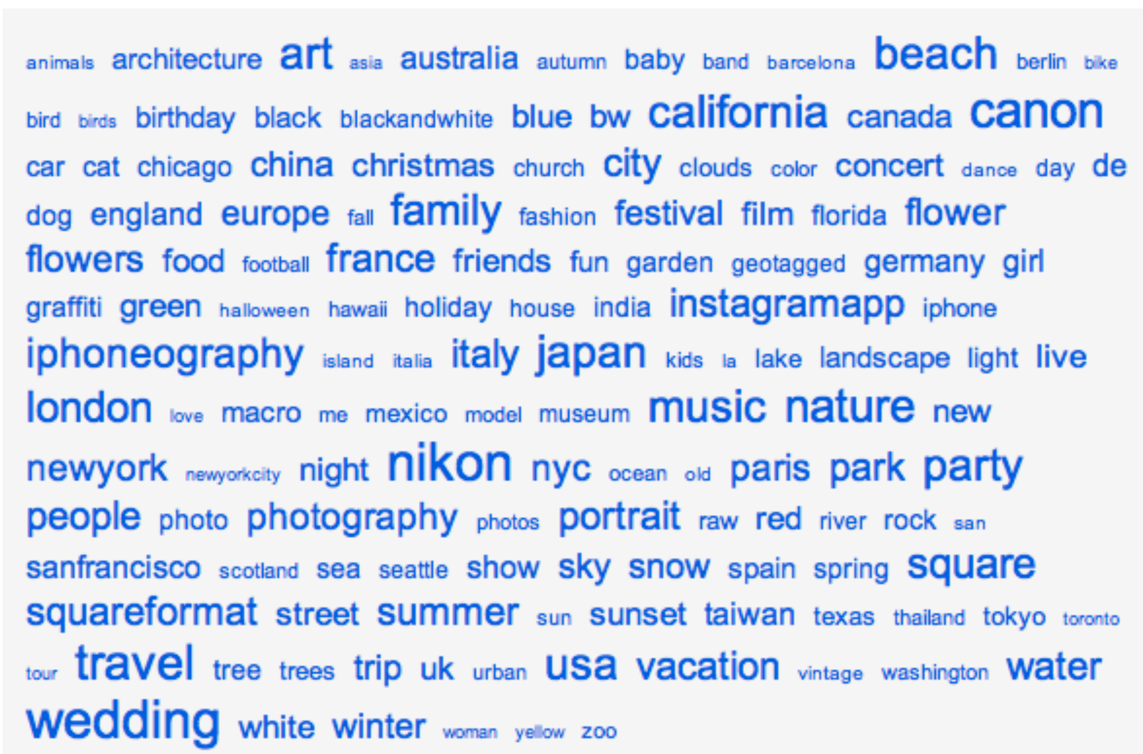

Figure 2-2. Tag cloud from Flickr.com displaying the "hot tags".

Tagging Systems in Education

There have been a few examples of tagging and annotation systems used in an educational context.

MobiTOP was a sytem to support university students in performing geology trials in the field. MobiTOP explored a different approach than the typical tagging sites by using hierarchical tagging to enable users to comment on photos by having a tag within a tag (Razikin et al., 2009). The tagging system received a mixed response from the college students who used it to collect data at geological sites while peers were in a lab suggesting tests. The complications of the tagging process and poor image quality made it difficult for students not involved with the collection process to understand the data.

Museums have used tagging as a way to empower visitors to interact socially around exhibits and to make sense of the content that they are exploring in the museum (Cosley et al., 2009). With MobiTags users could contribute tags to objects in an art museum and also vote on tags other users created to increase social awareness, social navigation, and engagement (Cosley et al., 2009). The visitors reported feeling more connected and engaged in the museum, though there were several cases where users voted down silly and unhelpful tags and suggested having a filtering system.

Highlights

The research I discussed on tagging systems was primarily carried out on adults in a social media or personal organization context. Tags can be a means to annotate and describe multimedia data, and can be used to find data and explore a data set, though its effectiveness appears mixed. While research on tagging systems was primarily done for adults outside of an educational context, several lessons can be learned from this work that are applicable to student usage in an educational context:

1. Create a common format for the tags students apply and encourage accurate, consistent (between students), and useful annotations in order to increase usefulness later in using data.
2. With data collection outside the classroom, it would make sense for students to utilize a narrow folksonomy with annotating the data. Only the student who collected the piece of data would have the initial defined sense of how they plan to use it (though back in the classroom other students may change their idea).
3. Tag clouds can provide a means of understanding the content in the data set, though alphabetically listing the tags may be more effective for using tags in retrieval tasks.

While there is a lack of tagging research to see how students in K-12 apply and use tags, an appropriately designed tagging system has the potential to support students in this context. However, due to the cognitive challenges students face in collecting data and later using it to construct explanations, an array of scaffolds and supports needs to be built around the tagging system to support them through these processes.

2.3.2 Voice Notes

There are also cases where audio recordings are used as a method of data entry for mobile devices. This has been used as a means to create notes while avoiding the challenges of typing on the mobile device.

Audio notes have been used to convert to speech. Dragon Dictation is an application that provides speech to text capabilities with high accuracy, advertised at being over 90% (Nuance communication). However, though quite accurate in solo usage by adults, the same accuracy may not be achieved with developing children's voices in a noisy environment.

Note taking tools such as the Livescribe Pen enable a user to record audio as they write with the pen, which links the audio to the particular note (Tsandilas & Mackay, 2010). These allow a traditional form of note-taking to be augmented with technology, providing for a richer form of data to review.

Several mobile educational systems have enabled users to record audio while collecting data outside the classroom. Myartspace allowed students to record audio notes in the museum as a means to record their thoughts on a particular exhibit (Vavoula et al., 2009). The SENSE project had students record videos of their data collection, where some of the students described how they were collecting the data or any conditions they noted (Fraser et al., 2005).

While voice notes can be effective at avoiding the challenges of data entry, they do not provide any organizational support for finding data later, as users are unable to search on the content within the notes. Another potential downside of using audio notes is that listening to an audio note can take longer than reading text, and also can increase the ambient noise level in a classroom and cause a disruption. Having students wear headphones could counteract the ambient noise, but that may have the effect of reducing collaboration between the group members by discouraging discussion.

Highlights

The previous research has shown that voice notes have the potential to be a useful means of annotating data. They are not a means of addressing organizational issues around information retrieval, but enabling students to record audio notes can address the challenges they may face in typing on a handheld keyboard.

2.3.3 Supporting Students Searching

In dealing with larger data sets, at some point it becomes infeasible to solely browse through the data. At this point, students will have to have mechanisms to search through and filter the data in order for the data set to be fully usable. The cognitive development of the student plays a large role in their success in searching for information, with even a two-year difference in age causing significant differences in ability (Druin, Foss, Hutchinson, Golub, & Hatley, 2010). Thus, tagging mechanisms, while they have been shown to support adults who are novices in a field (Kang et al., 2010), may fail when the novice is less cognitively developed.

To gain an understanding of how students could be supported to search for information, research on how they search for information on the web was examined. A variety of research has been done on understanding how students search for information on the web and supporting them in this process (Bilal, 2002; Druin et al., 2009; Hutchinson et al., 2007; Jochmann-Mannak et al., 2008).

Studies looking at how children search for information on the Internet have found them to have different strategies than adults; often children have chaotic search habits that employ a variety of strategies, though they typically prefer and are more successful at browsing for data instead of using keyword searches (Bilal, 2002; Hutchinson, Bederson, & Druin, 2006; Hutchinson et al., 2007). However, keyword searches have merit: Solomon found that children were most successful in keyword searches when using simple, concrete terms, which required little planning or else were provided to them (Solomon, 1993).

Difficulties with keyword search are partly due to the fact that typing and spelling are difficult for children (Borgman et al., 1995; Solomon, 1993). However, even if those issues are addressed, children still need sufficient knowledge to come up with a useful query, and even then the search engine may use different terminology and return no results (Abbas, Norris, & Soloway, 2002).

There have been a number of websites designed to support children's search behavior. These explored different interface and search strategies that students can use to find information. Relevant examples of interfaces used by students that overlapped with the age range of this study (between 11 and 13 years old) are discussed below.

Yahooligans was a web portal for Yahoo! that was specifically designed for children (aged 7-12) (Bilal, 2002). Bilal ran three different trials with children (aged 12-13) looking at how they approach doing fact-based, research-based, and defining their own search task (Bilal, 2002; Bilal, 2001). She found that students fared better with browsing strategies, as misspelling and search syntax issues hindered keyword searches. The students also had a greater interest and task completion when they were defining their own search task versus being assigned a task.

Hutchinson researched how students (ages 6-11) searched for books on the International Children's Digital Library (Hutchinson et al., 2006; Hutchinson et al., 2007). She explored different search interfaces to support both basic and advanced searches for books, trying to determine how to best support students with varying developmental abilities. The interface that she designed that was the most usable and preferred used faceted browsing, enabling students to filter data based on various categories to create a pan-able list of books to read.

There were several studies comparing the hierarchical category browser of the Science Library Catalog to more traditional keyword-based interfaces with 9-12 year olds (Borgman et al., 1995). These studies found that children's performance was comparable on directed tasks where keyword searches were easily articulated, but that children did better with the browsing interface when it was an open-ended task or the search required spelling out difficult words.

Highlights

These studies have found students face several challenges and research has found several different successful strategies for addressing the challenges. While how students search for information on the web or within various data catalogs may be different than how they search and explore their personal and peer-collected annotated multimedia data, there are several lessons and challenges that should be taken into account for developing such a system. These are:

1. Students are typically more effective at browsing a data set with options to filter the data than keyword searching.
2. Students have difficulties formulating complex keyword searches; spelling and search syntax can hinder them.
3. Students can perform keyword searches successfully that are simple to articulate and spell.

2.4 Summary

This chapter covered a variety of relevant literature towards designing the Zydeco environment and where Zydeco fits in with existing research.

Section 2.1 provided a background of the science inquiry practices and instructional strategies used in Project-Based Science, which is being used in this study. The challenges learners face in collecting data outside of the classroom and then using a large, student-collected data set were discussed and summarized, which scaffolds in the Zydeco system must address.

Following this, section 2.2 covered relevant systems from the literature that supported students in collecting and using data. Lessons learned from testing these systems and scaffolding guideline research was discussed that can help inform the Zydeco system design to address the challenges learners will face.

Finally, this chapter reviewed various research on how multimedia data collection and usage could be scaffolded, pulling in literature on tagging systems, using voice notes, and also how researchers have supported children in searching on the web (which they will need to do within the student-collect data set in Zydeco:UseData). From each of these areas, key points that relate to the Zydeco system were highlighted at the end.

In the next chapter I discuss the design of the Zydeco tools, which seeks to adapt appropriate techniques from previous work to support students in overcoming the challenges of collecting and using data.

CHAPTER 3

Design Rationale and Implementation of Zydeco

This thesis explores how to scaffold students in collecting multimedia data outside of the classroom and then use an aggregate, student-collected data set to construct scientific explanations. This research is being done within the Zydeco Project, from which this research is studying two integrated tools used in the Zydeco Project:

1. Zydeco:CollectData – running on a mobile device, (e.g. Apple's iPod Touch) this app scaffolds students engaging in multimedia data collection outside of the classroom for later use by themselves and their peers, and
2. Zydeco:UseData – running on a tablet, (e.g. Apple's iPad) this app scaffolds students in using large amounts (several hundred pieces) of personal and peer-collected multimedia data to construct scientific explanations.

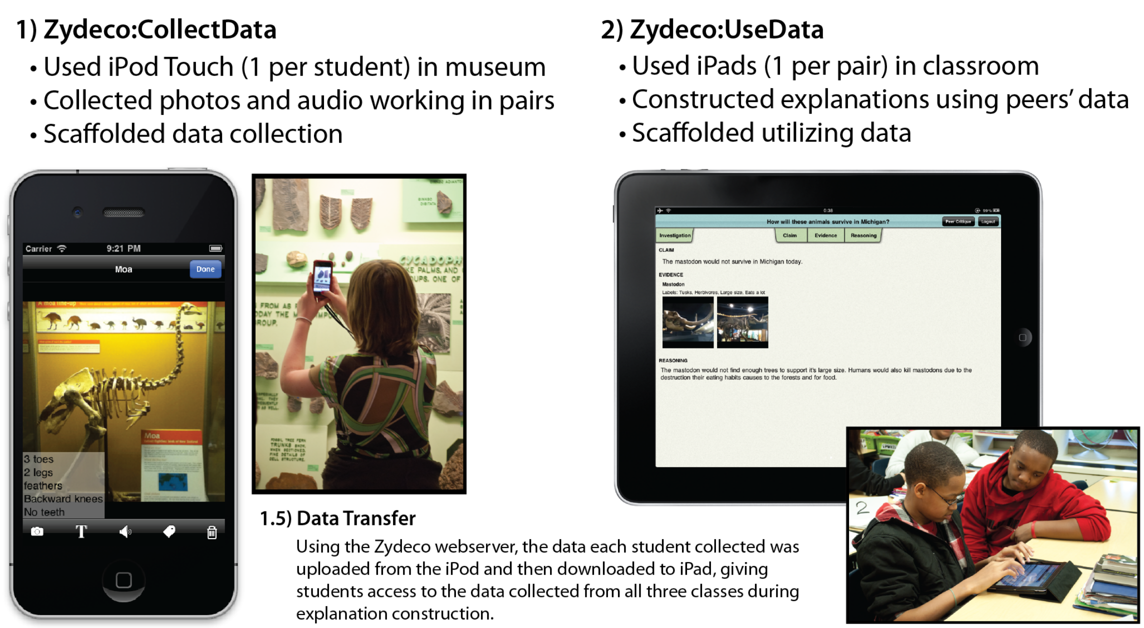

Figure 3-1. Depiction of how the Zydeco:CollectData and Zydeco:UseData tools were used to scaffold students engaging in science inquiry in the May 2011 trial.

This chapter describes the iterative design process that was used in developing the Zydeco tools to investigate this area of research. The current design rationale of each Zydeco tool is also discussed, detailing how the tool adapts existing scaffolding strategies from literature as well as techniques used in non-educational systems around collecting and using multimedia data.

3.1 Iterative Design Process

Zydeco has been designed by a team of researchers, who have tested different aspects of the system throughout the iterative design process (discussed below). The current versions of the Zydeco tools are developed in Objective C for iOS 4.2; the author was responsible for all of the development of the Zydeco tools.

The Zydeco tools were designed by utilizing an iterative learner-centered design process. The Zydeco:CollectData tool started development in November of 2009, while the development of the Zydeco:UseData tool was started in the summer of 2010. Together, they scaffold inquiry processes around data collection, data analysis and interpretation, and synthesizing the data to construct a scientific explanation. In the following sections we discuss how the iterations and trials with the Zydeco tools informed changes to the system. The design of the Zydeco tools presented in this chapter is the result of these iterations, and is the version tested during this field study.

3.1.1 Zydeco:CollectData Pilot Studies

The Zydeco:CollectData tool was tested in more than 10 pilot trials where 2-6 students used the system, helping to form the initial design. These pilots, as well as the later trials, all used the University of Michigan Exhibit Museum of Natural History as the site students used to gather data.

In the first version of the Zydeco:CollectData tool students took on roles of a specific type of scientist (such as a paleontologist) and had a directed inquiry experience, where they investigated a few exhibits in depth pursuing a guided investigation (Kuhn, Cahill, Quintana, & Soloway, 2010). Unfortunately, when we piloted this design with students we found the students were not engaging with the museum exhibits and became focused on completing the activity as fast as possible. This behavior was not desired and the Zydeco system was redesigned towards enabling students to choose what exhibits they want to explore in the museum and allow for more open-ended data collection (which the scaffolds implemented to support this are discussed in 3.2).

While supporting student choice in an open-ended investigation is beneficial to students, as it enables them to experience the richness this environment offers, it adds challenges regarding how to scaffold the process. It becomes less feasible to provide specific scaffolds and supports tailored to assisting students as they interact with a particular exhibit.

After redesigning the system, the next version was a multimedia data collection and annotation system where students were able to collect photographs or audio notes and title and tag them. However, beyond enabling data to be collected and annotated with titles, tags, and audio notes, the system offered no scaffolding for the learner. Initial piloting with students led to students being confused as to what sort of data they were supposed to collect and how to annotate it in a manner to be useful later.

To address these issues, we created scaffolds such as using a tagging system as supportive features, the stepwise guidance of the complex process. The tagging system integrated discussions between students and teachers on how they might annotate data, which were used as a list of potential tags while annotating data (discussed in 3.2.2).

In addition, during pilot testing it was noted that students wanted to explore the environment and collect data that might not directly apply to their investigation. To support students desire to engage with exhibits, we allow students to collect data in an "other stuff" category in limited quantities (5 pieces of data per student for the trial in this thesis) on topics of interest.

3.1.2 Zydeco:CollectData Initial Trials

The Zydeco:CollectData system was then used in two full trials at the University of Michigan Exhibit Museum of Natural History, one with 42 students using it in May of 2010 and another with 85 students in October of 2010. Iterations based on the results of these trials were implemented before the field study reported in this dissertation.

During the first large trial students were collecting data on animal traits to *Construct a new animal that could live near Michigan in the Cretaceous Period*. The author analyzed the factual accuracy and potential usefulness of the tags students applied to data, looking at if the information in the tags was helpful to the investigation (Kuhn, Cahill, Quintana, & Schmoll, 2011). Additionally, the author interviewed 36 of the students to learn how they think about tagging data. This study found students thought tags were useful for labeling and later recalling the data. Analyzing the tags showed that students were largely able to apply factually accurate (83% mean accuracy of tags) and potentially useful (95% of the accurate tags were deemed potentially useful) tags for answering their investigation. In this study, students used Flickr to review their data to construct their new animal by drawing it based on the evidence they reviewed.

The second large trial was held at the same history museum, though in this case a set of stations was set up in a history museum with docents manning each station. Clara Cahill investigated whether the Zydeco:CollectData tool promoted the heads down phenomena, spending time looking down at the device instead of at the exhibits (Cahill, Kuhn, et al., 2011), more than worksheets. That research found there was no significant difference in the attention students paid to exhibits and that Zydeco encouraged more on-topic social interaction with peers (Cahill, Kuhn, et al., 2011). This trial also explored how students used audio notes and tags as a sensemaking support, finding they both can support students in different ways (Cahill, Lo, et al., 2011). The interplay between audio notes and tags was investigated and it was discovered that some audio notes can encourage mindful tagging, while in other cases students had unrelated tags that were off-topic from the investigation and audio notes. However, the study was focused on the behavior of students in the museum and did not investigate how students were using their data after the museum visit.

In summary, prior trials found that:

- Zydeco:CollectData scaffolds appeared to enable students to collect data with accurate and potentially useful annotations, though how the data was later used was not investigated.
- No significant difference in heads down behavior between students using Zydeco and worksheets. Students using Zydeco were noted as having more on-topic conversations than their peers using worksheets.

3.1.3 Zydeco:UseData Pilot Studies

After these trials, the development of the Zydeco:UseData tool commenced. This component was iteratively designed with two children (aged 13-14) over five, two-hour sessions. This was followed by 4 pilot tests of the Zydeco:UseData tool with a total of 13 students using the system in two-hour sessions.

During this pilot testing, many usability issues arose around how information was displayed, the wording of labels within the system, and how users performed various actions, which were corrected through the iterative design. Additionally, the students desired to have additional information to be able to fact-check the student-collected data and learn more about the data in question. This led to the supply supplemental curated data scaffold, discussed in 3.3.1.

3.2 Zydeco:CollectData Design Rationale

Students are able to collect photographic data or record audio notes, annotate the data with a title, annotate the data with an audio note describing the data (if they collected a photograph), and annotate the data with tags in the Zydeco:CollectData tool. This data is stored locally and reviewable on the device, and the database can be synced to a website for later use.

The goal of the data collection system is to support students in two tasks:

- Determining what data is needed to answer their investigation.
- Describing and annotating the collected data for later retrieval by themselves and their peers.

The design and rationale behind the components that address these goals are each discussed in this chapter by first reviewing the cognitive task and associated challenges facing learners. The ways these tasks are addressed are then broken down to:

- *Scaffolding Strategies*: Existing scaffolding strategies from literature that could relate to this challenge are discussed, along with how this could be generally applied in this context.
- *Zydeco Scaffolds*: The scaffolds implemented in Zydeco are described, with relevant screenshots as appropriate, along with their benefits.

3.2.1 Determining Necessary Data

Students must overcome the cognitive task of determining what data is needed to answer their investigation. Students need to determine what data is needed to answer their driving question and collect this data in the field. However, students have difficulty knowing what data to collect to answer complex inquiry questions and struggle with planning how they approach the inquiry task (often not having any plan beyond what they are trying to do right now) and if they have a plan, they have difficulty monitoring their inquiry plans and activities to know what they need to do next (Krajcik & Blumenfeld, 2006).

Cognitive Task	Challenges Facing Learners	Scaffolding Strategies From Literature	Zydeco Scaffolds
Students must determine what data is needed to answer their investigation.	Students have difficulty planning and monitoring their inquiry plans (Krajcik & Blumenfeld, 2006). They struggle determining what data to collect to answer complex inquiry questions (Berland & McNeill, 2009; Griffin, 1998).	Provide guidance to facilitate productive planning and monitoring (Quintana et al., 2004). Describe complex tasks by using ordered and unordered task decompositions (Quintana et al., 2004).	Decompose the problem into more manageable parts: The inquiry problem was broken down into sub-questions in the pre-activities, which students need to collect data under these questions to direct their inquiry. Students can review the list of sub-questions; each question gives a visual indicator with the amount of data collected for it, assisting students in planning what questions they need to collect data on.

Table 3-1. Challenges for students determining data to collect and how Zydeco scaffolds the process.

Scaffolding Strategy

One existing science inquiry guideline is to describe complex tasks by using ordered and unordered task decompositions (Quintana et al., 2004). Project-Based Science [PBS] recommends that teachers help scaffold students by assisting them to develop sub-questions to breakdown the complex inquiry problem into more manageable portions (Krajcik & Blumenfeld, 2006). For example, if students were investigating *How is my animal related to other animals?,* the teacher could discuss with the class sub-questions to answer to help learn, such as *What are similar internal traits*? and *What are similar external traits?*.

When students engage in data collection during PBS, they tend to record data on worksheets to answer each of these sub-questions, and then synthesize the knowledge gained from each of these sub-questions to help them in answering their driving question. This mechanism can be applied to the digital environment. Students could collect data to specifically answer a driving question through its sub-questions. Another scaffolding strategy is to provide guidance to facilitate productive planning and monitoring (Quintana et al., 2004); by being able to monitor this list of questions and see which they have answered, the questions can assist students in focusing on the tasks they must accomplish.

Zydeco Scaffold

To scaffold this, the Zydeco:CollectData decomposes the inquiry problem into more manageable parts by breaking down the data collection task into sub-questions. Students and teachers discussed and broke down their inquiry question into sub-questions before entering the field, which the tool uses as the basis for how students begin to collect data, helping decompose the complex task of answering the overall inquiry problem into a more manageable task.

The data collection activity starts by students reviewing a list of all the sub-questions they need to answer for their driving question (Figure 3-2). Students decide which sub-question to collect data to answer by tapping on it, initiating the data collection process. Each sub-question gives a visual indicator of the number of data items under it to provide feedback on questions that might need more data.

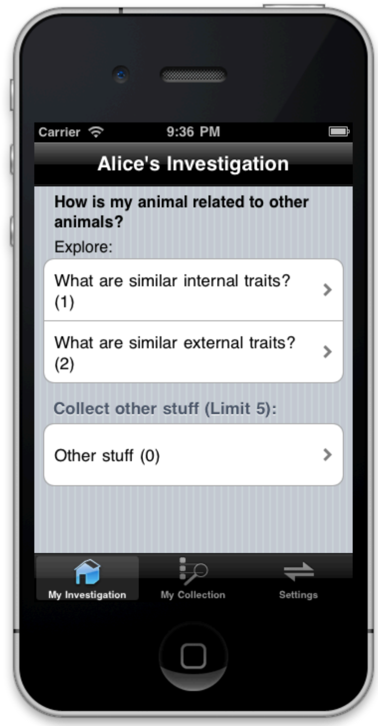

Figure 3-2. Zydeco:CollectData listing all the sub-questions the student needs to collect data to answer.

3.2.2 Annotate Data

Students must describe and annotate the collected data for later retrieval and use by themselves and their peers. Students lack the knowledge that experts have when engaging in inquiry and may not understand or reflect on how data needs to be annotated to be useful later. If students do not collect and annotate the data in a manner that is understandable by themselves and their peers they will be unable to use the data to construct a scientific explanation. Furthermore, students may be unable to search through the data set and find data later if the data is not properly organized (Vavoula et al., 2009).

All of these challenges are compounded by the fact that students can be cognitively overwhelmed outside of the classroom (Balling & Falk, 1980). The learner can get caught up in non-germane but difficult tasks like data entry, which can hinder their ability to collect data (Tossell et al., 2010).

Cognitive Task	Challenges Facing Learners	Scaffolding Strategies From Literature	Zydeco Scaffolds
Students must describe and annotate the collected data for later retrieval by themselves and their peers.	Students may be unable find data later if not properly organized (Vavoula et al., 2009). Students lack the knowledge experts have engaging in inquiry and may not understand or reflect on how data needs to be annotated to be useful later (Metz, 2000). Students may have difficulty connecting their past knowledge to the artifacts and phenomena they observe (Roschelle, 1995). Students can be cognitively overwhelmed in the field (Balling & Falk, 1980); non-germane but difficult tasks such as the mechanics of data entry can hinder their ability to collect data (Tossell et al., 2010).	Make disciplinary strategies explicit in the artifacts learners create (Quintana et al., 2004). Facilitate the organization of work products (Quintana et al., 2004).	Enable data objects to be annotated in multiple ways: students can link multiple annotations to multimedia data to be able to recall idea(s) when viewing that data object as a whole.
		Embed expert guidance to help learners use and apply science content (Quintana et al., 2004).	Integrate co-created tags that reflect expert discussion: The co-created tags are potential tags can provide structure to students in annotating data in a consistent manner. The consistent tag vocabulary improves the organization of the data for later use. The display of previously generated tags can prompt students to reflect on why they are collecting this piece of data and how the data can potentially be annotated by reminding the student of the discussion they had with an expert.
		Provide reminders and guidance to facilitate articulation during sense making (Quintana et al., 2004).	Guide Collection Step-Wise: Students are required to walk through the data collection and annotation steps (title, tag, audio note) in a set order and are given textual prompts at each stage of the process to instruct them on how to annotate the data.

Table 3-2. Summary of the challenges, scaffolding strategies, and scaffolds used to address annotating data for later recovery.

Scaffolding Strategy 1

Two scaffolding strategies that can be used in this case are to "make disciplinary strategies explicit in the artifacts learners create" as well as "facilitating the organization of work products" (Quintana et al., 2004). To employ these guidelines, the structure of data the students collect should be meaningful and in a method similar to how a scientist might use collecting multimedia data in the field. This data structure should be useful not only to organize their thoughts around the data, but also make later retrieval of the data easier.

Zydeco Scaffold for Strategy 1

To incorporate this strategy, Zydeco enables data objects to be annotated in multiple ways. The system enables data objects to contain an optional photograph, linked with a title, one or more tags, and an optional audio note (Figure 3-3).

By enabling students to link these multiple annotations together, it can help the learner organize their thoughts around a particular idea in a single data object, which can be easier to review as a single object than multiple items back in the classroom.

Another benefit of having multiple types of annotations is that each can be used for different purposes if needed. Students can use the title and tags to apply textual descriptions to the data, which can be searched later to retrieve and review the data. Students can use audio notes to record open-ended notes while avoiding the spelling and typing issues of data entry. These audio notes can describe the artifacts in question or ideas and questions concerning their investigation.

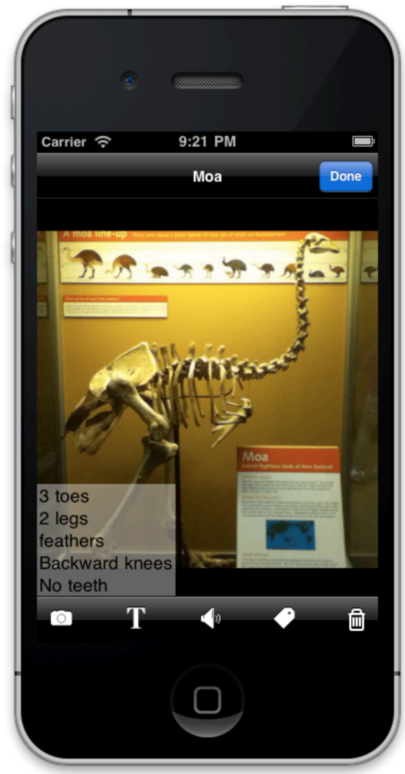

Figure 3-3. Viewing a datum titled "Moa", containing a photograph and five tags.

Scaffolding Strategy 2

Another scaffolding strategy for science inquiry is to "embed expert guidance to help use and apply scientific content and practice" (Quintana et al., 2004). This strategy does not detail how the guidance should be embedded into a mobile system. However, research has noted that embedded guidance needs to become part of the data collection process or else learners are apt to ignore the scaffold (Quintana, 2001). We also know that the device has a small screen that is easy to clutter and reduce the usability of the system unless scaffolding is streamlined (Luchini et al., 2004).

A goal for embedding the guidance is to help learners annotate the data in a manner that will help them answer their investigation later. Additionally, since the data will be shared with their peers, it is important for their annotations to be understandable by their peers so they can use it, as Section 2.3.1 highlighted issues tagging systems have where users are unable to understand the meaning of each other's tags.

Furthermore, if the annotations were to be used for searching upon the data set later, it would be helpful to have the annotations done in as consistent a manner between students as possible. In reviewing the tagging research in 2.3.1, it was noted that inconsistent tag wording could reduce the effectiveness of users to be able to search upon the data set.

Zydeco Scaffold for Strategy 2

We integrated co-created tags that reflect expert discussion to scaffold students to use in annotating the data. Students and teachers discussed and co-created a list of tags that consists of different labels that could be used to annotate the data. The tool integrates these co-created tags, displaying the list of co-created tags for potential use (Figure 3-4). To reduce the data entry mechanics around applying the co-created tags, learners only had to tap on a tag to apply it to the data.

By integrating the co-created tags that were discussed in class, students are able to understand the semantic meaning of the tags and can remember the expert guidance from the teacher on how they might apply the language. For example, students and the teacher discuss and decide that “four legs”, “large brain”, and “ears on top” are potential labels that students may use collecting data for their investigation on animal relatedness, and come to a shared understanding of what is meant by “top” and “large”.

When the learner is in the field, carrying out an investigation and finds an exhibit of a hippopotamus, the learner can get an idea of what useful annotations are by reviewing the co-created list of tags on their handheld during the tagging phase of data collection; this allows the student to access the expert guidance from the discussion with their peers and teacher. The learner can then reflect upon the possible annotations in the co-created tags list and examine the hippopotamus, tapping on an annotation in the list to annotate the data with “four legs” and “ears on top” upon reviewing the hippopotamus’s features.

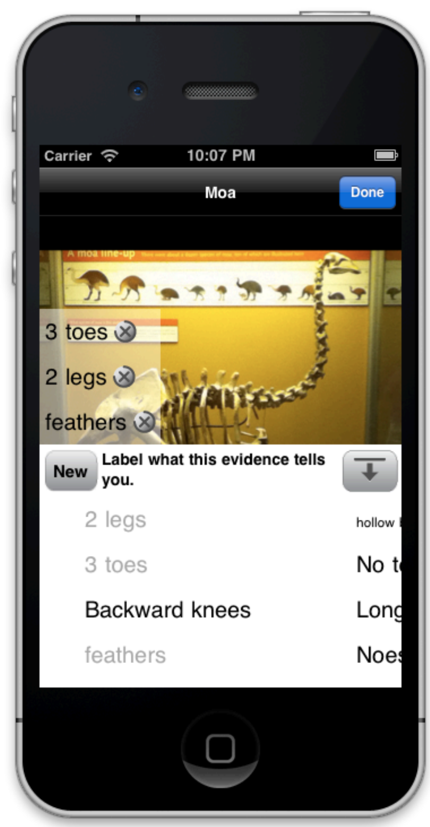

Figure 3-4. Zydeco:UseData displaying the co-created tags, listing the potential tags students can tap to apply to the date.

By students using the co-created tags, it can further encourage consistent annotations (e.g. everyone writes "four legs" instead of other methods such as "4 legs" or "legs (4)"), which can assist in the retrieval process as the more varied tag sets were found to be harder to search upon.

Scaffolding Strategy 3

A science inquiry strategy that can be adopted to address the challenge of annotating data in a useful manner is to "provide reminders and guidance to facilitate articulation during sense making" (Quintana et al., 2004). By guiding students through a set of tasks with prompts, it can encourage them to reflect on the task they need to accomplish at each stage; the goal of this being to encourage students thoughtfully annotate the data. If done in a step-by-step manner, this can provide a structure for ensuring students walk through the scientific process. Furthermore, textual prompts can be provided at each stage to contain either process management or reflective support, as needed.

Zydeco Scaffold for Strategy 3

Guiding the data collection process in a focused step-by-step manner was implemented in Zydeco by having the system flow and prompts walk the user through the data collection process. First, students choose a sub-question to collect data under. Then they choose the type of multimedia they want to collect on a second screen (audio note or photograph). Following this, the annotations each have separate steps in the process as well: title the data (Figure 3-5), include an (optional) audio note, and then lastly tag the data (requiring at least one tag on each data object, Figure 3-4).

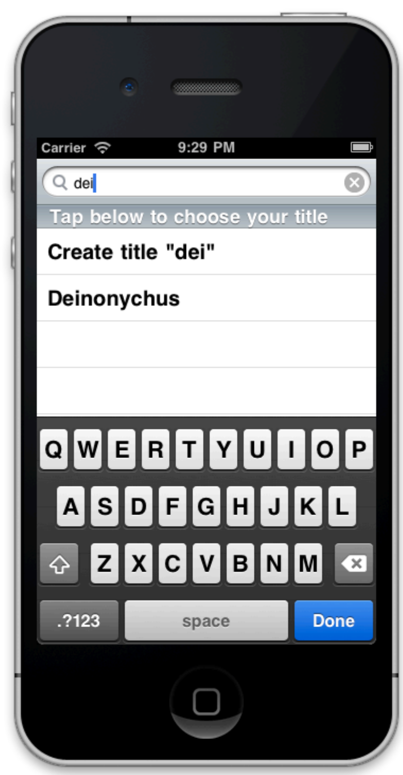

Figure 3-5. Prompting the students to title the data that was collected.

3.3 Zydeco:UseData Design Rationale

In the Zydeco:UseData component, the data the students collected is downloaded from a website where it was synced from the Zydeco:CollectData system, and students can browse through the data set, find data they are interested in, and then assess and select data to use in their explanation. The final explanation is in McNeill's Claim, Evidence, and Reasoning format, from which the data is being used as evidence (McNeill & Krajcik, 2011).

The explanation construction system relies upon the scaffolds in the data collection system to enable students to collect an accurate and useful set of annotations on the data; if the data set is not useful towards answering their investigation, the students will be unable to use the data to construct explanations regardless of the other supports.

The focus of this study is on how students find and assess the data for use in constructing their scientific explanation. Because of this, the focus is on the challenges students encounter concerning two tasks:

- Finding data by searching through and exploring the data set.
- Assessing the data they and their peers collected to judge whether it is useful for their investigation.

The scaffolds were developed to overcome the challenges of the tasks around using the data are discussed in detail below. As both of the tasks students must do are non-linear and there can be a lot of overlap in the challenges and scaffolds developed to address the tasks, they are discussed together, beginning with a description of the cognitive task and challenges facing the learners, and then describing the scaffolding strategy and Zydeco's implementation of the strategy.

3.3.1 Find and Assessing Data

To complete this activity, students need to search through and explore the data set and they need to find and assess the data they and their peers collected to judge whether it could be useful for their investigation.

Students can have organizational issues dealing with large amounts of multimedia data (Vavoula et al., 2009). When multimedia data is annotated students struggle with how to search for data and have difficulty constructing appropriate search syntax (Druin et al., 2009; Bilal, 2001). The difficulty students face with search syntax can be complicated further as students may be unaware of what types of data are present within the data set, making them unable to know what to search on.

Even if learners are able to find data, they struggle with using appropriate and sufficient evidence to support their explanation (McNeill & Krajcik, 2011). Even with curated information, students struggle with using appropriate and sufficient evidence; however students have the greatest difficulty dealing with large, student-collected data sets as this data can contain information that is inaccurate and/or not be useful for constructing a scientific explanation (Berland & McNeill, 2009).

Cognitive Task	Challenges Facing Learners	Scaffolding Strategies from Literature	Zydeco Scaffolds
Students need to search through and explore the data set. Students need to assess data they and their peers collected to judge whether it is useful for their investigation.	Students can have organizational issues dealing with large amounts of multimedia data (Vavoula et al., 2009). Students have difficulty searching for information (Druin et al., 2009) and have difficult constructing appropriate search syntax (Bilal, 2001). Children struggle with using appropriate and sufficient evidence (K. L. McNeill & Krajcik, 2011). Students have the greatest difficulty dealing to deal with large, student collected data sets as this data can be inaccurate and/or not be useful for constructing a scientific explanation (Berland & McNeill, 2009).	Provide structure for complex tasks and functionality (Quintana et al., 2004). Provide representations that can be inspected to reveal underlying properties of data (Quintana et al., 2004).	Enable assessment and filtering by data characteristics: Enable reviewing the data characteristics, which were collected with the Zydeco handheld scaffolds & supports to encourage accurate and useful annotations. Also enable filtering of the textual data characteristics; sub-question, tag, title, and by whether it was collected by their group, or their class. Enable exploratory search: Enable exploratory searching by allowing students to browse the entire data set to inspect the data collected as well as browse through the list of tags and tap on a tag to find data matching it, reducing the search mentality required and enabling students to more easily explore potential data. Expedite searching for similar data: Finding similarly titled or tagged data when focused on a particular piece of data. Supply supplemental curated data: Supply additional curated information on the most popular artifacts collected to provide supplemental information to lookup and fact check.

Table 3-3. Summary of the challenges, scaffolding strategies, and scaffolds used to address the searching for and assessing data in the student-collected data.

Scaffolding Strategy

This research seeks to scaffold students in being able to have different methods to inspect and analyze the data set. This follows an existing guideline: "Provide representations that can be inspected to reveal underlying properties of data" (Quintana et al., 2004). In doing so, the system needs to allow students multiple methods to search through and organize the data as they examine the data set, and view the data in different manners to organize their thoughts.

However, students have difficulties organizing and searching through data, Another inquiry guideline gives conceptual advice to address this, suggesting "providing structure for complex tasks and functionality" (Quintana et al., 2004). By providing structure for the task of searching through data, the complexity can be reduced to a level that is manageable for the students.

Zydeco Scaffold #1

Zydeco enables students to assess all the data characteristics and filter the data by any textual annotation or data characteristic. Data characteristics include the photograph, audio note, title, tags, sub-question, and whether the data was collected by the group or their peers. Students need to be able search through their personal and peers' data to find appropriate data for use as evidence. When students are reviewing the data, they need a means to assess the data object in detail and examine the media and annotations. Zydeco addresses these needs through filtering and review options.

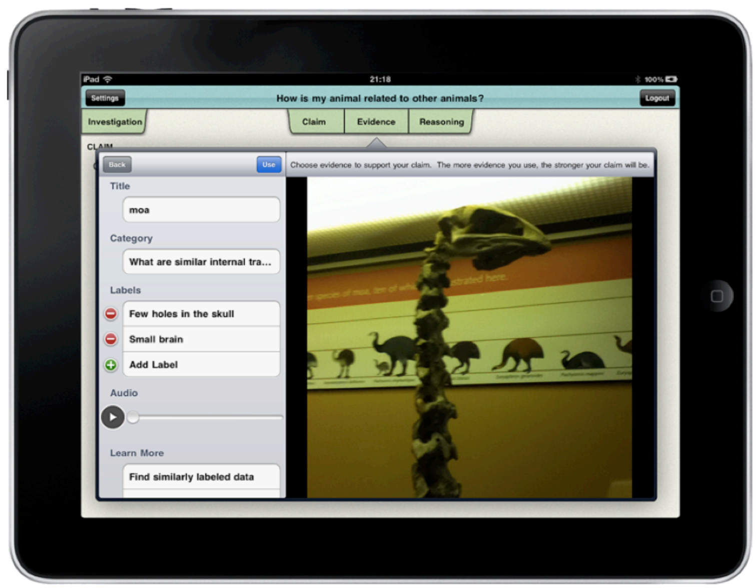

Figure 3-6. Assessing the various annotations and image on the data.

When examining evidence, students can filter the data by the tag, title, and sub-question it was collected under, and choose to view only data collected by them or to look at the collective data set. This enables students to condense the data set by whichever means they want to view it, analyzing a relevant subset of the data (e.g. looking at only data titled "Eagle" to view all the data on Eagles).

Zydeco enables students to review all the annotations on the data and also zoom in on the photograph to examine it in greater detail (Figure 3-6). This can enable the students to examine the linked ideas of the data object together, taking into account all the different annotations and examining particular features of the image by zooming in to a spot for further analysis.

Zydeco Scaffold #2

The second scaffold the tool employs toward this strategy is enabling exploratory search to view characteristics of the data set.

When searching for information, students should be able to openly explore data. The tool scaffolded learners through an open exploration of the data set by enabling them to scroll through thumbnail tiles of data with condensed metadata information (displaying the title, tags, and images).

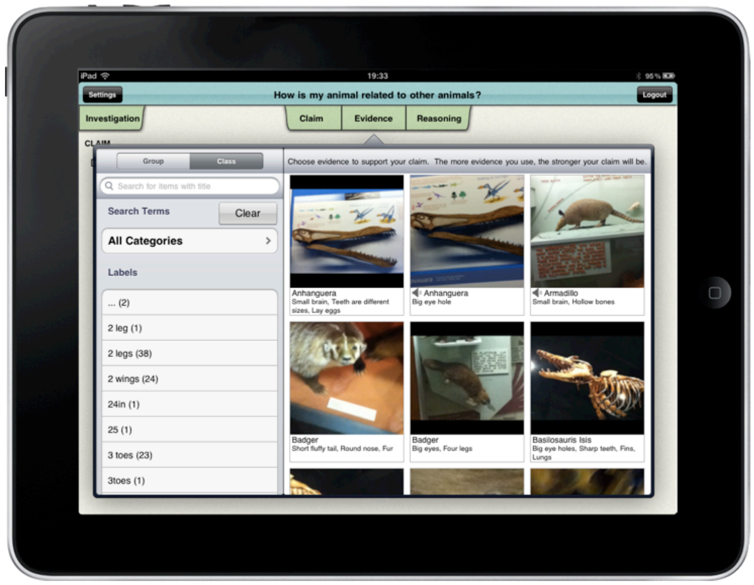

Figure 3-7. The data review page allowing students to view tiled images with titles and tags and see the list of tags on the displayed data.

Additional exploration was enabled by allowing students to view a list of all the tags the data was annotated with and tap on any tag to display all the data that matches that tag. Searching by tags has been shown to be beneficial for novices to do exploratory search (Kang et al., 2010). One method used on the web is a tag cloud to display a list of how objects are tagged on the site and the relative frequency each tag is used. However, previous research has found these tag clouds to be more usable for finding information when in the format of an alphabetical list (Hearst & Rosner, 2008) and this alphabetical list was implemented in the tool.

Finally, the evidence page aggregated all of the tags applied to the data in the currently displayed data set. For instance, if a student searched for data titled “Eagle” they would see a list of all of the tags applied to Eagle data, as well as how many times each tag was used to annotate the Eagle data. This aggregate list of tags and the count for each tag is displayed on the sidebar in alphabetical order, enabling students to get a sense of the data and apply additional filters by tapping on a tag (seen in the left side of Figure 3-7).

Zydeco Scaffold #3

The third scaffold Zydeco employs toward this strategy is expediting the task of finding similar data. If students desire to find data similarly annotated to a piece of data, the task can be challenging, though not germane to the overall inquiry task. Search engines like Google utilize features where a user can find similar sites (Google inc.). This feature helps expedite the process of retrieving more information that is similar to a site of interest, and could similarly be applied in this context. A similar feature was implemented in Zydeco - when a student is examining a single piece of datum in the detailed view, they can choose to find similarly labeled or similarly titled. This automatically applies either a filter for all the tags present on that data object, or a title search for the same title name (respectively).

Zydeco Scaffold #4

The final scaffold Zydeco employs toward this strategy is supplying supplemental, curated information. Pilot testing found that students might not gather all the data they need for answering their investigation-- in this context, students needed data regarding particular animal traits.

Zydeco provides supplemental information to enable students to fact-check the annotated data and fill in gaps of information that the student-collected data is lacking. In this study students were learning about animal relatedness, so students were given supplemental information on the various animals. This information consisted of an image of the animal as well as two or three paragraphs of age appropriate information about the animal's background and traits. Due to the wide range of animals students were able to collect data on and time constraints, the supplemental information was only provided for the more common animals students collected data on. In this trial, supplemental information was provided on any animal at least 5 students collected data on.

3.4 Summary of the Zydeco Design

This chapter discussed the iterations, design, and rationale of Zydeco's two tools, Zydeco:CollectData and Zydeco:UseData. Each tool was designed adopting scaffolding strategies from the literature to scaffold students through collecting and using large quantities (several hundred pieces or more) of multimedia data to construct scientific explanations.

These two tools connect the in class and out of classroom contexts by transferring information between the tools, enabling this inquiry activity to take place. One method to support this transfer are the co-created tags (CCT), which are reflected in both tools and developed with expert guidance, ensuring all students had an understanding of their semantic meaning. This allowed students to annotate data in a useful manner during the out of classroom context. When back in the classroom, the CCT then became a means of scaffolding students as they interpreted and evaluated the data annotated with these CCT as the CCT had a defined meaning.

These co-created tags are an instantiation of a co-created language, a language that the students and teacher agree upon to use to annotate the data and understood the meaning of. This scaffolding technique enabled students to overcome the cognitive tasks that they faced during the activity. A summary of the cognitive tasks, challenges facing learners, scaffolding strategies, and the resulting scaffolds used in each Zydeco tool can be seen in Table 3-4 and Table 3-5.

Cognitive Task	Challenges Facing Learners	Scaffolding Strategies From Literature	Zydeco Scaffolds
Students must determine what data is needed to answer their investigation.	Students have difficulty planning and monitoring their inquiry plans (Krajcik & Blumenfeld, 2006). They struggle determining what data to collect to answer complex inquiry questions (Berland & McNeill, 2009; Griffin, 1998).	Provide guidance to facilitate productive planning and monitoring (Quintana et al., 2004). Describe complex tasks by using ordered and unordered task decompositions (Quintana et al., 2004).	Decompose the problem into more manageable parts: The inquiry problem was broken down into sub-questions in the pre-activities, which students need to collect data under these questions to direct their inquiry. Students can review the list of sub-questions; each question gives a visual indicator with the amount of data collected for it, assisting students in planning what questions they need to collect data on.
Students must annotate and organize the collected data for later retrieval by themselves and their peers.	Students may be unable find data later if not properly organized (Vavoula et al., 2009). Students lack the knowledge experts have engaging in inquiry and may not understand or reflect on how data needs to be annotated to be useful later (Metz, 2000). Students may have difficulty connecting their past knowledge to the artifacts and phenomena they observe (Roschelle, 1995). Students can be cognitively overwhelmed outside the classroom (Balling & Falk, 1980); non-germane but difficult tasks such as the mechanics of data entry can hinder their ability to collect data (Tossell et al., 2010).	Make disciplinary strategies explicit in the artifacts learners create (Quintana et al., 2004). Facilitate the organization of work products (Quintana et al., 2004).	Enable data objects to be annotated in multiple ways: students can link multiple annotations to multimedia data to be able to recall idea(s) when viewing that data object as a whole.
		Embed expert guidance to help learners use and apply science content (Quintana et al., 2004).	Integrate co-created tags that reflect expert discussion: The co-created tags are potential tags can provide structure to students in annotating data in a consistent manner. The consistent tag vocabulary improves the organization of the data for later use. The display of previously generated tags can prompt students to reflect on why they are collecting this piece of data and how the data can potentially be annotated by reminding the student of the discussion they had with an expert.
		Provide reminders and guidance to facilitate articulation during sense making (Quintana et al., 2004).	Guide Collection Step-Wise: Students are required to walk through the data collection and annotation steps (title, tag, audio note) in a set order and are given textual prompts at each stage of the process to instruct them on how to annotate the data.

Table 3-4. Summary of the tasks, challenges, scaffolding strategies, and scaffolds in the Zydeco:CollectData tool.

Cognitive Task	Challenges Facing Learners	Scaffolding Strategies from Literature	Zydeco Scaffolds
Students need to search through and explore the data set. Students need to assess data they and their peers collected to judge whether it is useful for their investigation.	Students can have organizational issues dealing with large amounts of multimedia data (Vavoula et al., 2009). Students have difficulty searching for information (Druin et al., 2009) and have difficult constructing appropriate search syntax (Bilal, 2001). Children struggle with using appropriate and sufficient evidence (K. L. McNeill & Krajcik, 2011). Students have the greatest difficulty dealing to deal with large, student collected data sets as this data can be inaccurate and/or not be useful for constructing a scientific explanation (Berland & McNeill, 2009).	Provide structure for complex tasks and functionality (Quintana et al., 2004). Provide representations that can be inspected to reveal underlying properties of data (Quintana et al., 2004).	Enable assessment and filtering by data characteristics: Enable reviewing the data characteristics, which were collected with the Zydeco handheld scaffolds & supports to encourage accurate and useful annotations. Also enable filtering of the textual data characteristics; sub-question, tag, title, and by whether it was collected by their group, or their class. Enable exploratory search: Enable exploratory searching by allowing students to browse the entire data set to inspect the data collected as well as browse through the list of tags and tap on a tag to find data matching it, reducing the search mentality required and enabling students to more easily explore potential data. Expedite searching for similar data: Finding similarly titled or tagged data when focused on a particular piece of data. Supply supplemental curated data: Supply additional curated information on the most popular artifacts collected to provide supplemental information to lookup and fact check.

Table 3-5. Summary of the tasks, challenges, scaffolding strategies, and scaffolds in the Zydeco:UseData tool.

CHAPTER 4

Experimental Design and Methods

In this section the preparation of the research study, the data collected, and the methods used to analyze our data are presented.

4.1 Research Approach

4.1.1 Field Study

Since the goal of this research was to determine how the students used the tools given the scaffolds as they proceeded through a science inquiry investigation, a field study was chosen as the method of learning about student usage. The goal was to utilize the Zydeco tools and activity structure in a format similar to what a teacher would use. This involves activities performed in class as well as an out of class field trip to collect data, to try and be as naturalistic as possible while gathering data on how students used the scaffolded tools. Through the course of this field study, students had 6 days (1 hour a day) of pre-activities, 1 hour of data collection in a museum using Zydeco:CollectData, and 5 days (1 hour a day) of post-activities, four of which they were using the Zydeco:UseData system..

Field studies are one of the most common methods used to understand how children interact with technology. In a study analyzing ten years of published research on technology design for children, field studies were the most common approach (53%), followed by action research (42%) and lab experiments (37%) (Jensen & Skov, 2005).

The field study allowed us to discover how students are able to complete the task and interact with the scaffolds in the system. From this, we can note where the scaffolds were successful at enabling students to complete the task and also find areas where students need additional scaffolding. This knowledge paves the way for future research on comparative studies that further investigate particular scaffolds of interest.

4.1.2 Planning

In order to test Zydeco in a classroom, a set of curriculum first needed to be developed. This required picking out a field site, in this case a history museum (discussed in 4.2.2), and a school program that had learning objectives that would align with an investigation around the artifacts present at the museum.

The goal of the curriculum was to develop an investigation where students would learn about animal traits, which is a required state learning objective for 6th grade students in the state of Michigan. Learning about animal traits aligned well with the artifacts of animals at the history museum and so met our goal of taking advantage of the field site.

In order to develop the curriculum, three researchers had biweekly meetings with the 6th grade science teacher at the elementary school and the Education Director of the history museum over the course of four months. This was done to develop an appropriate driving question that would combine the state learning goals the teacher needed to pursue and be relevant to the exhibits at the museum, to plan out the pre-activities and topics that need to be introduced to the students, to determine what types of data and what exhibits students would collect data from in the museum, and to decide how the post-activities with the Zydeco:UseData tool would be run.

The scientific explanations students were constructing after the museum visit were chosen as being topics that the science teacher and museum educator felt would be appropriate for the students. The overall flow and content of the activity that was planned is described in the use case in section 1.5.

4.2 Research Sites and Participants

4.2.1 Elementary School

This study was conducted at a diverse suburban school in Southeast Michigan where 62% of the students receive free or reduced lunch. This study followed 54 students (aged 11-13) from three 6th grade classes at the school, and the letters M, K, and B in the results chapter identify the individual classes they attended. While we worked with three classes that had three separate teachers, one teacher was responsible for teaching science to all of the classes. This primary contact teacher was involved with the curriculum planning and field study and had previously been involved in a Zydeco trial where (different) students went on a field trip to the Ann Arbor Exhibit Museum of Natural History. For the trial the primary teacher was in charge of the three science classrooms due to her familiarity with the Zydeco and the curriculum. This had the advantage of having a constant teacher influence for each class, though the teacher was less familiar with the students who were not in her daily science class.

Participants were surveyed on their usage of technology. Eighty-eight percent had used a mobile touch-based smartphone before (65% of those students use it daily) and 74% had used a touch-based tablet (33% use it daily). None of the students had used Zydeco before. The school does not teach any form of Project-Based Science.

During the trial, a researcher acted as the teacher for describing the activities and leading the instruction for the three classes, while the teacher acted as supplemental teaching support and handled any behavioral issues. During the six days of pre-activities, one or two additional researchers were present each day to assist with research data collection. In the five days of post-activities, four additional researchers were present: two of these researchers were there to give in-process interviews and handle audio recording students, while the other two researchers took field notes and were present to deal with any technical issues that arose or answer any questions related to inquiry that students had.

4.2.2 Ann Arbor Exhibit Museum of Natural History

The Ann Arbor Exhibit Museum of Natural History is located on the campus of the University of Michigan in the heart of Ann Arbor. Founded in the 1920s, the museum hosts permanent exhibits on anthropology, geology, Michigan wildlife, and evolution, including Michigan's largest collection of dinosaur skeletons. The Zydeco curriculum needed to align with the state standards for 6th grade and the natural history museum provided a great opportunity to study the animal relatedness aspects of the required curriculum, so the trial was scheduled during the time of the year when that subject was normally taught.

During the field trip, students were instructed to visit two of the exhibits halls: The Hall of Evolution on the second floor- which houses Michigan's largest display of prehistoric life including dinosaurs, prehistoric whales, and mastodons- and the Michigan Wildlife Gallery on the third floor- which features local birds, mammals, reptiles, amphibians, and plants. At any given point six research staff and one teacher were available on the floor to answer questions from the students.

4.3 Data Collection and Analysis Procedures

Throughout the field study, eight data collection procedures were used. Parental consent forms were handed out for each student in advance of the trial, and could indicate several different degrees of consent for data gathering [see Appendix A]. Permissions for interviews, surveys, photography, video, and classroom performance were requested.

Students who did not wish to participate in the research trial could opt to take part in the activity and field trip while not having any of their work analyzed.

4.3.1 Pre Surveys

The day before the field trip to the museum each student was given an individual survey to determine their thoughts on science, their previous experience with science, and their prior use of mobile and tablet technologies [Appendix C].

Student survey responses were used to determine if there was any correlation between their perceptions and experience on science and their activity patterns and artifacts produced. These were also taken into account in the qualitative analysis of the data to see if these initial thoughts on science might factor into their work in the field or post-activities.

4.3.2 Video Recordings

While participating in the field trip to the museum, three students wore a headcam. This headcam provided video and audio feed of whatever the student was looking at while they explored the museum collecting data. Due to the bulky nature of the headcam, volunteers were requested and the students were randomly selected from those volunteers who had permission to be videotaped.

The pre and post activities were videotaped in the classroom with two video cameras. Both of these cameras covered portions of the classroom where students who had permission to be video recorded were working. Whenever the teacher was providing instruction, one of the video cameras was focused upon the teacher. These videos were reviewed to examine behaviors and how students interacted with the scaffolds.

4.3.3 Student Collected Data Objects

Each data object that students gathered during the trial and its corresponding annotations was evaluated. In total, the three classes collected 676 pieces of data. These data included work from students in the classes who were not part of this research study and was excluded from further analysis, leaving 474 pieces of data. Forty of the remaining data were audio notes that we requested of the students at the end of their field trip to tell us what they learned. As these audio notes were requested of the students after the activity and were of a specific content, they were also removed from the analysis of the main body of data presented in the results.

This data was used to determine a variety of averages and preferences: type of data that was collected, number of data collected per sub-question, number of annotations per data object, type of annotations applied, and whether they applied preset titles and tags or created new annotations in the museum.

Annotations were also analyzed and coded for accuracy to the data object and usefulness to the investigation. Each annotation type (titles, tags, or audio) was coded on different factors to get a sense at their usefulness in the activity.

Tags: Each of the tags was coded as to whether it was inaccurate or nonsensical in its representation of the data (e.g. "scales" as a tag for data on a fox), accurate to the data (e.g. "fur" as a tag for data on a fox), or if there was uncertainty regarding whether it was accurate or not. The accurate tags were then analyzed as to whether they could potentially be useful to the investigation or not. For instance the tag "New to me", while accurate from the student's perspective, is not helpful in ranking animal relatedness or determining survivability in the investigation.

Titles: Since some of the methods for searching through data relied upon title searching, and titles would be a means in which students identified the data, each title was coded by what was being described: whether it was giving the name of the animal (duck, T-Rex, dino mastodon), a general family name for the animal (bird, dinosaur, dino), a general descriptor (predator, large), or gibberish (amm, ch). All the titles (except those that were gibberish) were analyzed as to whether they were accurate (either naming the correct animal, family, or the description was accurate), not accurate, or unclear. Lastly, the spelling of each title was analyzed as misspelled titles can hinder the students in being able to search on the title, each title being categorized as correctly spelled or not.

Audio notes: Each audio note was labeled as to whether what was said was accurate (no factual statements made that were inaccurate), partially accurate, or not accurate. Additionally, these notes were analyzed as to whether they were potentially useful to the investigations or not useful at all (e.g. "I found the T-Rex on the second floor of the museum next to the Mastodon" is an accurate, but not useful audio note as this information does not help answer their investigation). Since the content of the audio note can also be related to the usefulness, we sought to code each audio note into the themes of what was discussed. One area we wanted to see was if they mentioned tags they applied to their data in the audio note, but in order to determine a coding scheme beyond this, the audio notes were reviewed, recording themes that emerged. These themes were used to develop the rubric for which audio notes content was coded into (one note may be in multiple categories): 1) summarizing labels or text in the museum ("it says the T-Rex has ..."), 2) interpreting labels, text, or displayed specimens ("I think the T-Rex could eat our school bus", "The coyote looks like the fox"), 3) reading labels or text in the museum verbatim, 4) relating exhibit to the tags they annotated the data with, 5) affect ("the Mastodon bones are soooo cool!"), and 6) off topic.

To ensure the rubric for coding the data was consistently understood and ensure the reliability between the researchers, we calculated the inter-rater reliability between samples of the data. For each annotation category that was analyzed, two researchers first coded 20% of the data independently. In cases where each annotation could only be coded into one category we used Cohen's Kappa, a statistical measure of inter-rater agreement for categorical items (Cohen, 1960). However, Cohen's Kappa does not apply when a data object can be coded into multiple categories, and in this case the overall percent agreement between the two reviewers was calculated. A benchmark scale for interpretation of kappa values for inter-rater agreement is presented in the Table 4-1 below (Altman, 1991). In all cases the researchers coded 20% of the data with "very good" agreement, showing that the coding could be done consistently. After the researchers checked the inter-rater agreement, one researcher coded the rest of the data.

Kappa value	Strength of agreement
<0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1.00	Very good

Table 4-1. A benchmark scale for interpretation of kappa values for inter-rater agreement.

4.3.4 Daily In-Process Interviews

Two researchers interviewed three pairs in each class (eighteen pairs in total), speaking with the same pairs for 5-10 minutes each day. The teacher chose one group to be interviewed from each class based on which group the teacher believed would be able to provide the most informative feedback. The other 15 the groups were randomly selected (five from each class). Since the goal of this study was to get a sense of how students used and thought about the system supports and their data, we wanted to increase the chance of having groups that would be able to articulate their thoughts on how they approached the task and issues they encountered during the process.

Before starting the interview process, the interviewers were instructed on how they should prompt students for additional information to questions to promote consistency, including different verbal responses and actions (such as pointing to areas of the screen).

Interviewers were given a packet with instructions detailing a semi-structured interview protocol and the questions that should be asked during each session [see Appendix B]. In addition to the unique questions asked at relevant points of the inquiry process, each interviewer asked four daily questions. These questions related to how students thought about the science inquiry process, using the system, and interacting with the data. The interview was recorded on audio for later transcription, and the interviewer was asked to take notes of things they noted during the interview and filled out a survey on the attitudes and progress of the interviewed pairs at the end of each session [Appendix B]. Some groups were unable to be interviewed on every question due to time constraints.

The interview questions' content and wording was discussed with the teacher and an external educational research evaluator to ensure students would understand them. The interview was piloted with three students before the trial to further vet the questions.

The interview responses and notes were transcribed were organized by group by day. In the analysis we evaluated themes and suggestive trends in various interview questions and placed representative quotes into the results.

4.3.5 Voice Recordings

Certain students were audio recorded while in the museum and back in the classroom. Before the museum trip, 18 groups were chosen for the in-class interviews during the post activities (15 randomly and 3 by the teacher, as discussed in 4.3.4). In the museum, one student in each of these groups to wear an audio recorder hung around their neck.

Back in the classroom, each of the 18 pairs that were interviewed were asked if they would wear an audio recorder while they worked. 17 of them agreed to this and wore an audio recorder for the duration of their time using the tablet; the one group that did not consent to wearing the audio recording for the duration of the activity did agree to have their interviews be audio recorded.

Additionally, the teacher and researchers all wore audio recorders while in the classroom. This was used to later transcribe the student interviews.

Audio recordings were transcribed and analyzed for discussion that displayed emergent themes or difficulties (e.g. students needing to be able to collect data under multiple sub-questions) and to help us determine how students were responding to the supports implemented in the system. (e.g. student discussing as they are titling their data, "Wolverine? It'll show up… w-o-l there it is").

4.3.6 Usage Logs

As the students used the Zydeco:UseData system each touch event was recorded. Due to an error in the logging system the touch events from the Zydeco:CollectData system were not usable for analysis. Each touch event was time stamped and the usage logs were compiled to get the amount of time each group spent on various activities by day, which was compiled to get the total time spent on various tasks, such as: how long students spent searching, what types of searches they preformed, when they used different annotations to perform searches, and what data they applied to their explanations.

These logs were used to assess the activities each group performed and to note trends between all of the groups. The usage logs were also referred to when listening to the in class audio recordings that were taken as groups constructed explanations, to track how they were interacting with the system.

4.3.7 Final Explanations

Screenshots were taken of the final explanations that the students produced for each investigation. Claims, evidence, and reasoning grades from these final explanations were assessed for each investigation according to a specific rubric, which was modified to fit the investigation from a template rubric (McNeill & Krajcik, 2011). For each investigation, two reviewers independently graded 20% of the explanations and checked inter-rater reliability by percent agreements.

4.3.8 Post Survey

Every student was given a post survey [Appendix D] to assess their thoughts on various aspects of using Zydeco and suggestions for improvement. These responses were analyzed to assess potential areas where more support is needed in Zydeco to assist the students.

4.3.9 Additional Qualitative Data

In addition to the notes taken during the interviews, each researcher also kept field notes during the activity. After each day there was a debriefing with the teacher to discuss thoughts on the activity, students' progress, and issues noted.

These notes and discussions were reviewed for trends- for example it was noted by several researchers that groups zoomed into the photos and looked intently at different aspects of the photographs.

CHAPTER 5

Results

Overarching Research Question: What scaffolds can mitigate the challenges for students inherent in collecting multimedia data as well as using a large, student-collected, multimedia data set to construct explanations?

Research Question 1: How do students collect and annotate multimedia data given the system's scaffolds?

Research Question 2: How do students use data from the student-collected data set given the system's scaffolds?

The results in this chapter are presented by analyzing how students collected data outside the classroom using the Zydeco:CollectData tool, followed by how they used data with the Zydeco:UseData tool. Following this, students' explanations are evaluated to determine the quality of the final product that was made with the data collected and used during the activity. Lastly, a summary of how students used each scaffold is presented; this scaffold usage is discussed in Chapter 6 to address the Overarching Research Question.

5.1 Students Using Scaffolds of Zydeco:CollectData

Two researchers graded 20% of the data to determine inter-rater reliability, then the rest of the data was graded by one researcher following the rubric and format described in section 4.3.

Research Question 1: How do students collect and annotate multimedia data given the system's scaffolds?

27 pairs of students using Zydeco:CollectData for an hour in a history museum collected 434 annotated data objects. The scaffolds were overall successful at enabling students to collect multimedia data with accurate and useful annotations:

- The average accuracy of tags per group was 88% and all but 1 accurate tag were potentially useful to the investigation.

- The average accuracy of titles per group was 83%, and an average of 85% of titles were correctly spelled per group.
- The average accuracy of audio notes was 70% per group, and 64% of these audio notes per group were potentially useful to the investigation.

The breakdown of data collected and annotations made by each group along with the averages and medians across all the groups can be seen in Table 5-1.

Group	Data Objects (#)	Tags per Data Object (#)	Co-Created Tags Used (%)	Accurate Tags (%)	Audio Notes (#)	Accurate Audio Notes (%)	Potentially Useful Audio Notes (%)	Accurate Titles (%)	Correctly Spelled Titles (%)
B1	26	2.7	83%	100%	2	50%	100%	100%	100%
B2[1]	-	-	-	-	-	-	-	-	-
B3	20	3.0	82%	88%	3	100%	100%	90%	90%
B4	13	2.5	59%	78%	9	56%	89%	62%	62%
B5	16	3.3	94%	89%	11	91%	64%	100%	94%
B6	7	2.7	89%	89%	3	33%	67%	57%	86%
B7	22	2.0	84%	84%	1	0%	0%	91%	82%
B8	13	3.8	65%	94%	1	100%	100%	92%	100%
B9	10	3.4	100%	91%	3	33%	67%	90%	80%
K1	6	4.3	100%	100%	6	100%	100%	100%	100%
K2	15	2.8	71%	95%	8	100%	75%	100%	93%
K3	11	3.7	93%	80%	0	-	-	91%	91%
K4	14	1.9	73%	92%	10	100%	70%	100%	79%
K5	10	1.1	73%	91%	3	67%	0%	60%	80%
K6	34	3.2	90%	88%	3	67%	67%	68%	68%
K7	30	1.9	60%	89%	8	63%	25%	80%	87%
K8	4	1.8	71%	71%	3	67%	67%	50%	50%
K9	25	3.1	95%	83%	7	86%	71%	92%	92%
K10	8	2.8	100%	95%	3	33%	33%	50%	100%
M1	12	1.2	14%	64%	3	100%	67%	100%	100%
M2	11	2.5	86%	96%	2	100%	50%	82%	100%
M3	15	1.6	62%	83%	0	-	-	47%	53%
M4	21	2.8	95%	83%	4	100%	100%	81%	90%
M5	15	1.2	100%	94%	4	75%	25%	87%	80%
M6	23	2.5	74%	88%	5	40%	60%	91%	91%
M7	15	4.0	93%	95%	2	50%	100%	93%	67%
M8	36	2.6	87%	96%	3	67%	33%	94%	97%
Average By Group	16.62	2.63	81%	88%	4.12	70%	64%	83%	85%
Median By Group	15.00	2.72	85%	89%	3.00	67%	67%	90%	90%

Table 5-1. A breakdown of collected data and annotations by group, graded by researchers to evaluate their accuracy and usefulness.

[1] There was a software bug that caused B2 to lose all of the data that they had collected.

This section begins with a breakdown of the overall characteristics of the data sets students created during data collection and annotation. It then provides a breakdown of how students went about collecting data and applying annotations on the data when out of the classroom, given the system's scaffolds.

Each piece of data the students collected was either a photograph or audio note annotated with an additional title, tags, and an optional audio note (if they took a

Figure 5-1. Example data collected and annotate by students.

photograph). Examples of data students collected can be seen in Figure 5-1. Of the 434 pieces of data whose annotations were analyzed, 327 were photos (75%), 9 were audio notes (2%), and 98 (23%) were a photo that had an audio note appended to it. On average, each pair collected 16 pieces of data.

There was an even distribution of data collected under the sub-questions for *"What are similar in internal traits?"* (42%) and *"What are similar external traits?"* (44%), and far less collected under *Other Stuff* (14%) category (a sub-question for data that did not fit as an internal or external trait). There were a total of 1134 tags applied to the data (2.6/ object) and out of these tags, 945 were tags that were part of the co-created language (83%).

The following sections discuss how students used the scaffolds around capturing the data and provide a detailed breakdown of the results on how students annotated these data.

5.1.1 Data Capture

Zydeco:CollectData scaffolds learners through data capture by guiding the data collection stepwise, beginning the data gathering process with choosing a sub-question on the handheld to collect data that will help answer that question (in this trial, determining whether the data should fall under *What are similar internal traits?* or *What are similar external traits?*). Decomposing the problem into more manageable sub-questions was useful to students as a way to focus their data collection. Student conversations shows evidence that they used sub-questions in the intended method, to guide them as they began their data collection, an example being a student who said, "Look, get the piranha. Is it internal or external?" The sub-questions also influenced how they regarded the phenomena they were viewing and entered the discussion between partners:

> Student 1: "Is it external or internal?"
>
> Student 2: "It's internal, it's inside."

However, it was not possible to collect data under both questions at the same time. This led to the need to gather the evidence twice if it applied to both internal and external traits.

> Student: "Look- get the piranha! Is this internal or external?"
>
> Student 1: "Is this internal?"
>
> Student 2: "This is internal, but you might wanna get it as external because you're getting that too."

Students were able to differentiate what was applicable to their investigation, and use the "Other Stuff" collection category appropriately.

> Student 1: "Look at this huge snail, I'm gonna take a picture. That snail is too big to not have a picture taken. We are going to external traits. Hold on… I can't take a picture. They gotta be related to our animal. But they aren't related to our animal, because different animals…"
>
> Student 2: "Right and she said we could take pictures of things that doesn't even relate to our animal."
>
> Student 1: "Mmm yeah we have to put it under "other" things."
>
> Student 2: "Mmk so 'other stuff'?"
>
> Student 1: "Yes, 'other stuff.'"
>
> -------------
>
> Student 1: "Should I put this under other stuff?"
>
> Student 2: "No, because it's vicious just like our animal."

While being guided through the steps of data collection and annotation via the software, the students were standing in front of the exhibit (e.g. an Eagle figure) to examine the features of the exhibit and associated labels on the exhibit. However, an unanticipated behavior occurred during the tagging process. From audio recordings and the head cam video, we are able to determine that several students moved on from an exhibit shortly after titling the data, tagging it while on the go. Once these students had a photograph of the animal, they appeared to use the digital image to review while tagging the data, instead of lingering at the physical exhibit while applying the tags, showing an unexpected reliance on the digital image to annotate the data (as opposed to the artifact in the museum).

From audio:

> Student 1: "Isn't that a wolf?"
>
> Student 2: "No that's a coyote. Coyote. Umm ears on top, eyes in front, four legs, paws, fur. Ok. Umm. Ooo look at that! Look at the badger!"

From head camera:

> Student 1: "Eww! I'm gonna take a picture of that." Selects external traits SQ. Takes picture and moves on to the title page. Looks around at signage. "What is this? I don't know what kinda animal this is." Titles data, "old animal teeth" and begins to walk away from exhibit. Pauses walking to type tag, "Dirty an big" and then continues on.

A focus on the quality of the data captured, here the picture quality, is a theme that is noted throughout the entire Zydeco system usage, which assisted students in making sense of the data for annotating and later use. From both the audio recordings and head cam videos, we noted many conversations about positioning the camera, retaking the image, and comparing results during data collection. Zoom was a feature that was important to students, as was clarity of the photograph.

> Student: [Takes a blurry image of a grey wolf.] "Nah." [Cancels data collection, and does not try again.]
>
> -------------
>
> Student: "How can I zoom? Oh."
>
> -------------
>
> Student: [Walks to the possum exhibit.] "I gotta take a picture of this to teach all the kids not to mess with all these. That's a good picture, look at that." [Shows iPod to another student collecting data on the same exhibit.] "You got a better one, cuz you zoomed in."

5.1.2 Data Annotations

To determine how students annotate the data and the quality of the different annotations we analyzed how groups used the Zydeco system to show cases where students were successful or areas where students need more support. After assessing how each pair collected and annotated data[2], we looked at the aggregated student-collected data set as a whole since this is what students would be using back in the classroom to construct their scientific explanations.

Each tag was analyzed by a member of the research team and coded as to whether it was inaccurate or nonsensical in its representation of the data, accurate to the data (e.g. "fur" as a tag for data on a fox), or if there was uncertainty regarding whether it was accurate or not. The accurate tags were then analyzed as to whether they could potentially be useful to the investigation or not. For instance the tag "New to me", while accurate from the student's perspective, is not helpful in ranking animal relatedness or determining survivability (this was the only tag that was graded as accurate but not useful).

Every group had applied at least 64% of their tags accurately to the data. Of these accurate tags, all but one were judged as potentially useful ("New to me" was the tag that was accurate but not useful for the investigation, 100% agreement between reviewer for IRR) on grading usefulness).

Groups as a whole used the co-created tags for at least 59% of their tag annotations (median group used co-created tags for 85% of their tags), except for M1 (Table 5-1). For this reason, M1 was an interesting case to look at their usage in detail.

The M1 pair used only two co-created tags out of the 14 they applied (thus using co-created tags for 14% of their tags), and also had the lowest accuracy of any group (64%). However, the lower factual accuracy comes from the increased number of tags that were graded as uncertain by the reviewer (4 of their 5 tags that were not accurate were labeled as uncertain). This example demonstrated that some of the tags the learners create in the museum are harder to analyze due to the open-ended nature. Some of their uncertain tags were labeling information about exhibits that would not be useful in their investigation such as "Female" for a Mastodon skeleton and "Dirty an big" for a photo of teeth titled "old animal teeth". They also had a tag of "bio" for an audio note describing a five-lined skink (that was titled "skinks") that may be intended for labeling the idea behind the data. The use of a tag to potentially summarize the contents of the audio note was a usage characteristic

[2] "Student-collected data"

that was not noted in other groups, though could be useful to encourage given the appropriate activity context (all the co-created tags were internal or external traits).

The M1 pair only created 1.2 tags per data object, barely above the required 1 tag per object and far less than the average of 2.6 tags per data object (only one other pair had less tags per data object than M1, which was K5 with 1.1 tags per object). This low usage of tags may be due to the increased effort needed to type in the tags versus using the co-created tags already in the system that are tapped to apply.

Beyond M1 and K5, all of the other groups had over 70% of their tags be factually accurate. Annotating data is a difficult task for students, and this can indicate they were reflective in applying tags to the data to accurately describe characteristics of artifacts they are observing.

While 24 (89%) of the groups recorded audio notes, these were of mixed accuracy and usefulness, discussed below in the audio notes section.

Next, we break down the collective statistics for each type of annotation (tags, titles, and audio notes).

Tags

When tagging data, students were able to apply up to five tags either by tapping from either their co-created language of tags or by typing in a new tag. Throughout data collection, students applied an average of 2.6 tags to every piece of data they collected. 17% of these were tags the students added on the fly, the remaining 83% of the tags were from the available bank of tags co-created back in the classroom.

When looking at the set of tags as a whole, 1016 tags (89.6%) were accurate, 91 (8%) inaccurate or nonsensical, and 27 (2.4%) uncertain (Cohen's Kappa=0.82). If the data was looked at without the potential off-topic tags from the "Other Stuff" question, there was a small increase in accuracy; there were 943 (90.8%) accurate tags out of 1039 total tags, which shows a 1.2% increase in accuracy for the data set.

To see if there was any significant correlation between the tags previously generated through class discussion (co-created tags) and the user-created tags, their accuracies were compared, excluding again the data in the "Other Stuff" category. For the previously generated (co-created) tags applied, 866 out of 945 tags were accurate (91.7%), while 108 of the 128 user-created tags were accurate (84.4%). To analyze whether there was a statistical difference between the accuracy of previously generated tags versus user-created tags, I used a Wilcoxon signed-rank test. The Wilcoxon signed-rank test is a non-parametric alternative to the paired student's t-test and can be used for numerical values. The null hypothesis for this test was that there would be no difference between the accuracy of the two kinds of tags. The results of this test were statistically significant at the 0.1% significance level (p=0.0005), thus rejecting the null hypothesis and showing that previously generated tags used by groups were significantly more accurate than user-created tags.

Overall the data was tagged well, the vast majority (~90%) of the tags were both accurate and potentially useful to the investigation; though that left a portion of tags (~10%) that students would be searching through that was incorrect or nonsensical.

Titles

Both of the supports around titling data, auto completion of text and using a preset browse-able list, were used by students. From the audio recordings we noticed several cases where students found the list of preset titles (which is filtered as they type) to be helpful, and from head camera footage we were able to see a student repeatedly scrolling through the listing of the names of the animals in the museum to find an appropriate title.

> Student 1: "Wolverine?"
>
> Student 2: "How do you spell that?"
>
> Student 1: "W- It's going to show up. W-O-L- There it is."

It was interesting to learn that students only used the preset titles -applied either through the auto completion of text or the browse-able list- 35% of the time, which was far lower than the co-created tags usage, which was applied 83% of the time. This difference could be in part due to the students creating the tags through class discussion, while the titles were preset in the system without discussion. It was noted that generally students would construct titles that are a more descriptive name for the animal, such as "Gray wolf" instead of "Wolf" or "Owosso Mastodon" instead of "Mastodon". Other times they used abridged animal names as titles (T-rex instead of Tyrannosaur), recorded an animal's scientific name when it was not present in the list of preset titles (the preset titles were primarily composed of more common names for the animals), or titled an animal that did not have a preset option in the system.

Since some of the methods of searching through the data in the explanation construction system relied on accurate and correctly spelled titles, the titles were coded as to whether the title accurately described the animal and whether it was spelled correctly.

All of the titles were first coded as whether they were the name of the animal (duck, t-rex, dino mastodon), a general family name for the animal (bird, dinosaur, dino), a general descriptor (predator, large), or gibberish (amm, ch), and had a high inter-rater agreement (Cohen's Kappa=0.93). The vast majority (80%) of the data was titled with the name of the animal, which we had encouraged this consistency given the activity since it increased the ease of searching for the data. Of the remaining data, 5.3% were titled with a more generic classification of the animal (fish, canine), 9% were either named after a trait or short description of the animal, and 5.7% of the titles were gibberish (amn, ch). Most of the titles were correctly spelled, with 371 (85.4% of the total data set, 90.7% of the titles that were not gibberish) of the titles with the correct spelling (100% agreement between reviewers). The objects that had incorrectly spelled titles would be less likely to show up in the search results if looking for a specific animal, though this only affected 9.7% of the titles that were not gibberish.

Next, all the titles (except the 5.7% that were gibberish) were analyzed as to whether they were accurate (either naming the correct animal, family, or the description was accurate), not accurate, or unclear. In assessing the overall accuracy of titles in the data set, 367 of the titles were accurate (84.6% of the total data set, 89.7% of the titles that were not gibberish), 22 were unclear (5.1% of the total data set, 5.4% of the titles that were not gibberish), and 20 of the titles were not accurate (4.9% of the titles that were not gibberish, 10.4% of the total data set when including gibberish titles in the inaccurate count) (Cohen's Kappa=0.86); all the non-gibberish titles that were not accurate were titled with the wrong animal name (for example, a possum was titled "owl"). Some of these incorrect names could be due to confusion over signage at the museum; in one case the student titled a data object with the name on the placard in the museum, but the exhibit had multiple animals on display and it did not match the animal they photographed.

Audio Notes

Students created 107 audio notes during data collection (98 were attached to a photograph) and the audio notes were created with different types of information in them (Table 5-2). Summarizing (54) and interpreting (42) the labels and text they found in the museum were the most frequently occurring uses, followed by relating the exhibit to the tags they had used to annotate the data (36). If an audio note was used for multiple purposes, it was coded under the corresponding multiple categories (92% agreement between raters, 100% agreement after discussion).

Audio Note Type	Frequency
Summarizing labels or text *(It says the mammoth has…)*	54
Interpreting labels or text *(I think the t-rex could eat our bus.)*	42
Relating exhibit to tags they had annotated the data with	36
Reading labels or text in the museum verbatim	21
Affect *(I like this!)*	16
Posing a question	12
Off-topic	10

Table 5-2. Types of audio notes and their frequency of occurrence.

In the example below, the data collected was labeled with “big eye holes”, “big teeth”, “few holes in skull”, “tusks”, “lungs”. Counts in the chart were added to *Relating exhibit to tags they had annotated the data with*, *Interpreting labels or text*, and *Affect*.

> Student 1: “A mastadon? I took pictures of both.” [re:mammoth]
>
> -------------
>
> Student 1: [Takes an audio note.] “This is a mastadon. It's internal cuz it's bones. They have holes a lot of holes and they have big tusks. It's huge I had to take the pic from up here. So ok. Byeee. “

The pieces of data that had an audio note were analyzed to determine the accuracy and potential usefulness of the data. Each audio note was labeled as to whether what was said was accurate (no factual statements made that were inaccurate), partially accurate (containing both factually accurate and inaccurate statements), or not accurate. Additionally, these notes were analyzed as to whether they were potentially helpful to the investigations or not helpful at all. For the audio notes, 81 (76.4%) were accurate, with 14 (13.2%) partially accurate and 11 (10.3%) inaccurate (Cohen’s Kappa=0.79). From these notes, 69 (67%) were found to be potentially helpful in answering the investigation (Cohen’s Kappa=0.89).

Because the audio notes had lower accuracy and usefulness compared to titles and tags, it may be less helpful for students to review and listen to audio when evaluating data. Additionally, while the act of taking an audio note mentioning characteristics that the student will tag the data with (done on 34% of all the audio notes) can prompt reflection on what tags they will apply to the data, these audio notes might be less useful for later review back in the classroom as the same information is repeated in the tag (though they were coded as potentially useful in this).

To examine whether audio notes had any effect on tag accuracy, the data was broken into two categories: tags with an audio note attached to the data object, and tags without an audio note attached to the data object. For aggregated data set, 90.3% of the tags without an audio note were accurate, while 90.1% of the tags with an audio note were accurate. To analyze the effect of audio notes on the accuracy of tags at the group level, I ran a Wilcoxon signed-rank test at the group level and a Mann-Whitney U test at the dataset level to see if there was any significance. The null hypothesis was that there would be no difference in the accuracy of tags irrespective of the inclusion of audio notes. While the aggregated percentages shows a slightly reduced tag accuracy with audio notes, the difference was not statistically significant at the group or data set level, indicating that the presence or absence of an audio note with a data object did not have a statistically significant effect on the accuracy of tags.

5.2 Students Using the Scaffolds of Zydeco:UseData

Research Question 2: How do students use data from the student-collected data set given the system's scaffolds?

In the process of using data, the students need to perform several steps, which are:

- Find: Students must find data they want to look at, potentially filtering through the large data set.
- Assess: Once they find potential data, students need to analyze the data, assessing if the annotations are factually accurate and useful for the explanation they have in mind, or potentially revise their idea about their explanation due to the data they viewed.
- Select: Upon finding useful data, they must make a decision on what data to select for use as evidence in their explanation.

The aforementioned tasks were analyzed to assess how the students use data. While students have to do these tasks to use data, their final product of the inquiry activity is an explanation they construct. Many factors contribute to the overall explanation beyond data usage, but the final explanations can be an indicator of whether the students were able to successfully use the data they collected and is the final area this research analyzes as a metric to see if they were able to use data in a successful manner.

In summary, we found that students were able to use data to construct scientific explanations with the data they and their peers collected in the field. Students were able to take advantage of the scaffolds around using the data set to find data and then to assess it for use in their explanation. The different data characteristics (titles, tags, and audio notes) were used for a variety of useful purposes: tags were most often compared to images as a sense making support to learn about the image and fact check the data, titles were a means to identify the animal and search upon the data set, and audio notes were a method to learn more about the animal and remember why the student collected the data later. Twenty-two of the 27 pairs selected data primarily from the collective data set and the main means students used to find data they selected was through title searches (accounting for 47% of all the data selected to use as evidence).

How students used the data is detailed below, broken down into the finding, assessing, selecting processes and then the student final explanation quality is discussed.

5.2.1 Finding Data

The different search strategies that the system's supports made available to the students in this trial included: searching by the title of the data, searching by tag(s) present in the data, searching by sub-question, looking at their own data, or looking at the entire data set. The groups employed a mix of these strategies and spent radically different amounts of time viewing data (ranging from 16 to 113 minutes, median of 48 minutes).

The sub-question filter caused confusion; several groups did not realize how the filter worked or that they had it on at all, indicating a usability issue with the interface. The filter was applied by 12 of the 27 groups, with five of them browsing for data with the filter on for more than five minutes (the longest was 15 minutes).

The filter options for tag, title, and sub-question could be applied simultaneously. This was not commonly done; only 1% of the total time students spent looking at data was using multiple filters, and only 3% of the time groups had search filters applied (excludes browsing the unfiltered data set) did they have multiple filters. The rest of the methods students used to filter and browse data used a single filter or no filters (exploring the unfiltered data set).

Applying a tag filter was the least common method of searching. While all but 2 groups tried tag searching, only 12 groups spent at least 5 minutes using this method. Only two groups spent over 25% of their time looking at data by tags (the heaviest user spending 38% of their time looking at data using tags, a little over 22 minutes).

The next most common method that students used to search for data was browsing through their own data. Because of the small amount of data collected by each group (average 16 pieces), most or all of the data could be viewed onscreen. Two groups spent the majority of their time browsing their own data, with a total of 14 groups spending 5 or more minutes each browsing through their own data. All but one group spent time reviewing their own data, with 20 of the groups spending at least two minutes reviewing their data. The class average for time spent looking at their own data was 17% of their total time looking at data.

Following searching through their own data, the most often method students used to inspect data was title searching, with 17 of the groups spending 5 or more minutes searching by the title and a class average of 22% of their time looking at data being employed this way. Seven groups used title searching as their primary means of searching for data (spending up to 42 minutes looking at data in this fashion). However, not all groups took this approach, one never using this feature and another five groups spent less than two minutes searching in this fashion.

Browsing through tiled thumbnails of the entire data set without any filters applied was the most popular method of reviewing data, with an average of 54% of their data browsing time being spent in this way, Seventeen groups used this as their primary means of perusing data, and all groups spent at least 5 minutes using this tactic. Nineteen groups spent 20 or more minutes reviewing data via this method (two groups spending over an hour). This method did not use annotations for search purposes. It was observed that some students browsed through the entire data set to see what their peers' collected and enjoyed seeing unrelated data, which could account for the high amount of time spent here.

While finding data students could choose to find similarly titled or labeled data when they are focused on a particular piece of data. The find similar data scaffold was designed to assist students in applying the search filters quickly, however only four groups even tried using this scaffold, and no group used the feature more than twice. From qualitatively analyzing the usage logs, it was noted that several pairs would browse on the unfiltered data set, focusing on data objects of interest, then manually make title searches on similarly titled data instead of using the scaffold that would automatically perform the same search.

To see where students struggled with searching through the data, we looked at when a search failed to return any results. These were coded to look at reasons for the failure and assessed to see what were the frequent causes of failure (Table 5-3).

Frequency	Reason
100	Applied filters in addition to filtering within own data
34	Misspelled
18	Searched for a tag in the title search
17	Gibberish (ie. “abm”)
14	Too many search terms, not counting filtering own data
11	Extra data in search / additional specifics (ie. “white tailed deer”)
9	Data does not exist
7	Pluralized entry that exists
3	Searched for user name

Table 5-3. Frequency and reasons for searches that returned no results.

The two most common reasons that students’ searches returned no results were:

- Having too many search filters applied, particularly the filter for looking at their own data. The issues around students filtering within their own data was a usability issue; it was unclear to some students what filters were applied and they forgot that they were searching their own data, particularly as search terms stayed applied when they left the evidence page and later returned to the evidence page.
- Misspelling titles, where students were off by a character or two in the proper spelling of the title, resulting in no data being returned.

5.2.2 Assessing Data

Once they find potential data, students need to analyze the data, determining if the annotations are factually accurate and if the data is useful for the explanation they have in mind, or to potentially revise their idea about their explanation due to the data they viewed. To assess the data, students examined the different data characteristics (titles, tags, audio note, image) on their own and peers data in order to analyze the data and see if it would be useful as evidence in their explanation.

In order to get a sense at how students used the characteristics to assess the data, we asked them, “How do you determine if the information you and your classmates collected is good?” Here, the tag annotations were commonly used. Nine of the twelve groups responding to this question indicated that they compared the tags to the associated images in order to evaluate their peers’ data.

> Student 1: “It has good tags.”
> Student 2: "It has good tags and we kinda like look at the picture and see if it’s like if its right.”
> Student 1: “If it goes with the animal.”
> Student 2: “Yep.”

It was observed during the trial that groups zoomed into the photos and looked intently at different aspects of the photographs. During the interview with student groups, two groups mentioned desiring high-resolution photographs to view of animals so they could really zoom and examine the animal's features. The system did not support this in that zoom did not stay at the desired level (zoom was capped at 2.5x magnification, though students could pinch the image to temporarily make it zoom in further), and that the images pixelated at this high zoom:

> Student: "In this case if you want to get a closer up look the zoom doesn't stay so you have to continue holding your fingers you would have to be close and the quality of the picture gets ruined so its very hard to look at the tiny details cause like these are tiny animals you have to look at tiny details and you can't really tell."

After comparing tags to images, discussion between partners (4 groups) and using tags to find traits that were similar to their own animal (3 groups) were the most common approaches toward determining how good their classmates' data was.

> Student 1: "You should check the features of the animal the same traits that one animal has over the other…"
>
> Student 2: "You use the picture. You click that you're gonna use it, look for the traits that the picture has. If it has the same trait as your animal then that those are the right traits."

To further look at the utility of tags (called labels to the students), and any effect they had on analyzing the data, we asked the students, "Did you find the labels describing individual animals helpful? Why?" All but one of the 14 groups responding to this question reported the tags were helpful, primarily to see similarities in their animal (6 groups) or get more information than would be present in the image alone (5 groups). The group who didn't find the labels helpful felt this way, "*because, most of the stuff I would see by yourself so you don't really need labels*."

> Student: "If there were no labels I would think what's the picture like what's the like what's all the things about it cuz I don't know anything about if there's no labels. Like turn that picture to the animal. Ok, then I didn't know what it was, there's no label on it so I don't know what it is still, and I don't know any facts about it or nothing."

This implies that the tagged data assisted some students in evaluating their evidence, with some students believing they could not evaluate the data without tags being present.

However, audio notes were not mentioned in how students determined good. In looking at the usage logs, we discovered that while all but one group listened to audio notes, there was a high variance in audio playback. During the trial we had noted many students seemed to enjoy listening to audio notes other students collected, often wanting to identify who made the note. Additionally, there were two pieces of off-topic audio data that several students found entertaining (one being a photo of a duck titled "Aflack" with an audio note that said Aflack, mimicking a TV commercial). Several students repeatedly played these audio notes; one group listened to the Aflack audio 85 times. This caused a disruption to students nearby at one point, and the group was instructed to stop playing it by the teacher.

The breakdown of how many times each pair listened to audio notes can be viewed in Figure 5-2, which also separates out the two off-topic audio notes. Off-topic playback only became a disturbance in the K class (notably the class where the students recorded both off-topic audio notes) as once students initially played the off-topic audio clip, it encouraged other students to do the same.

When students were asked, "Did you find audio playback on this page useful? Why?," several students found it useful, one student mentioned "because it explained details" and another group liked their own audio notes "because it kinda refreshed me why I kind of took it, the picture". However, one group mentioned not finding them useful because they say the same thing as the tags, showing a mixed impression of the utility of this annotation.

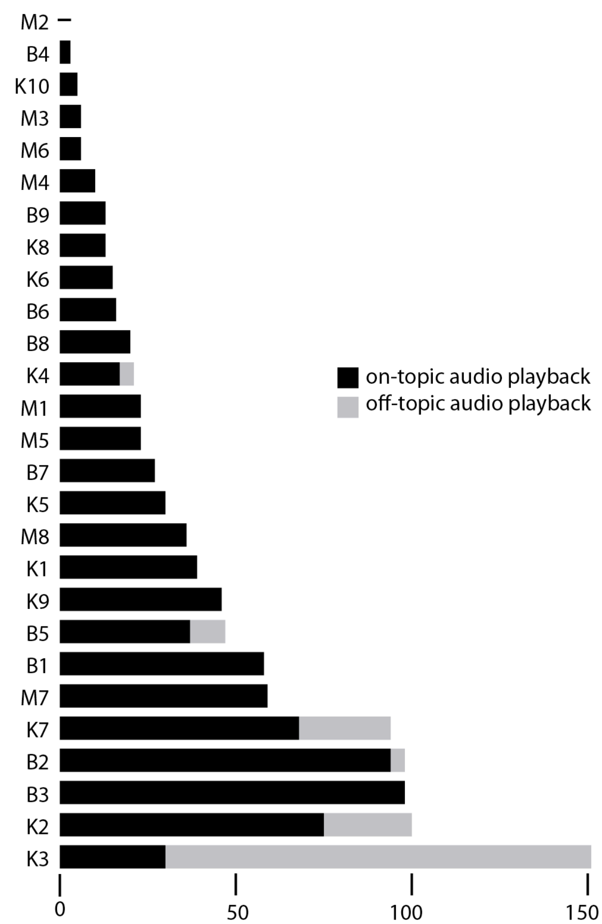

Figure 5-1. Number of times each group listened to audio, separating out repetitive off-topic playbacks.

Students also had the option to look up the supplemental, curated information on any animal that at least 5 pieces of data were collected on. The page containing the supplemental information (labeled "*More Information*") was discovered by 19 of the groups (70%). Eight groups reviewed every animal they could receive extra info on, but never spent more than 20 seconds on a particular animal (showing less serious perusal), many of these groups just tapping on each piece of data to see what was displayed on each page. Only six groups (22%) spent more than 20 seconds on a particular extra info page (to exclude those with less serious perusal). These six groups looked at one to three different animals and would typically spend between one to five minutes on a page. For each group, at least one of the animals they reviewed was included in their explanation. Three groups made edits that appeared to be directly related to the information they had just read. In the following example, a group that had been perusing and listening to audio about 'Turtle" data then read about Turtles in the supplemental information. They then went on to modify their reasoning, incorporating information they had read:

Initial Reasoning:

The turtle's shell helps it survive because it helps it protect itself from predators that may attack it by giving the turtles a layer of protection.

Excerpt from Supplemental Information:

A turtle is a type of reptile that may live on sand or in the water. Turtles have shells and are the slowest moving reptiles. *Most turtles that live in the water have webbed feet like a duck.* Tortoises have thick legs like elephants and sea turtles have flippers.

Modified Reasoning:

The turtle's shell helps it survive because it helps it protect itself from predators that may attack it by giving the turtles a layer of protection. *A turtle also has webbed feet that help it swim faster from predators or swim faster to get let's say a fish.*

Several groups mentioned that the supplemental information page was not useful as the animal they needed was not present (curated data was available for 33 animals while students collected on 153 different animals), and there were requests to either have additional data present on all the animals, or be able to Google for more information (which students had done in the pre-activities). Though the students from the conversation below had discovered the supplemental data two days previously and explored several different pages of information, they still desired to have access to more information:

Interviewer: "What is one thing you would change about the system that you noticed today?"

Student: "If you wanted to search an animal like you have to pop up some links and maybe search on Google or something so you could get more results (…) if you can't find anything that somebody found out about your animal then you wouldn't be stuck without evidence."

5.2.3 Selecting Data

To understand how students selected the data they used in their explanation and how they ultimately chose each piece of evidence, we examined how they discovered every piece of evidence. Some students added the same piece of evidence twice, so only the first occurrence of finding the evidence was added to look at students' method for finding each unique piece of data.

In order to understand what they found important about the pieces of evidence included in their explanation, we asked them (1) how they might choose data they would want to use and, later, (2) how they chose the data that they did use in their investigation.

How Each Piece of Evidence Was Discovered

The 187 unique pieces of data students chose to use in their explanation were categorized by which search or browsing method they used to find the data (breakdown of the methods by group in Figure 5-3).

Only a few groups showed any preference toward evaluating their own data, and only five groups used more than 50% of their own data in their explanations while nine groups never used any of their personally collected data. While most groups focused on the collective data set for the majority of the time, these five groups spent a substantial amount of time viewing their personally collected data and had heavy utilization of their own data, showing it is valuable to support the filtering between personal and peers' data.

Finding data by searching titles was the most common means of adding data to the explanations, which groups did for 47% of the data they added to their explanations (88 pieces of data were used that were discovered by title searching). Though students spent more time browsing the unfiltered full data set, the groups only used 57 (30.5% of all data used) pieces of data from this search method; this can be partially the result from the combination of using unfiltered browsing students performed before following up on a lead by applying search filters to the data set, as well as the off task behavior that was noted when some students are browsing the unfiltered data set. As with time spent searching, tags were not used to discover much data that students used in their explanations, with only 12 (6.4% of all data used) pieces selected in this fashion.

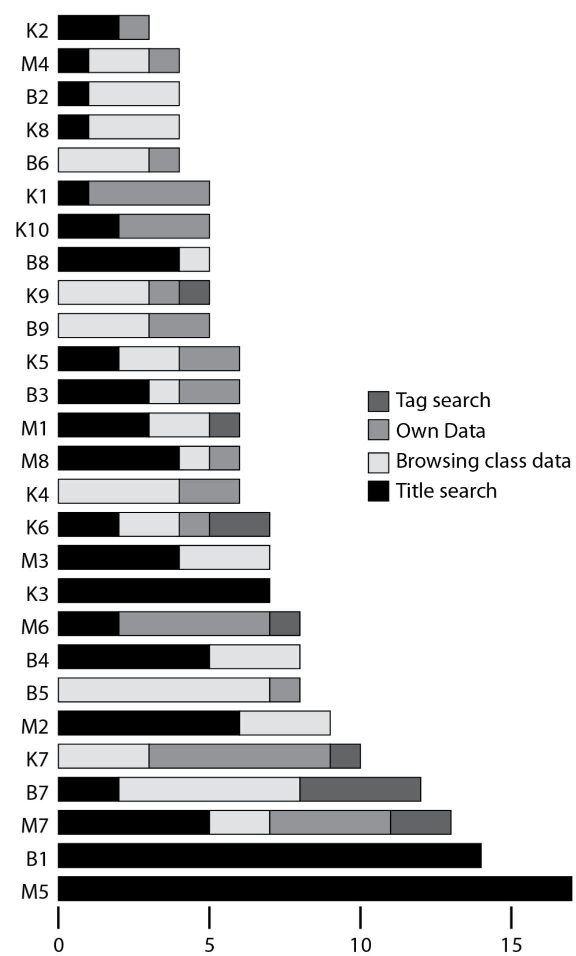

Figure 5-2. Search filters students had applied when they found data they used in their explanation.

When we looked at the characteristics of how students searched for data when they then chose to use a piece of data, several trends emerged. Whenever students found a piece of data to use with a tag applied, the tag was always a trait their favorite animal had. This is an ideal search behavior for discovering related animals, as the number of shared traits in common is a strong predictor for animal relatedness.

For title searching, every search that resulted in a student adding data to their explanation began by looking up the animal name except for five cases. Three separate groups made queries on a more generic animal type (animal, fish, birds) and two different groups only typed in part of the animal name to retrieve the result (Q to find Quetzalcoatlus and gir for Giraffe). The system evaluated the search term looking to see if it matched the prefix of any word in the title, which in this case helped the student avoid spelling out Quetzalcoatlus (and avoid misspelling it in the process).

Student Perceptions on Selecting Data

In order to understand what students found important about the pieces of evidence included in their explanation, we began by asking them, "How do you pick out data you would want to use?" While comparing pictures to tags was the main mode of judgment for students in determining what "good" data was (5.2.2), the criterion for what data they wanted to select as evidence in their explanation shifted. Though comparing pictures to tags was still a notable strategy, the more prevalent strategy was comparing the tags to the traits of their chosen animal. Seven of the fifteen pairs interviewed explicitly named this as their strategy.

> Student: Points to the labels under an image and explains that, "Basically what we trying to find is similar traits like how many legs and how many toes it gots like the salamander (the pair's assigned animal)."
>
> Interviewer: "So you're matching the labels?"
>
> Student: "Yea."

We expected that if students were being reflective about their work, then these methods of comparing the labels to the image, or to their chosen animal, would be their main strategies. While image quality and asking an expert (such as a teacher or researcher) received mention as ways to determine good data and pick out what data to use, reviewing audio was never mentioned as a potential tool for evaluating the data. This may be due to the lower usefulness of the audio notes (due to providing redundant data or being off-topic) or the students' interpretation of the question, but it could also indicate that audio notes are not seen as a useful means to interpret the data in this context.

An interesting phenomenon occurred when the students were asked, "How did you choose the specific pieces of data that you used in your explanation?" Three groups reported that picture quality became an important factor, even potentially surpassing the quality of the annotations on the data.

> Student: "Well, its really one thing like this salamander picture it's actually a good picture… you can actually get better like you can get a bit more detail. Some other animals may be more closely related but still this gets back to the quality of the pictures."

These students were willing to choose a less closely related animal because of higher image quality, which suggests that it may be helpful to have a method for the student to swap the image that is associated with the annotations if a goal is to encourage using data with the best annotations.

Trend on data selected for use in explanations

Trend noted that some students gravitated towards using data they collected in the field. Two of the groups used at least one piece of their own data for every animal they had in their explanation. However, when looking at whether students have collected data on each animal they used in their explanation (whether or not they used their own data on each), seven of the pairs chose all animals that their group collected data on in the field; this suggests that the data students collect on in the museum can influence the type of data they will use in their claim.

5.2.4 Evaluating Student Explanations

To gain another indicator as to whether students were able to collect and use the data successfully in this inquiry activity, I took a look at the performance of the learners and evaluate the explanations they constructed.

Determining the Quality of the Students' Explanations

Previous research has shown that students typically can articulate a claim and use some supporting evidence (with difficulty), but greatly struggle with providing reasoning that connects evidence to their claim (Berland & McNeill, 2009). We were working with students who have never engaged in project-based science inquiry investigations in school and this was their first time using the Claims, Evidence, and Reasoning framework, and the students were working in a difficult context using large amounts of peer collected data to construct these explanations.

A rubric was created for each investigation in Zydeco (Table 5-4 and Table 5-5) modeled off McNeill's generalized claim—evidence—reasoning grading format (McNeill & Krajcik, 2011). In the first investigation students must construct an explanation ranking how three animals are related to their favorite animal. The second investigation asked students to explain why or why not two of the related animals they chose in the first investigation would survive in Michigan.

Claim	Evidence	Reasoning
0: Nothing/ not a claim	0: No evidence/ evidence not related to claim	0: Does not provide reasoning/ reasoning does not connect the evidence to the claim/ does not make sense
1: Partial claim/ ranks fewer than three animals	1: Some ranked animals without evidence	1: Provides reasoning that links the claim and the evidence by repeating the evidence
2: Complete claim- includes three ranked animals	2: One evidence for each ranked animal	2: Provides reasoning that explains why the evidence (traits) supports the claim
	3: Multiple evidence for some ranked animals	3: Provides reasoning that includes why the animals share traits through a scientific principle (i.e. shared ancestors)
	4: Multiple evidence for all animals	

Table 5-4. Grading rubric for investigation 1.

Claim	Evidence	Reasoning
0: Nothing/ not a claim	0: No evidence	0: Does not provide reasoning/ reasoning does not connect the evidence to the claim/ does not make sense
1: Makes a claim related to the question, but it is incomplete or does not directly answer the question	1: Does not provide relevant evidence	1: Provides reasoning that links the claim and the evidence by repeating the evidence
2: Makes a complete claim related to the question	2: One relevant evidence	2: Provides reasoning that explains why the evidence (traits) supports the claim, but does not address all pieces of evidence
	3: Multiple relevant evidence	3: Provides reasoning that explains how the evidence supports the claim and addresses all the pieces of evidence

Table 5-5. Grading rubric for Investigation 2.

For each investigation, two reviewers independently graded 20% of the explanations and checked inter-rater reliability by percent agreements. The inter-rater agreement was 100% for claim and 100% for evidence on each explanation. Because reasoning was harder to judge, the reviewers each graded 40% of the explanations independently, having 75% agreement on the first investigation and 92% agreement on the second investigation. The reviewers discussed how they evaluated the reasoning on the explanations where they disagreed until they came to 100% agreement. After this, one reviewer graded the remaining explanations.

Investigation 1: How is My Animal Related to Other Animals?

In the first investigation, "*Determine how three other animals are related to your animal,*" the average scores for each claim, evidence, and reasoning for each class are presented in Figure 5-4. When the claim, evidence, reasoning grades are averaged across all the classes, the students had a 1.70 for claims, 2.30 for evidence, 1.73 for reasoning.

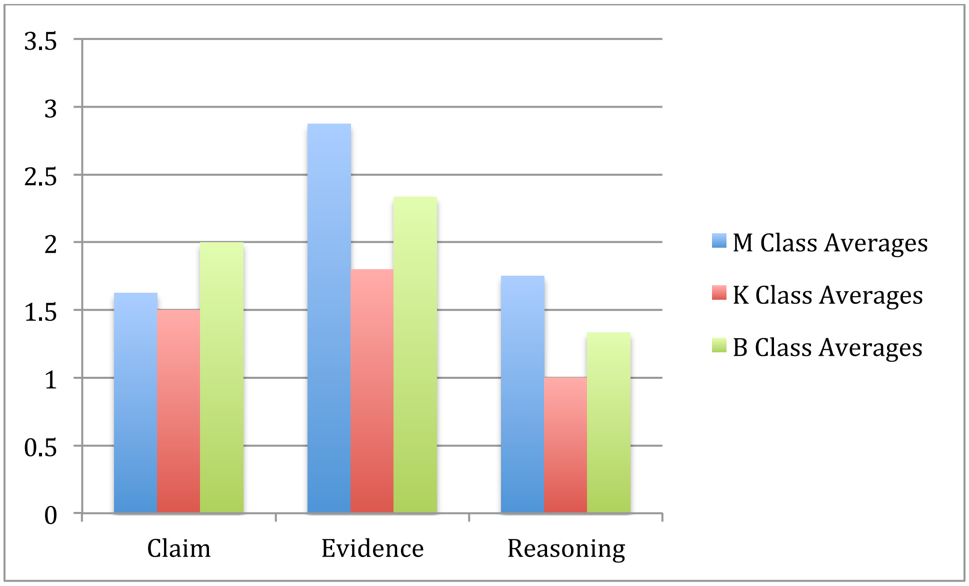

Figure 5-3. Investigation 1 results for each explanation component by class.

The results overall were positive, with 23 of the 27 pairs using evidence for each animal, and 24 of the 27 pairs providing reasoning that at least repeats the evidence on animal traits. Also, five pairs failed to make a complete claim, not ranking three animals.

It was noted that 5 of the 9 pairs who had issues with the claim, evidence, or reasoning were in K's class, which had numerous behavioral problems throughout the experiment.

To provide further understanding of the explanations students constructed, a representative example of a great explanation as well as an average explanation that students constructed is given below:

Representative Great Explanation

A representative example of a great explanation created by a student in this trial can be seen in Figure 5-5, which had a score of 2 on the claim grading due to having a complete claim, a 4 score on evidence for multiple evidence in all the categories, and a 3 on reasoning for linking the traits of the animals to the idea of internal and external traits in common as a means of relatedness, as well as showing understanding that the coyote and wolf are in the same family.

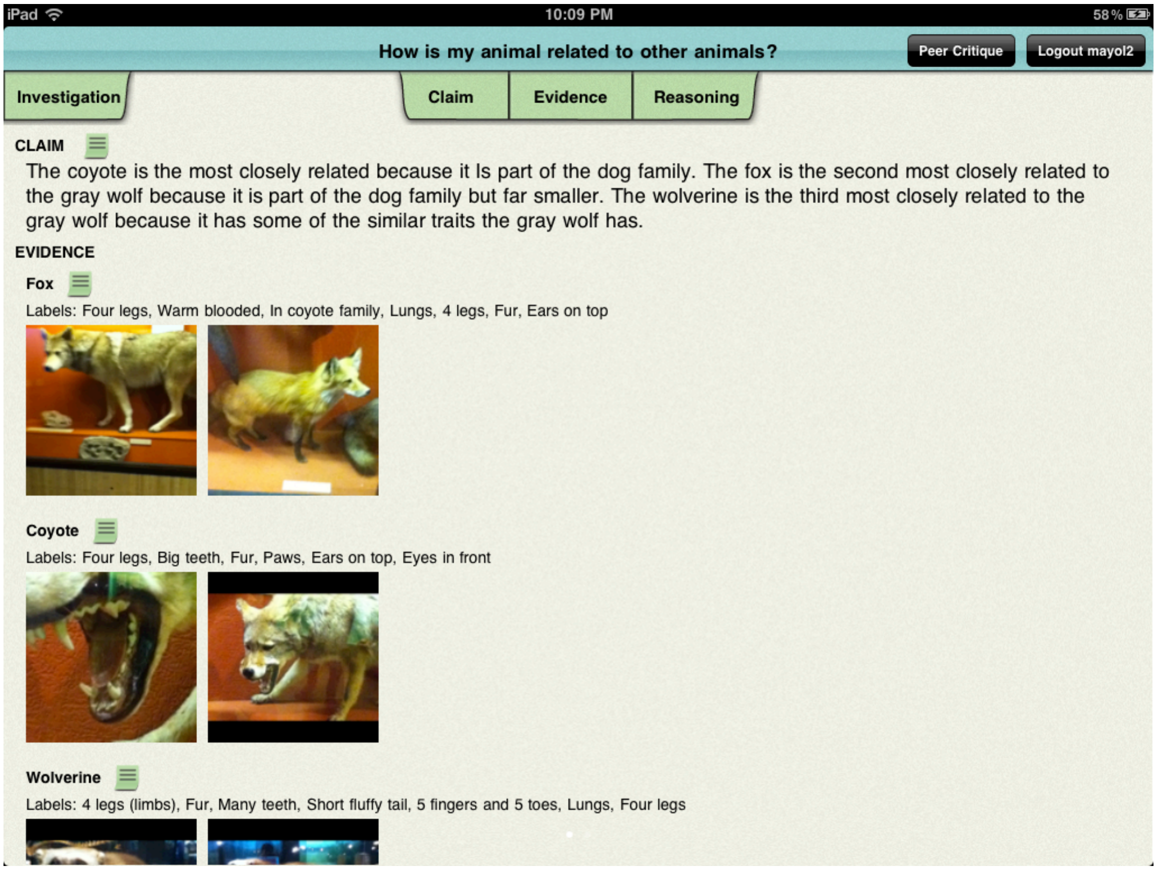

Figure 5-4. Representative Great Explanation generated from this trial.

> Reasoning (not visible in the screenshot): "The coyote is most related to the Gray wolf because it has most of the external traits and internal traits and because it is in the wolf family and it can survive in winter just like the gray wolf can and they are hunters. The fox is the second most related to the gra wolf because the fox is warm blooded and has many of the same ex and internal traits and has many of the same habits. The wolverine is the third most related to the gray wo because we don't really have any background in the wolverine and we know a lot aotu the coyote and gray wolf and the wolverine doesn't have many of the same traits as the gray wolf."

Representative Average Explanation

Representative example of an average explanation can be seen in Figure 5-6, which had a score of 2 for the claim due to having a complete claim, a 3 score on evidence for multiple evidence in some categories, and 1 on reasoning for just repeating the evidence without connecting it to the claim.

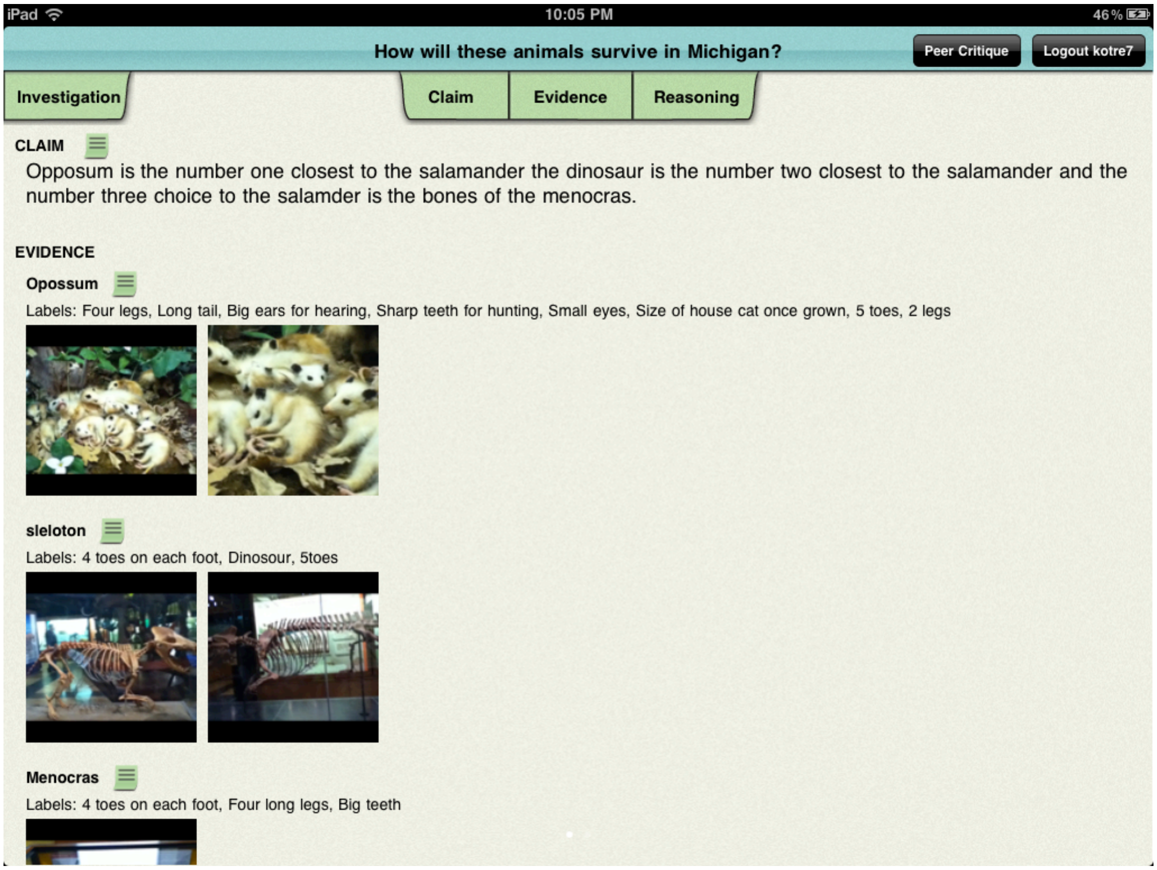

Figure 5-5. Representative Average Explanation generated during this trial.

> Reasoning (not visible in the screenshot): "These animals have 3 things in common 1.long bodies 2.long tails and 3.four toes the opossum is the closest to the salamander because of the long tail long body and the sharp teeth for hunting prey or other animals"

Investigation 2: How Will These Animals Survive in Michigan?

To provide a breakdown of the overall results for investigation two, "*Describe how these animals could survive in Michigan*," we averaged both explanations together (if they had two), and then constructed the class average from that, as seen in Figure 5-7. When the claim, evidence, and reasoning grades are averaged across all the classes, the students had a 1.50 for claims, 2.18 for evidence, 1.43 for reasoning.

However, lack of time to complete the assignment was detrimental to student performance; if we judge completion by students reaching the stage where they asked for a peer critique, only 14 of the 27 pairs completed this investigation. The students who did not ask for peer critique had explanations in various states of partial completion.

A look at only the scores of students who received a peer critique can be seen in Figure 5-8. When the claim, evidence, reasoning grades are averaged across students who received a peer critique, the students had a 1.75 for claims, 2.22 for evidence, 1.64 for reasoning.

These results excluding students are biased towards the students able to finish (possibly better students than the others). It was noted that 7 of the 8 of the pairs in M's class were able to finish, while only 4 of the 10 in K's class and 3 of the 9 in B's class finished. From teacher interviews on student performance, it was noted that K's class had numerous behavior issues, a reoccurring issue for those students, and B's class was going at a slower pace, taking more time to complete the activity, "seemingly" taking tasks seriously.

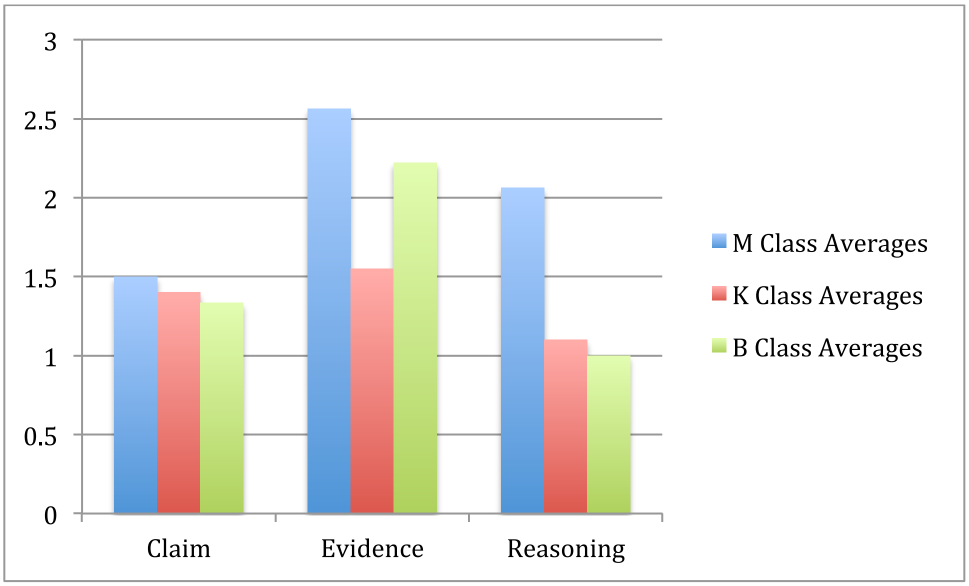

Figure 5-6. Investigation 2 results for each explanation component by class (with all pairs).

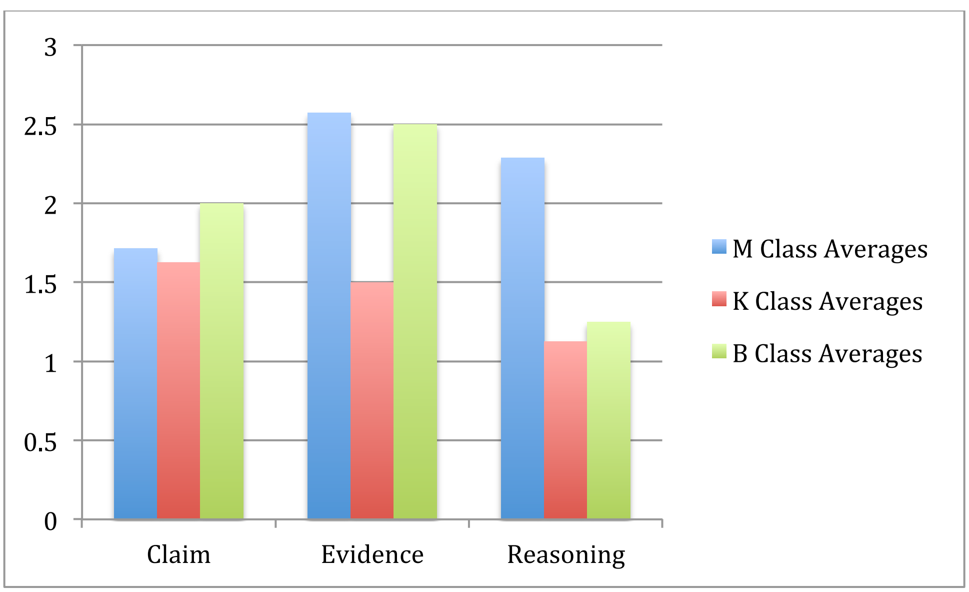

Figure 5-7. Investigation 2 results for each explanation component by class (pairs that received a peer critique).

Overall Explanation Performance

Overall, most students in the first investigation made explanations using multiple pieces of evidence with reasoning that spoke about the evidence, a task difficult for students.

For the second investigation, lack of time prevented many students from finishing. This could be due to the complexities of looking at the data and making an explanation, or also related to the students attitudes and behavior issues. However, students who did finish were able to incorporate multiple pieces of evidence and provide reasoning that related to the evidence. The literature has shown that students struggle with creating reasoning, even with scaffolding, which was similar to what we saw with the explanations made in this study (McNeill & Krajcik, 2008; McNeill & Krajcik, 2011).

During post-class interviews with the teacher, she said she was impressed by students' performance, stating that students overall had a much better work output and put in more effort than they typically do in her class. Though this may be in part from the novelty effect of the technology, it is encouraging that students are performing better than usual at this challenging inquiry task.

5.3 Summary of Zydeco Scaffold Usage

This chapter has covered how students collected and used data, given the scaffolds in the Zydeco:CollectData and Zydeco:UseData tools, respectively, and answered our two research questions.

In order to address the overarching research question on how the scaffolds were able to mitigate the challenges around data collection and annotation, I summarize how students used each scaffold by the tool (Zydeco:CollectData and Zydeco:UseData). In the following chapter, I discuss how the scaffolds were able to mitigate the challenges around collecting and using data and areas for improving the scaffolds.

Zydeco:CollectData Scaffolds

A summary of the results noted around each scaffold can be seen in Table 5-6, which comes from the full set of results in section 5.1.

Scaffold		Collecting Data
1) Decompose the problem into more manageable parts		Sub-questions directed students approach to data collection on different artifacts and encouraged reflection on what they were about to collect.
2) Enable data objects to be annotated in multiple ways	Audio Notes	25% of all the data had an audio note, 98% of the time the audio was associated with an image. 76% audio notes were accurate, 67% potentially useful. 34% of the audio notes content overlapped with the tags associated on the data.
	Titles	85% accurate, 35% titles came from preset list. 80% of titles were animal names. 9% of titles were descriptions of animal or its traits 85.4% of the titles correctly spelled.
	Tags	Average of 2.6 tags per data object. 90% accurate, 99.9% accurate tags were potentially useful.
3) Integrate co-created tags that reflect expert discussion		83% tags applied were from the co-created tags. Statistically significant increase in accuracy when students use the co-created tags versus new ones they create out of the classroom.
4) Guide collection step-wise		Students exhibited thoughtful behavior during each collection stage, reflecting and articulating appropriate information on the data at each stage.

Table 5-6. How students collected and annotated data using the scaffolds in Zydeco:CollectData.

One pair out of the 27 had difficulties using the app Zydeco:CollectData tool to record accurate and useful annotations. This breakdown occurred due to the group ignoring the co-created tags and primarily creating their own, constructing data artifacts with annotations that were less accurate and more difficult to determine by the researcher if they were useful to the investigation due to their ambiguous annotations, such as “bio”.

Additionally, audio notes had a large amount of off-topic information and overlapping information on tags that made reviewing them less helpful for students. The audio notes could still have a benefit in these cases by helping students refine their thoughts as they take an audio note, though there was no significant difference in tag accuracy when the data did or did not have an audio note.

Zydeco:UseData Scaffolds

Table 5-7 details how students used a variety of approaches to find, assess data, and select data for use in constructing an explanation (full results in section 5.2).

Scaffold		Finding Data	Assessing Data	Selecting Data
1) Enable data assessment and filtering by data characteristics	Titles	Most used method of filtering the data, 17 pairs spent 5 minutes or more title filtering.	Used to identify the animal.	Title search was most common method used to find data selected as evidence (47%).
	Tags	25 pairs applied tag filters but only 12 pairs spent at least 5 minutes tag filtering.	Main methods of assessing the data; used to learn about how the data can be used and fact-check the data by comparing tags to the image.	Main method to assess and select data was comparing the traits in the tags to their animal to determine relatedness. While 25 groups used tags to explore the data set, only 7 pairs found data in this manner (pairs did this by searching on traits in common with their animal to find related animals).
	Audio	N/A	Helped some students remember information and learn more, while others found it had redundant information and was off-topic.	Never mentioned as a means to determine data to select.
	Images	N/A	Examining and zooming into the image was done to identify features of the animal - used in conjunction with tags.	Three pairs interviewed considered image quality most important when selecting data to use.
	Sub-Questions	Least common filtering method; noted was applied in error by several groups. Applied by 12 groups, 5 groups used for more than 5 minutes.	Not noted as being used.	No data used as evidence was found by sub-question filtering.
	Own Data	Second most common method of filtering the data, 14 pairs spent more than 5 minutes filtering to view their own data.	N/A	Only 5 pairs selected more than 50% of their own data, the rest primarily using the collective data set.
2) Enable exploratory search		Students spent most of their time browsing the unfiltered data set. Some students browsed the unfiltered data to find data of interest, then used filters to investigate further. Exploratory tag searches prevented searches that return no results, which happened during title searching.	N/A	Second most common method of selecting data was via exploratory search through the unfiltered data set (31% of all the data used).
3) Expedite searching for similar data		Barely used; 4 students discovered, never activated more than twice.	N/A	No data students selected was found using this scaffold.
4) Supply supplemental curated data		N/A	Discovered by 18 pairs, used to review data on animals by 6 pairs, 3 pairs made relevant edits after reviewing this data.	N/A

Table 5-7. How students found, assessed, and selected data using the scaffolds in Zydeco:UseData.

However, several issues were encountered around a few of the scaffolds (in addition to audio notes being of mixed usefulness, noted previously above):

- Difficulties using filters: some students had issues spelling title searches, confusion finding data when applying multiple filters and confusion understanding the search filters in effect.
- Searching for similar data not being attempted and not used successfully by the two groups who did.
- Supplemental curated data not having the content that many of the groups needed.

In the next chapter we synthesize these results to discuss the implications of how students used the scaffolds, describing the benefits and areas for improvement.

CHAPTER 6

Discussion

This research investigated how to scaffold students through the processes of collecting annotated multimedia data using mobile devices when out of the classroom and how to scaffold students in using their personal and their peers' data to construct evidence-based scientific explanations back in the classroom. This study focused on how students used the software scaffolds around collecting and using data within two integrated tools, Zydeco:CollectData and Zydeco:UseData (Figure 6-1).

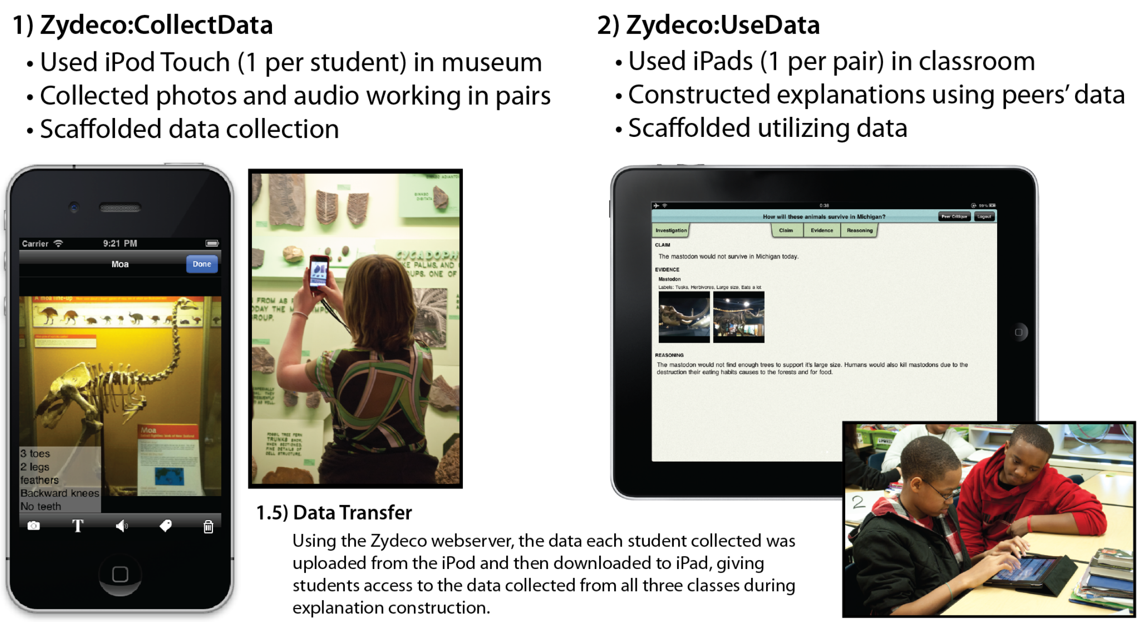

Figure 6-1. Overview of the two integrated tools in Zydeco and how students used the system.

In order to address this research goal, we sought to answer the following overarching research question:

> **Overarching Research Question:** What scaffolds can mitigate the challenges for students inherent in collecting multimedia data as well as using a large, student-collected, multimedia data set to construct explanations?

To learn this, in Chapter 5 students' usage of Zydeco was evaluated to determine:

> **Research Question 1:** How do students collect and annotate multimedia data given the system's scaffolds?
>
> **Research Question 2:** How do students use data from the student-collected data set given the system's scaffolds?

The end of Chapter 5 summarized how students used the scaffolds in both tools. This chapter synthesizes that information to discuss how the scaffolds were able to mitigate the challenges around collecting and using data, and areas for improving the scaffolds to better mitigate the challenges. Following this is a discussion of general factors that seem to have improved the success of the integrated tool and a summary of guidelines for other designers seeking to support similar data collection and use tasks.

Discussion Overview

The Zydeco tools integrated several complex inquiry tasks into a multi-step process, scaffolding students through their investigation and making the daunting task of collecting and using data manageable for students. The benefits of the scaffolding in the Zydeco tools were reflected in the students final products, which were explanations supported with evidence pulled from a large, multimedia data set that the students and their peers created through their data collection.

In particular, the system scaffolded inquiry activities that fell under the Next Generation Science Standards' practices of *carrying out investigations*, *analyzing and interpreting data*, and *constructing scientific explanation* (National Research Council, 2012). This research found that:

1) The Zydeco:CollectData system enabled the collection of hundreds of pieces of annotated data, with the vast majority of the annotations applied to the data being factually accurate and useful for answering the investigation.

2) The Zydeco:UseData system facilitated using this data set to find, assess, and select data to construct scientific explanations.

As described in this chapter in detail, by and large all the scaffolds were beneficial – and many had areas where they needed to be improved to better address the challenges students faced. An overview of the key lessons learned from evaluating how students used the Zydeco:CollectData system are listed below in Table 6-1 and those learned from the Zydeco:UseData system can be found in Table 6-2.

Challenges	Zydeco Scaffolds	Benefits	Areas for Improvements
Students have difficulty planning and monitoring their inquiry plans (Krajcik & Blumenfeld, 2006). They struggle determining what data to collect to answer complex inquiry questions (Berland & McNeill, 2009; Griffin, 1998).	Decompose problem into more manageable parts.	Sub-questions directed students approach to data collection.	Support adding data to multiple questions as some students desired data to be classified under multiple questions.
Students may be unable find data later if not properly organized (Vavoula et al., 2009). Students lack the knowledge experts have engaging in inquiry and may not understand or reflect on how data needs to be annotated to be useful later (Metz, 2000). Students may have difficulty connecting their past knowledge to the artifacts and phenomena they observe (Roschelle, 1995). Students can be cognitively overwhelmed outside the classroom (Balling & Falk, 1980); non-germane but difficult tasks such as the mechanics of data entry can hinder their ability to collect data (Tossell et al., 2010).	Enable data objects to be annotated in multiple ways.	Multiple linked annotations enabled students to record interlinked ideas as they carry out their investigation.	
	Integrate co-created tags that reflect expert discussion.	The co-created tags assists students in reflecting on and articulating how the data should be suitably described during collection.	Because the system does not force students to use the collaborative language, some students may avoid these scaffolds (occurred in 1 of the 27 groups). Have a discussion to co-create titles to increase their usage and data consistency (potential titles were applied to 35% of the data).
	Guide data collection step-wise.	The step-wise process requires students to focus on each stage of the data collection process.	Provide additional scaffolding around audio note collection promoting notes that are useful.

Table 6-1. Summary of the benefits and areas of improvements noted on the Zydeco:CollectData tool in answering research question 1.

<table>
<tr><th>Challenges</th><th colspan="2">Zydeco Scaffolds</th><th>Benefits</th><th>Areas for Improvement</th></tr>
<tr><td rowspan="9">Students can have organizational issues dealing with large amounts of multimedia data (Vavoula et al., 2009).

Students have difficulty searching for information (Druin et al., 2009) and have difficult constructing appropriate search syntax (Bilal, 2001).

Children struggle with using appropriate and sufficient evidence (McNeill & Krajcik, 2011). Students have the greatest difficulty dealing to deal with large, student collected data sets as this data can be inaccurate and/or not be useful for constructing a scientific explanation (Berland & McNeill, 2009).</td><td rowspan="6">Enable assessment and filtering by data characteristics.</td><td>Titles</td><td rowspan="6">Data characteristics were used for a variety of purposes to find and assess (sense making, identification of artifact, getting more information).</td><td rowspan="6">Students need additional support constructing search terms and understanding how the data is filtered, as some students struggled spelling terms and understanding how to search.</td></tr>
<tr><td>Tags</td></tr>
<tr><td>Audio</td></tr>
<tr><td>Images</td></tr>
<tr><td>Sub-questions</td></tr>
<tr><td>Own data</td></tr>
<tr><td colspan="2">Enable exploratory search.</td><td>The list of all the tags that can be used to apply filters can assist students in searching on the data set and avoid searches that return no results.

Being able to browse through condensed views of the information in the data set assisted students unable or unsure of what search filters to apply to look at the data set.</td><td>Expand exploratory search to list out and be able to filter all textual data characteristics.</td></tr>
<tr><td colspan="2">Expedite searching for similar data.</td><td>Not used.</td><td>Improve scaffold visibility.</td></tr>
<tr><td colspan="2">Supply supplemental curated information.</td><td>The supplemental information helps some students interpret and use data in their explanation.</td><td>Supplemental data needs to encompass the full variety students can collect on (may need access to Internet resources).</td></tr>
</table>

Table 6-2. Summary of the benefits and areas of improvements noted on the Zydeco:UseData tool in answering research question 2.

Sections 6.1 and 6.2 of this chapter, expanding on Tables 6-1 and 6-2, discuss the benefits and areas for improvement for the scaffolds used in the two tools in detail. Overall factors of the integrated tool design may have contributed to the overall utility of Zydeco are discussed in 6.3. This chapter concludes with a summary of design recommendations for designers based on the outcomes of this work.

6.1 Zydeco:CollectData Scaffolds

Research Question 1: How do students collect and annotate multimedia data given the system's scaffolds?

As outlined in table 6-1 above, the scaffolds of Zydeco:CollectData were able to support students in numerous ways as they collected and annotated data outside the classroom. In answer to RQ1, we evaluated 27 pairs of students using Zydeco:CollectData for an hour in a history museum during which the students collected 434 annotated data objects. As discussed in our results, the scaffolds were overall successful at enabling students to collect multimedia data with accurate and useful annotations:

- The average accuracy of tags per group was 88% and all but 1 accurate tag was potentially useful to the investigation.
- The average accuracy of titles per group was 83%, and an average of 85% of titles were correctly spelled per group.
- The average accuracy of audio notes was 70% per group, and 64% of these audio notes per group were potentially useful to the investigation.

Providing useful annotations to data gathered outside the classroom is a difficult task for newcomers to inquiry (refer to Table 6-1 above) and these high levels of accuracy and usefulness imply that the scaffolds around data collection were successful at mitigating the challenges surrounding carrying out an investigation outside the classroom.

In the following sections each scaffold used in Zydeco:CollectData is discussed, stating why the scaffold appeared to have utility and identifying areas of improvement to the scaffolding design.

6.1.1 Decompose the Problem Into More Manageable Parts

Using sub-questions is a technique adapted from Project-Based Science (PBS) that has been successfully used to decompose an inquiry question into manageable parts that students can answer (Krajcik & Blumenfeld, 2006). This technique was adapted to the handheld platform by incorporating it into the first stage of the data collection process. While not all inquiry investigations may have sub-questions, particularly if the students are investigating a simple question for their main investigation, many investigations have enough breadth to be broken down into smaller sub-sets. In this study the driving question was decomposed into two sub-questions regarding whether traits observed were similar to their animal's internal or external traits. These sub-questions provided a starting point to direct and guide learners on artifacts to look for during data collection. Students were required to review and select a sub-question in the system to then collect data that can work toward answering that sub-question.

Benefits

Students were carrying out open-ended investigations with Zydeco and were expected to determine the data they would collect to answer their investigation while out of the classroom. The students were using the sub-questions as a means of determining what data they needed to collect and how artifacts they found related to their investigation. Zydeco supported this by requiring students to review and then select a question to collect data toward before moving on to capturing the data.

By requiring students to look at their questions and choose one question to collect the data under, students can be encouraged to reflect and focus on the purpose of the data they are collecting and how it fits into their investigation. Zydeco:CollectData listed these sub-questions on the collect data screen. Audio recordings taken while students were collecting data in the museum revealed a number questions and discussions between students to determine what sub-question the artifacts they saw would fit under. For instance, in asking, "Look, get the piranha. Is it internal or external?," the sub questions of *What are similar internal traits?* and *What are similar external traits?* influenced how students regarded the phenomena they were viewing, having entered the discussion between partners.

Ultimately, for investigations that can be broken down into sub-questions, these sub-questions can initiate how students choose to begin collecting data to promote student-choice while guiding the learner.

Areas for Improvement

Outside the classroom students saw artifacts that they wanted to collect which applied to two sub-questions (e.g. an Eagle, capturing it under both internal and external traits sub-questions). In the system, students were limited to collecting a piece of data under only one sub-question. This forces students to either 1) collect two pieces of data, one under each sub-question and often taking the same picture twice, or 2) collect the data under one question but apply tags that related to both questions.

However, collecting under one question has an additional drawback. To provide relevant information to the students, the co-created tags that were incorporated in the system was different for each sub-question, giving each sub-question a different set of suggested tags (e.g., collecting under the internal traits sub-question would display only internal trait tags and not any external trait tags). In situations where students wanted to collect data that addressed multiple sub-questions, the system did not allow them to view all of the potentially relevant annotations. In cases where students collected the data that applied to both sub-questions under a single sub-question, the student would only have access to part of the co-created tags.

Ultimately, not allowing data to be collected under multiple sub-questions led to students having to spend additional time and energy to record all of an artifact's relevant information. Both time and energy are a limited resource, thus potentially detracting students from being able to collect more useful data.

For these reasons, scaffolds around decomposing the task into more manageable parts using sub-questions should allow students to collect data under as few or many sub-questions as desired.

6.1.2 Enable Data Objects to be Annotated in Multiple Ways

Some previous systems that enabled data collection in field did not allow data objects to have a variety of annotations (Rogers et al., 2004; Vavoula et al., 2009). This caused difficulties managing and locating data later in the inquiry process. Zydeco addresses the organizational issues between unlinked ideas that arose in previous work by enabling data objects to consist of a data object annotated in multiple ways: with titles, tags, and an audio note. We found that enabling the data objects to be annotated in multiple ways can benefit the later retrieval and use of the information.

Benefits

These multiple annotations, typically applied with a photograph (98% of the time), enabled students to collect a mix of annotations and photographic data around a particular idea that was organized for easier retrieval and use. For instance, a piece of photographic data students collected was an image of an eagle titled "Eagle". The eagle data had five tags associated with it, describing traits of the eagle such as "nostril in front" and "four toes on foot". The student included the audio note, "This is an eagle eating a fish on a rock". By having these annotations on the photograph, it can aid students in locating and reviewing ideas and information about the eagle (benefits of this fully discussed below in the scaffold *Enable data assessment and filtering by data characteristics* in the Zydeco:UseData system).

An additional benefit to allowing students to annotate their data in multiple ways linked was that they used these different annotations (title, tags, and audio) to record different sorts of information. Students typically used the title to name the artifact collected (e.g. Eagle) and the tags to describe the characteristics and traits of the artifact (e.g. four toes on foot). Audio notes were used as a means of recording additional information (such as reading museum placards) without having to deal with typing on a handheld device. Many students were using the audio notes to this effect, recording observations on the animals or information presented on displays in the museum.

6.1.3 Integrate Co-Created Tags that Reflect Expert Discussion

It is difficult for learners to annotate data as they must reflect on what an accurate and useful annotation is for the data and ensure they have enough information to answer their investigation (Metz, 2000). Given the open-ended nature of Zydeco investigations where students can collect a variety of data and that data is shared with their peers, if the students do not annotate the data in a useful manner they will be unable to synthesize an explanation with the resulting data set.

To overcome these challenges of annotating data in a useful manner, students and teachers discussed and co-created tags that were able to be used to annotate the data in the system. For example, students and the teacher discuss and decide that “four legs”, “large brain”, and “ears on top” are potential labels that students may use collecting data for their investigation on animal relatedness, and come to a shared understanding of what is meant by “top” and “large”. When the learner is outside the classroom, carrying out an investigation, and finds an exhibit of a hippopotamus the learner can get an idea for what useful labels for the data are by reviewing the co-created tags on their handheld during the tagging phase of data collection, accessing the expert guidance from the discussion with their peers and teacher. The learner can then reflect upon the possible annotations in the co-created tags and examine the hippopotamus, tapping on an annotation in the list to annotate the data with “four legs” and “ears on top” upon reviewing the animal’s features.

Benefits

The integration of the co-created tags was beneficial for students in applying accurate and useful annotations (tags) on their data because the language had been vetted through discussion with their peers and an expert. Students had an idea of how they might apply the tags and annotate the data, instead of having to create all the annotations outside the classroom on the fly. Due to the discussion and vetting, the co-created tags embodied in the tool should be meaningful if accurately applied by the student, which the learners were able to do with over 91% of these tags being factually accurate and potentially useful towards the student investigation. It was noted that students had a statistically significant increase in accuracy and usefulness when they used the co-created language versus making new tags in the field (p=0.0005), which provides further support to the benefit of this scaffold.

The co-created tags were created so that all of the potential annotations were relevant to the investigation. By having the tool reflect the language, students were able to review the language when out of the classroom, helping them determine what potential traits were relevant for the data they were capturing and annotating it accordingly. An example of this could be seen as a student tried to determine how to annotate a photograph they were capturing, reviewing the list of annotations in order to decide:

> Student: "What do I put? Elephant, or what? Fur? Wait. Hair."

Data entry can be an issue with handheld devices and provide another barrier to using the co-created language to describe data. Having the system provide a list of all the co-created tags and applying them by tapping reduced the mechanics of data entry. This may have the effect of encouraging students to use the co-created tags without enforcing it explicitly; students' preference towards the co-created tags was noted in the trial (83% of all the tags used were from the co-created tags).

Additionally, integrating co-created tags into the system encourages consistency of tagging between students as all potential tags that are synonyms or simply different spellings of each other (such as "four legs", "4 legs" and "quadruped") can be merged into a single option ("four legs"). By having the system integrate the co-created tags, it prevents ambiguous and redundant word choices to describe the same concept, which can hinder finding and assessing data. As 83% of all the tags applied to the data were from the co-created tags, this implies that most of the data students collected were consistently tagged.

Areas for Improvement

Integrated co-created tags that reflect expert discussion is ineffective for students who do not desire to use these co-created tags and follow the same format as the expert discussion. In this study, one pair barely used the co-created tags and had the tags that were judged the least factually accurate and also the most that were labeled as unclear when coded by the researcher, such as "Bio".

The more students that do not use the co-created tags and have unclear or factually incorrect annotations can decrease the ease of using the aggregated data set and could hinder the data sharing aspect. However, students not using the co-created tags were a rare case, as all the other pairs had the majority of the tags they applied were from the co-created tags.

The benefits of applying the co-created tags could be extended to include having students co-create titles for the data through discussion with their teachers. In this trial students were provided a list of potential titles in the system, but the researchers provided them and the students did not discuss these titles. Application of these potential titles was far less frequent than the co-created tags (35% potential titles versus 83% co-created tags were used on all the created titles / tags), which could be due in part to students not feeling the same ownership over the title choices having not been part of the creation process. Students tended to make more descriptive and exact titles instead of using the potential titles, such as "Gray wolf" instead of "Wolf" or "Owosso Mastodon" instead of "Mastodon".

Due to the set up of this investigation, it was not possible for students to know what potential ways that they may title the data in the museum if they desired to use the titles as the name of the animal. This is due to most of the students never having visiting this history museum before and being unaware of what exhibits it had. If students were more familiar with the setting and what sort of data they may find, co-creating the titles could work.

An area where the co-created titles may help was noted in this trial. Some students had spelling difficulties on titling some of the data (15% incorrectly spelled), which can make the data harder to retrieve. If the students generated these descriptive titles in advance and used these titles more, it could reduce the incorrectly spelled titles, making the data set easier to search and filter when finding data to use in an explanation.

A part of the benefit of co-creating the titles with expert guidance can be seen in this study though, as the teacher did discuss how students the data could be titled (through providing animal names) and the system provided potential titles of animal names. In this study students tended to title the data following the same general format as the potential titles (giving the name of the animal in 80% of all the titles), showing the students still made relatively consistent and correctly spelled titles without co-creating the titles (but having the system integrate possible titles to choose from).

6.1.4 Guide Data Collection Step-wise

While outside the classroom, students were using handheld devices with limited screen real estate while in a novel environment filled with stimulus and potential distractions (Balling & Falk, 1980). To help learners overcome the challenges inherent in data collection and the additional challenges using a handheld device in the field brings, the system broke down data collection into distinct stages and guided the learner through each stage. This breakdown was done through collection and annotation step (choose a question the data will answer, choose to take a photograph or audio note, title, audio note, tagging) and guided with a prompt at each stage regarding what the learner should accomplish. The learner could not advance until they had accomplished each step (except for the optional audio note).

Benefits

Breaking down the data collection process into steps and guiding students through each step in the software requires students to focus on one task at a time. These steps also reduced the amount of information that was required to be present on the screen.

Students were able to navigate through each step's screen and follow the instructions to focus on each aspect of the annotation process. Audio recordings gathered during the out of classroom data collection captured students as they spoke aloud through this process, such as this student collecting data on a Mastodon:

> [Takes photo] "A mastodon? I took pictures of both." (re: a mammoth)
>
> [Records audio note] "This is a mastodon it's internal cuz it's bones. They have holes a lot of holes and they have big tusks. it's huge I had to take the pic from up here. So ok. Byeee."
>
> [Tags data] "Big eye holes, big teeth, few holes in skull."

This scaffold works in conjunction with *enable data objects to be annotated in multiple ways* (discussed above) around a single datum object. By having the system require the students to title and tag the data they collect, it also ensures that the student takes advantage of the multiple ways data can be annotated. As the only way to find and assess the data set is by the photograph and annotations, having more annotations on the data (if they are accurate and useful) improves the chance of the student or their peers being able to find and understand the data object when they review the data set later.

For these stated benefits, it is beneficial to guide the students step by step through the data collection process, focusing on one aspect or annotation of collection at each step when students are collecting data outside of the classroom.

Areas for Improvement

More support and guidance is needed to assist students in collecting audio notes. This additional support could involve being more explicit in the prompts or providing a format for what a useful audio note might contain through either software or instructional scaffolding. The audio notes were the least accurate (76% notes contained all factually accurate information) and useful (67% of the notes were potentially useful) annotation students applied. The lower accuracy and usefulness is partially due to the amount of information that students place in an audio note, which can bias the result. An audio note may have students mentioning multiple characteristics or ideas about an animal, each of these ideas capable of being inaccurate or accurate, as opposed to a tag or title that is focused on one characteristic or idea, which leads to more partially accurate audio statements (13% of the audio notes were partially accurate). For instance, the earlier audio note about the Mastodon contains multiple pieces of information that may be factually accurate or not:

> Student: "This is a Mastodon it's internal cuz it's bones. They have holes a lot of holes and they have big tusks. It's huge I had to take the pic from up here. So ok. Byeee."

There was a redundancy in a third of the audio notes where students were mentioning the tags they had annotated the data with. Usage of the audio notes back in the classroom indicated that while all but one pair of students listened to audio notes, when interviewed the students claimed the use of the audio to analyze and interpret the data was mixed: the redundant information and off-topic data were reasons students did not find them useful, while others felt audio notes helped them learn more information or remember why they collected data.

6.2 Zydeco:UseData Scaffolds

Research Question 2) How do students use data from the student-collected data set given the system's scaffolds?

The scaffolds of the Zydeco:UseData tool enabled students to analyze and interpret their peers' data and use the data to construct an evidence-based scientific explanation (refer to

Table 6-2 above). Analyzing and interpreting data is a difficult task for students; these tasks were made more difficult by having a large, student-collected data set that must be searched through and examined to find relevant information. The scaffolds appeared to be able to mitigate the challenges around these tasks, as students were able to find, assess, and then select data to construct evidence-based scientific explanations.

In the following sections each scaffold from the Zydeco:UseData activity is discussed, covering the benefits of each scaffold and any areas we noted that could improve the scaffold design.

6.2.1 Enable Data Assessment and Filtering by Data Characteristics

Students need to search through their personal and peers' data set to find appropriate evidence, though this is a difficult task, especially to interpret the data collected by a peer (Berland & McNeill, 2009). Then, the students had to assess the data to determine if the data they found should be selected for use in their explanation.

The Zydeco:UseData tool enabled students to review all of the characteristics of the data in the data set and filter by any textual annotation, through keyword searches and tap-able filter options. Students viewed the thumbnails displaying the image and compared the image to the title and tags of the data object to perform a quick analysis of the data's quality. They could also tap on the thumbnail to get access to a larger, zoomable image to examine it further and playback any audio notes.

Benefits

Enabling assessing and filtering the data by the data characteristics was useful for various reasons, which varied by the characteristic.

In interviewing students about how they assessed data and reviewing how they used the annotations, we found that the semantic meaning of tags was understood and that many students compared tags to images to fact check the tags for accuracy and used the tags to perform sensemaking on the data. The use of the tags for sensemaking is a different function than most tag use from the literature, being typically used only to describe and organize data (Ames & Naaman, 2007; Bischoff et al., 2008; MacGregor & McCulloch, 2006). As students had an educational task and were working with a data set annotated by students with the same task, the tags were constructed more with this goal in mind (particularly the co-created tags that reflect expert discussion). As all the users were working towards the same educational task on data analysis, it can account for the different use of tags than normally seen in the literature.

While tags were the primary source students mentioned using for analysis, audio notes also were used. Some students claimed audio notes served as a reminder of why the data was gathered and as a means of learning additional information.

Titles were less of a sensemaking support and instead used as a means of identifying the artifact in question. Searching on titles was the primary means students searched and found data to use as evidence in their explanations (47% of all the data added as evidence was found while title searching). Students predominately used the title searching option to filter the data set, applying simple keyword searches that tended to be the name of the animal they desired to find (e.g. "Eagle"). Most of the students were able to do this, which is due to the combination of 1) having the data set with enough consistent annotations that were the names of animals (80% of all the data objects were of animal names), and 2) being able to construct the simple keyword searches to filter the data. The ability to overcome the search challenges with simple keyword searches reflects what has been seen in the literature in section 2.3.3, as students tend to be able to make simple keyword searches but otherwise are better able to browse the data set if it requires complex search tactics.

Because each annotation type served a different but useful purpose, there seems to be no one size fits all annotation type when performing inquiry activities. Other systems could benefit from supporting a variety of data annotations, enabling students to express more information and create a data objects with these multiple types of annotations to improve the ability of students to use the data.

Areas for Improvement

Difficulties applying search filters were discovered when students were trying to find data for use in their explanation; some students encountered spelling issues and had difficulties applying search filters and finding data, at times being unable to find any data (searches returning no results).

In order to address issues students had with spelling when students were performing title searches, future designs could have a list of all title names (filterable as the user types) and suggest corrections to fixing a search that will return no results, similar to search techniques used on search engines such as Google, which provides prompts to direct the user to similar search terms that could be due to a spelling error, "Did you mean ____"

(Google inc.). This can help address the spelling difficulties that many novice learners have, which are non-germane to the science practices the students are learning.

Beyond these spelling problems, a usability issue arose when students were applying search filters. Students would not realize they were searching through their own data and would apply additional filters (e.g. trying to find an animal titled "owl" when looking at their own data set), making them repeatedly not find any results. To a lesser extent students would also apply too many filters when looking at the collective data set and also not find any results (e.g. trying to find data with a tag "fur" and a title of "snake"). While searches that return no results are not inherently bad, students can become confused at how to get on track filtering the data set in a manner that returns results and get frustrated and waste time. This issue was likely made worse by the interface design, which had the filters applied in different visual styles. All filters were placed in a consistent location and had a uniform visual style, except the filter that enabled students to search on their own data, which had a different style and location.

In order to address the issues some students had with having multiple filters applied and having searches return no results is to 1) ensure the existing filters are clearly visible and consistently and suggest removing search filters if they receive no results when applying multiple search filters, and potentially to 2) not allow students to apply additional filters when looking at their own data (students never used search filters to find data within their own pool, and had at most 34 pieces of data in this trial), though this could be limiting in extended usage trials when students gather more of their own data.

6.2.2 Enable Exploratory Search

When searching for information, people will sometimes perform directed searching (such as wanting to find information on an Eagle and giving a keyword search "Eagle"), but also need to engage in more open exploratory searches or general wandering through information. The system scaffolded learners to be able to wander through the data set by enabling the user to scroll through thumbnails of data with titles and tags displayed and by viewing an alphabetically sorted list of all the tags used, along with the number of data objects annotated with each tag.

Benefits

By reviewing the list of all the tags on the data, students are able to get a sense of how the data is annotated and choose particular traits to filter the data set to find data that matches those traits. Since all the potential searches are listed out (along with providing the additional information on the frequency of data with each tag), students are guaranteed to have search results and know how many data objects match a tag.

Even if the student was aware of what data they might desire, there were cases where students were unable to apply search filters to filter the data set (discussed above in *assess and search upon data characteristics*). Using the listing of tags present on the data and tapping on a tag to filter the data prevents the students from spending extra time constructing a search query from which no data would match, reducing the need for students to struggle with the difficulties of spelling and search mechanics as they find data.

Being able to browse through the data set to get ideas may prompt students to follow up with search filtering to find more relevant data, or use it as their sole means of discovering data (if they struggle with searching or do not desire to search). Both tactics were noted by students who did exploratory browsing and followed up with a title search, as well as other students who chose to use the browsing as their primary means of finding data.

This scaffolding around having the tag list and filtering via tapping on tags was used by a portion of the students to assist them in finding a number of pieces of evidence they used in their explanation. This indicates that the tag list appears to have merit in the educational context where learners struggle more with searching.

Areas for Improvement

The idea of having a list of the tags that can also be used to filter the data can be expanded to other annotations. In this inquiry context, title searching was the most common search method, so having a list of the other textual annotations (such as the titles) could further reduce searching difficulties and give students more avenues to get a sense of the data set.

Expanding the ways the data set can be analyzed would follow the scaffolding guideline to "Provide multiple ways to inspect data", which suggests having different ways to inspect the data set could assist learners with analysis, as well as being a search scaffold (Quintana et al., 2004). Having this combined search and analysis function would also follow Luchini et al.'s scaffold strategy to "Design double-duty scaffolds" guideline for maximizing usage of the screen real estate on mobiles (Luchini et al., 2004).

6.2.3 Expedite Search for Similar Data

The Zydeco:UseData system enables students to filter the data set through keyword searches and also do more exploratory searches (via choosing from the tag list) or browse through thumbnails of the unfiltered data set. Students have difficulty searching and sometimes need to view data that is similar to a piece of data they discovered (whether through exploratory or directed searches). Finding similar data is a common task in analysis and so a scaffold was created to find similar data by matching the data characteristics in common (titles and tags). For example, upon having a student discover a data with the title "Eagle" that has the tag "four toes on foot", the student can find similar data to find data annotated similarly, pulling up other Eagles and animals with four toes on their foot.

Areas for Improvement

Though some students searched on similar data after browsing the unfiltered data set, the scaffold to support the task of finding similar data was barely used, only accessed by four pairs. It was noted that the functionality provided by the scaffold was being applied manually by students, for example students were browsing the data set and seeing an animal of interest to search on, then typing in the title.

As students were manually applying title filtering similar to how the scaffold would expedite the process, this lack of use comes down to two factors: 1) the method of implementation caused a lack of discovery and understanding of the purpose of the scaffold, or 2) students chose to filter the data through the standard methods and did not need the task to be expedited. The issue with a lack of discovery has been noted in previous scaffolding literature, which found that scaffolds should be very visible as learners will not usually trigger scaffolds, and more essential than optional scaffolds should be created otherwise learners may bypass them (Quintana et al., 2002). The implementation for triggering the finding similar data scaffold was through a button at the bottom corner of the page when focused on a datum object; the button was not prominently placed and this feature is not essential to their task, rather it was designed to make the process of searching easier. Future work is needed to see if having this scaffold more visible can assist some students, or whether the students prefer applying filters to the data set themselves.

6.2.4 Supply Supplemental Curated Information

The students are primarily working with the student-collected data set, which may not contain all the information a student needs to answer their investigation and some data is labeled with inaccurate annotations. Because of this, it is useful for students to have a means to fact-check the data and fill in some gaps in knowledge that the student-collected data set lacks.

Benefits

Having supplemental curated information served this need for several of the groups in the study. By reviewing the supplemental information they were able to revise and expand their reasoning about the data in question, enabling the students to gain more knowledge about the data and construct a better explanation.

For example, students read more about turtles in the supplemental data to learn how they might survive, adding, *"A turtle also has webbed feet that help it swim faster from predators or swim faster to get let's say a fish."* to their reasoning immediately after reading about similar information in the supplemental information.

Areas for Improvement

Several students mentioned in interviews that they did not use the supplemental information because it did not contain information they needed for their investigation. One way to address this is to enable students to be able to search online or in other resource banks to find additional information. This is relevant in that the curated information may not cover all the information that students need.

However, collecting and interpreting information from a website can be difficult to novices in inquiry and students require guidance. Several science inquiry systems have sent students to curated list of sites and provided prompts on what information to look for (Quintana & Zhang, 2004; Linn et al., 2003). A curated list of information may not be enough for students who are collecting data outside the classroom and gathering a diverse collection of artifacts and phenomena, as the curated list of information may fail to encompass the range of artifacts.

Research on software called IdeaKeeper investigated how to scaffold students through collecting information from websites and worked with more open-ended searching by providing prompts and scaffolds to gather information from websites (Quintana & Zhang, 2004). Similar scaffolding could be integrated into the Zydeco learning environment to scaffold students at gathering supplemental data from websites and relevant databases.

Having a system scaffold students in gathering additional information from the Internet would ease the task of educators providing the full range of content that the learners may need to support the data they and their peers collected. While student data collection may encompass all the information they need for the investigation, this supplemental information is useful for fact-checking and also for filling in missing information that a student forgot to collect in the field.

6.3 Co-Created Language Contributing to System Utility

This work has discussed how the scaffolds implemented in Zydeco provided benefits toward the system utility around data collection and use, including how integrating co-created tags such as "big eyes" that reflect expert discussion and have a shared, mutually understood meaning helped overcome the challenges of annotating data in a useful manner (6.1.3). The co-created tags provided additional benefits to the students and were a contributor to the overall utility of the system as an example of a distributed synergistic scaffold. Specifically, the system integrated the social scaffolding of the teacher discussion with the software scaffolding by integrating the co-created tags into the task of data collection and annotation.

6.3.1 Extending Distributed Synergistic Scaffolding

Tabak initially proposed the idea of distributed synergistic scaffolding (DSS) and looked at how the social scaffolding of the teacher and peers and the software scaffolding could work together to be more effective than either form of scaffolding individually (Tabak, 2004). Additionally, Tabak's work on DSS claimed the scaffolding was more effective when the teacher and software shared a consistent language, enabling them to reinforce the meaning of each (Tabak, 2004). The synergy in the scaffolding was due to having the teacher give lessons about science topics and using and reinforcing the same framework, task structure, and language (including definitions and vocabulary) as the software system that students were using to complete the activity.

Tabak did research with high school students in an introductory biology classroom who were studying evolution and completing their curriculum using desktop software as part of the BGuILE project (Tabak, 2004). The system and teacher both used a language agreed upon by the researchers and teachers to talk about animal survival and evolution. While Zydeco similarly used the software to reinforce the language and lessons of the teacher, it extended Tabak's work by providing an example of a system that is used both inside the classroom and outside the classroom with less teacher support. Additionally, with Zydeco the students and teachers were able to construct the language together that students would use to annotate and talk about the data, which can further motivate students to be involved in the activity and make it more personally relevant to the students (Pintrich, Marx, & Boyle, 1993). Together, these changes provide an example extending how distributed synergistic scaffolds may be implemented.

6.3.2 Tags: An Instantiation of a Co-Created Language Mediating Contexts

Tags were a fundamental piece of the annotated multimedia data that formed the artifacts that students brought with them between the out of classroom and in classroom contexts. These data acted as connecting elements and enabled students to seamlessly transition their inquiry activities and continue working between contexts. This research has found that these data and annotations, particularly the co-created tags, were able to communicate the meaning of the artifacts and phenomena students observed outside of the classroom in a way that allowed themselves and their peers to later assess and interpret the data for use in constructing a scientific explanation back in the classroom. There are several high-level factors surrounding the development, integration into the system, and use of data annotated with the co-created tags that may have contributed to these positive outcomes in this trial. These tags are the instantiation of a co-created language that the teachers and students use to discuss data.

The co-created tags are an example of a distributed synergistic scaffold: the co-created tags were developed through teacher and student discussion in the planning stage of the inquiry activity, a method of social scaffolding, and were incorporated into the software scaffolding of the Zydeco system. Incorporating these elements into the technology benefited the activity in several ways:

1) Providing students with access to expert support via the teacher feedback.
2) Giving the students ownership of the tags, which research has shown can motivate students and make them work harder at the task (Pintrich et al., 1993).
3) Ensuring the tags makes sense to the students and they shared a common understanding of how to annotate and use the data.

The benefits of the co-created tags could be witnessed across the various stages of the inquiry activity. For instance, having the Zydeco:CollectData tool continually display and use the co-created tags while students annotated data seemed to assist students during the difficult task of data collection and annotation by incorporating the earlier expert support of the teachers. By having students take ownership of the language, it not only could motivate students in the activity, but also encourage application of the language to create a more consistent and understandable annotations on the data set (83% of the tags applied were the co-created tags). The consistent use of the tags can help overcome some of the issues users typically face in tagging systems, as users often do not agree on the semantic meaning of words or use varying words to specify similar ideas (applying spelling variants or synonyms), which can reduce the ease to search upon and interpret the data (MacGregor & McCulloch, 2006). The students had a defined purpose for the visit and a sense of how to annotate the data through class discussion on the tags they created, which are potential reasons for the levels of reflection and effort students put toward the annotation tasks--levels that were exhibited by the accurate and useful annotations made on the data.

Because the data was annotated with the collaborative language, students understood the meaning of the tags and were able to use them to interpret the data that their peers had collected. This was reflected in interviews with the students, where they discussed how the tags were able to give them more information than the image alone and assisted students in finding data to use in their explanation. One pair even mentioning that, "(without tags) I don't know any facts about it or nothing." Data was not only interpreted by using tags, it was also evaluated by comparing the tags to the images. For instance, students would use the tags as an indicator of what to examine in the image, then zoom in on the image to fact-check the tags and ensure they are accurate. The utility of using tags in analyzing and interpreting data for use in students' explanations was potentially improved because students annotated the data with the co-created language and the students understood the semantic meaning of the language (e.g. students understood what "ears on top" meant).

This work found that the co-created tags in the Zydeco tools seems to have contributed to student's ability to address the challenges of collecting and using large amounts of peer-collected multimedia data to construct a scientific explanation. Through developing co-created tags and integrating the tags as scaffolds in the Zydeco tools, we provide a scaffolding solution that fits a variety of inquiry settings. This ability to adjust the inquiry activity to fit the context of the curriculum and environment is important when dealing with out of the classroom experiences since the settings can be vastly different. Learners will be visiting local sites of interest related to an individual investigation, which may vary across contexts such as studying animals in a history museum, energy transfer in a science museum, or water quality in a nature park. Additionally, students are not forced to rely on having an expert be present to obtain guidance while performing these out of classroom activities because the co-created tags are a result of discussions with an expert, and have shown an ability to support learners in collecting a data set that contains accurate and useful annotations.

The benefits of the co-created language suggest that the methods in which the teacher's support and instruction can be reinforced by the software scaffolding and the software likewise reinforce the teacher instruction and discussion should be investigated further and that some of the scaffolding guidelines found within the literature should be expanded. In particular, existing guidelines should consider the use of distributed synergistic scaffoldings and provide guidance on how to combine social scaffolding with software scaffolding, such as this study performed with integrating the co-created tags that reflect teacher discussion into the tool. The Quintana et al. scaffolding framework, which was used as conceptual guidance for part of the design of the Zydeco tools, was focused entirely on software scaffolding (Quintana et al., 2004). The framework could be extended to provide advice on how to develop distributed synergistic scaffolding that helps software and social scaffolds reinforce each other.

6.4 Guidelines For Designers

Through this work I have developed several guidelines for designers seeking to support learners engaging in inquiry activities that span in classroom and out of the classroom contexts where the learners engage in multimedia data collection and share their data with their peers (Table 6-3).

A summary of the guidelines developed from the lessons this work presented throughout sections 6.1 and 6.2 follows below in Table 6-3.

Design Guidelines

Integrate Expert and Software Scaffolding

1. Students and teachers should work together to co-create data descriptors that can then be integrated into the software as tags that students apply to the data they collect for later use in finding and assessing their personal and peers' data.

Designing Software Scaffolding to Support Students in Collecting Multimedia Data

1. Provide the means to incorporate multiple types of annotations (such as titles, tags, and audio notes) into multimedia data collection and use to enable students to collect data that can be useful later.
2. Software scaffolds should mirror the language being used between students and experts in the classroom to help students overcome the difficulties of reflecting on and annotating data.
   - [LIMITATION] Students might not follow the general model of the co-created tags and lose out on scaffolds around data collection, which may lead them to annotate the data in a less useful manner.
   - [NOTE] Reducing non-germane data entry mechanics, through methods such as tap to annotate, can encourage use of the co-created language without forcing it.
3. Software scaffolds should guide students through data collection stepwise, focusing on one aspect of the collection or annotation process at a time to prevent students from getting cognitively overwhelmed. When possible, breakdown the inquiry data collection into sub-questions to provide additional direction on what sort of data to collect.
   - [NOTE] Students need additional scaffolding (through social and/or software) on how to record audio notes that will be useful in their investigation

Designing Software Scaffolding to Support Students Using a Large[1], Student-collected Data Sets

1. Provide interface elements to help students review and filter the data set by the different data characteristics so the data does not overwhelm them.
   - [NOTE] Some students may require additional scaffolding on how to search through the data. Some options may include different 'clouds' of textual annotations to assist with filtering, as well as search suggestions and search corrections.
2. Incorporate scaffolds to support students in viewing summaries of the data's annotations and in browsing through the data set in order to assist them in exploring the data set and determine what data they may want to investigate in greater detail.
3. Provide and scaffold students through finding additional information to supplement or fact-check the data they collected, such as through curated information or Internet resources.

Table 6-3. Guidelines for designers seeking to support students engaging in inquiry activities that span inside and outside of the classroom where the students collect data and share their data with their peers.

[1] Large data sets referring to sets containing to several hundred pieces of data or more.

6.5 Broader Applications

The guidelines presented in this work have the potential to apply more broadly to assist designers in building any system that requires the acquisition and utilization of digital artifacts, and do not solely pertain to middle school science inquiry settings. This work may be beneficial toward any open-ended task where a group needs to come to a consensus on how to structure and share data for collective use.

The guidelines of this work could benefit multimedia story sharing systems. Sharing stories is an established way in which people share their collective experiences, culture, and knowledge; it is a way to integrate and communicate data in areas such as politics, public health, and education as well as in individual lives. An example of this is communicating stories about an infectious disease such as tuberculosis or HIV: how they spread, whether there is a geographical influence, measures people should take to prevent further spreading of the disease, and how to effectively treat the disease. These stories can be enhanced through using multimedia to communicate complex ideas, such as with photographic images and video.

Designers of systems that facilitate these activities could benefit from the guidelines developed in this work by incorporating software scaffolds that facilitate collecting and annotating multimedia information that is shared. This can enable users to collect and share large quantities of multimedia information, which can assist the storytellers in finding the right data for the story they want to tell. In addition to making the multimedia information easier to find, by annotating the information that is being utilized to build stories with a language that is co-created between users, the annotations could have the potential to help users communicate as they build stories together and reflect on others' ideas to construct better stories. Families could use such software to create stories about where they immigrated from and share the media and information they have collected over the years—including photographs, copies of immigration documents, and family videos— with the children to learn about their family history and heritage.

Expanding on this idea of informational stories, another such area that could benefit from the guidelines is developing systems for sharing information between investigative journalists. These individuals may work with a team of other journalists, sharing the information they have discovered in the forms of photographs, videos, interviews, etc. Using information they gathered, they develop reports or videos for publication and broadcasting backed by this collectively generated content. As these teams can be distributed around the globe as they work on their projects, it is important for them to be able to collect data and communicate their ideas and the information they have collected effectively and efficiently.

In this situation, the guidelines this work posed around collecting and using data can be used to design software to assist these journalists in structuring their data collection process to make it more communicable. If data was collected and annotated using software following these guidelines then the gathered multimedia information would be more accessible to other investigative journalists, particularly if the data could be reviewed and filtered by discipline specific data characteristics such as the journalist who collected it, what sort of information the data contained, or the country in which it was collected. Having the shared information be annotated by the journalist that collected it would supply others reviewing it with the context and ideas surrounding collection. In this way, investigative journalists would benefit from using a system that incorporates the guidelines proposed in this work as they share the multimedia information and interviews they capture in a manner that is easy to find, evaluate, and utilize by their peers.

These are two examples in how these guidelines could be applied to broader applications beyond science inquiry. There are many other applications that can utilize these guidelines to design systems scaffolding users in collecting and later using shared multimedia data.

CHAPTER 7

Conclusion, Limitations, and Future Work

The goal of this research has been to understand how to design software scaffolds for students engaging in science inquiry as they 1) collect multimedia data outside the classroom and then 2) use a large, student-collected, multimedia data set to construct scientific explanations in the classroom.

Two software tools were built with scaffolding solutions for this inquiry context, Zydeco:CollectData and Zydeco:UseData. The scaffolds implemented were initially informed from literature concerning the challenges of collecting and using data and from similar systems that supported data collection and use. An iterative design process further developed the versions of the tools tested in this study.

Through evaluating student use of the Zydeco tools, we discussed areas where the scaffolds were beneficial and areas of improvement to their design. From this analysis design recommendations were presented for other designers seeking to support similar multimedia data collection and use activities.

This chapter presents a summary of the conclusions drawn from this work and how they contribute to the literature. This is followed by a discussion on limitations of the study and future work.

7.1 Conclusions

The Zydeco tools were able to mitigate the challenges students faced during data collection and when using the data to construct scientific explanations. With the Zydeco tools, students were able to construct scientific explanations with data they and their peers collected in the field. The scaffolds in the Zydeco:CollectData tool enabled students to collect data with annotations that were factually accurate and useful toward answering their investigation. Zydeco:UseData had scaffolds that incorporated the collective data set that resulted from the out of classroom activity, and students were able to use the data to find, assess, and select the data for use in their explanation.

This study has demonstrated the ability to support a new form of educational activity with the scaffolds built in to these two Zydeco tools. This has implications for others seeking to support similar science inquiry activities, as well as other contexts such as supporting collecting data and sharing data for multimedia storytelling or investigative journalism.

7.2 Contributions

This research presented three contributions:

1. The design of two integrated tools with scaffolds for collecting and using data.
2. An evaluation of how the students were able to use the tools, given the scaffolds, to complete the tasks, including areas where students needed additional support.
3. Design guidelines for other designers to follow in developing scaffolds to assist learners in similar tasks.

Throughout the course of the Zydeco project, a number of publications have been published or accepted for publication disseminating various lessons learned:

- **Kuhn, A.**, McNally, B., Schmoll, S., Cahill, C., Lo, W., Quintana, C., and Delen, I. (2012). How students find, evaluate, and utilize peer-collected annotated multimedia data in science inquiry with Zydeco. In Proceedings of the 2012 annual conference on Human factors in computing systems (CHI '12). ACM, New York, NY, USA, 3061-3070.
- **Kuhn, A.**, Cahill, C., Quintana, C., and Schmoll, S. (2011). Using tags to encourage reflection and annotation during nomadic inquiry. In Proceedings of the 2011 annual conference on Human factors in computing systems (CHI '11). ACM, New York, NY, USA, 667-670.
- **Kuhn, A.**, McNally, B., Cahill, C., Quintana, C., and Soloway, E. 2011. Constructing scientific arguments with user collected data in nomadic inquiry. In Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (CHI EA '11). ACM, New York, NY, USA, 2167-2172.
- **Kuhn, A.**, Cahill, C., Quintana, C., and Soloway, E. 2010. Scaffolding science inquiry in museums with Zydeco. In Proceedings of the 28th of the international Conference Extended Abstracts on Human Factors in Computing Systems (Atlanta, Georgia, USA, April 10 – 15, 2010). CHI EA '10. ACM, New York, NY, 3373-3378.
- Clegg, T., Bonsignore, E., Yip, J., Gelderblom, H., **Kuhn, A.**, Valenstein, T., and Druin, A. (2012). Technology for promoting scientific practice and personal meaning in

life-relevant learning. In Proceedings of the 11th International Conference on Interaction Design and Children (IDC '12). ACM, New York, NY, USA, 152-161.

- Cahill, C., Lo, W., **Kuhn, A.**, Quintana, C., McNally, B., Schmoll, S., and Krajcik, J. (2011). Student use of multimodal data and metadata tools during nomadic inquiry. In Proceedings of the 10th World Conference on Mobile and Contextual Learning (mLearn '11).
- Cahill, C., **Kuhn, A.**, Schmoll, S., Lo, W., McNally, B., and Quintana, C. (2011). Mobile learning in museums: how mobile supports for learning influence student behavior. In Proceedings of the 10th International Conference on Interaction Design and Children (IDC '11). ACM, New York, NY, USA, 21-28.
- Cahill, C., **Kuhn, A.**, Schmoll, S., Pompe, A., and Quintana, C. 2010. Zydeco: using mobile and web technologies to support seamless inquiry between museum and school contexts. In Proceedings of the 9th international Conference on interaction Design and Children (Barcelona, Spain, June 09 – 12, 2010). IDC '10. ACM, New York, NY, 174-177.

7.3 Limitations of the Study

This was an initial field study to get a sense of students' usage of the two tools and see how the scaffolds were able to support the students and identify areas where students need additional support; however, this research was identifying emergent ideas, and additional work is needed to verify the design recommendations and effectiveness of the scaffolds used in this trial.

The generalization to different populations and activities is open to question- small changes in system design, student ability, activity structure, and maturity of students can lead to radically different outcomes. Spelling issues and search ability would change dependent on the students' abilities, which could change how they approach the search process. Variance in the pre-activities leading up to the museum experience, different museum environments, the teachers' experience acting as a facilitator, and the overall inquiry tasks could all modify the results. These changes could lead students to annotate the data in a vastly different matter, and find and retrieve the data using different methods than we found and struggle with different areas in the process of using data to construct their explanation. While I presented the results of one such activity and population, more studies need to be performed to see the commonalities across activities and student populations.

Methodological issues also limit the generalization of the students' performance. The act of interviewing the students each day by nature could have changed the activity- our questions could prompt them to reflect and also change how they utilize the system.

7.4 Future Work

Many routes of future research have emerged from this work. This thesis has presented scaffolding guidelines to scaffold student's through collecting data outside of the classroom, and then using this data to construct explanations, as well as laid out areas where additional scaffolds are needed. One of the next steps, which is currently ongoing, is to redesign the system utilizing the knowledge gained from this study. Once the redesign is finished, Zydeco will be released to the App Store in order to get feedback on how educators and students utilize the system in the wild.

Future work planned for Zydeco falls into three general directions:

- Further assessing the design guidelines and supports and scaffolds developed from these guidelines.
- Extending the scope of activities Zydeco supports.
- Integrate explanation construction with sequential storytelling.

Further Assess Design Guidelines and Scaffolds

Future work will need to investigate how usage patterns change for different inquiry contexts and as students become more experienced at inquiry. This will enable us to determine what new scaffolds are needed as students engage in more diverse and complex inquiry activities. The changes in work pattern over longitudinal use is also of interest, as studying usage over time can provide knowledge on how the design must change between novices and more experienced learners. Work towards a longitudinal study is beginning and in the 2012-2013 school year three science classrooms will be using Zydeco for the entire year. Additionally, several other test sites around North America will begin using Zydeco in the following year to see how the scaffolds in Zydeco work in different contexts.

Beyond looking at more contexts, additional work is needed to narrow down the interplay between scaffolds. For example, having the option to record audio notes may be encouraging more thoughtful tagging of the data for later usage. A future trial investigating the differences in groups of students who are prompted to record audio notes, required to record audio notes, and do not have the option to record audio would be interesting to see how it affects the quantity and quality of data the students collect.

Another area to explore is an analysis of the benefits of using different scaffolds on the learning gains students achieve. This will be done by examining the learning gains of students using the scaffolds in the tools and comparing these learning gains when students use different sets of scaffolds; this will enable us to determine how different degrees and types of scaffolding affect learning across the activity as well as how it effects use patterns.

Extending the Scope of Activities Zydeco Supports

Another area that is being explored currently is how to enable students to construct scientific explanations with the data while they are outside of the classroom. In this scenario there was a separation between data collection and use due to time constraints and to prevent cognitive overload in the field. However, a properly scaffolded environment could potentially support the entire process and enable this activity to be carried out entirely in the field.

There is also room to extend and build upon the system, adding in additional support for more complex data analysis and utilizing numerical data, collected via various sensors or probes. This will be needed to support the full range of inquiry activities that students might perform in the field. Because the numerical data brings in additional challenges, additional research is needed in this area to understand how to support students in collecting and then using this data.

Furthermore, there is more room to enhance and prompt students to analyze their data. For example, the system could be expanded and tested with reflective prompts asking the student why they are using each piece of data in their explanation and see if this additional scaffolding is effective. Different visualizations of the data set can also be utilized to aid students in getting a sense of the overall data set and better being able to compare two or more pieces of data.

Integrate Explanation Construction with Sequential Storytelling

An area of research that begun in August 2011 after completing this study was using Zydeco in new environments while investigating adding in storytelling elements. In collaboration with researchers at the University of Maryland at College Park, Zydeco has been used as part of a Kitchen Chemistry curriculum (Clegg et al., 2012).

In Kitchen Chemistry, students learn about science and experiment by adjusting ingredients in the food they are making and observing the differences in a controlled fashion. A study was done where students compared using Zydeco and StoryKit (a mobile story writing system) (Clegg et al., 2012) in various activities observing and detailing their cooking experiments. This study has revealed that the data collection interface in Zydeco is helpful and supports students, but students enjoy and prefer the sequential storytelling abilities that StoryKit offers. Because of this, future work with Zydeco is seeking to integrate sequential storytelling as another option in addition to using the Claim-Evidence-Reasoning framework and explore how to scaffold the challenges of scientific storytelling.

Appendices

APPENDIX A: Parental Consent Forms

CENTER FOR HIGHLY INTERACTIVE CLASSROOMS • CURRICULA • COMPUTING IN EDUCATION
University of Michigan • 610 East University • Ann Arbor, MI 48109
734-476-5419 • claracah@umich.edu

Dear Parent/Guardian:

We would like to invite your child to be involved in a project focused on using new technologies to help students do science inquiry in museums, in order to help them connect what they are learning in the classroom to what they experience on their field trips. The overall project will enable your middle school student to do a scientific investigation using evidence they gather through a museum field trip. During the visit, students will each use an iPhone-based program to enable them to collect data in the form of voice notes, photographs, and observations, and to guide their learning and investigations in the museum. When they return to the classroom, they will be able to explore, summarize, and analyze the information and data the class collected during the museum field trip, in order to make and support scientific discoveries based on their collected data. In order to evaluate the effectiveness of the program in helping students build conceptual understanding and science inquiry skills, we will give students a short written or oral assessment before and after their museum visit, audio- and video-record students during their field trip and pre- and post-visit activities, and we will collect copies of student work related to the field-trip. We will also ask students for feedback about the experience to help improve the program.

We ask your permission for your child to be involved in the information-gathering component of this project. Please review and complete the attached forms regarding this project, and have your child return those forms to their teacher when completed. You may keep this letter, should you have any questions regarding this program.

We will keep students' names and identities confidential at all times. If you or your child decide against his/her involvement in the information-gathering component, you may withdraw your child at any time by informing your child's teacher and we will not include him/her in any of audio or video recording and/or collect copies of his/her work in any manner. Only voluntary participants will be filmed, recorded, and/or interviewed.

The University of Michigan Institutional Review Board has reviewed and approved this project. If you have any questions or concerns, please call or email Clara Cahill (phone: 734-476-5419, email: claracah@umich.edu).

Thank you for your consideration,

Clara Cahill
PhD Candidate, Science Education
School of Education
University of Michigan

IRB: Behavioral Sciences   IRB Number: HUM00035979   **Document Approved On: 1/25/2010**

Page 1 of 3

CENTER FOR HIGHLY INTERACTIVE CLASSROOMS • CURRICULA • COMPUTING IN EDUCATION
University of Michigan • 610 East University • Ann Arbor, MI 48109
734-476-5419 • claracah@umich.edu

PLEASE CHECK **ALL THAT APPLY**, SIGN YOUR NAME, AND ENTER THE DATE. PLEASE HAVE YOUR CHILD RETURN THIS PORTION OF THE LETTER TO HIS/HER TEACHER BY THE FOLLOWING DUE DATE:

________________. ONCE AGAIN, THANK YOU!

PLEASE CHECK THE FOLLOWING TWO BOXES IF YOU **DO** GIVE RESEARCHERS PERMISSION TO INTERVIEW AND ASSESS YOUR CHILD.			PLEASE CHECK THE FOLLOWING BOX **ONLY IF YOU DO NOT** GIVE RESEARCHERS PERMISSION TO INTERVIEW AND ASSESS YOUR CHILD.	
________ CHECK HERE	I give my permission for hi-ce to **interview, videotape** and **audiotape** my child for research purposes to improve teaching and learning.	**OR**	________ CHECK HERE	I **do not** give my permission for my child to take part in the documentation efforts of hi-ce to improve teaching and learning.
________ CHECK HERE	I give my permission for hi-ce to **collect classroom tests and other class work** for research purposes to improve teaching and learning.			

Students will be allowed to participate in the curriculum activities whether or not they agree to be interviewed and audio-taped. Should you have questions regarding your rights as a participant in this research, please contact Clara Cahill (phone: 734-476-5419, email: claracah@umich.edu). The University of Michigan Institutional Review Board has reviewed this project, and has classified it as exempt from further oversight.

______________________________ ______________
SIGNATURE, PARENT/GUARDIAN DATE

____
SIGNATURE, STUDENT

____
STUDENT'S NAME (PLEASE PRINT)

PLEASE RETURN THIS FORM AS SOON AS POSSIBLE.

Permission to Photograph or Record Electronically (Parent/Guardian)

I give my permission to the University of Michigan to record a photographic image and/or audio or video of my child for educational, academic, or research purposes. I affirm that I am 18 years of age or older.

Child's Name____

My Name____

My Relationship to Child____

Event____

If the faculty or staff of the University judges that education or research may benefit from the use of the photographs and/or recordings, the University may publish or sell (not-for-profit) them for academic purposes, or use them in any other professional manner that the University believes is proper, including, but not limited to: print publications, video streaming on U-M web sites, podcasting, and broadcast media.

I understand that the pictures and recordings belong to the University, and we will not receive payment or any other compensation in connection with the pictures and recordings.

I have had a chance to discuss this form with the University of Michigan staff and have received complete answers to all of my questions.

I release the University of Michigan from any and all liability that may or could arise from the taking or use of the pictures and recordings.

Signed____

Date____

Address____

____

For more information, contact the Outreach and Communications Office at the University of Michigan School of Education at (734) 615-1409.

APPENDIX B: Interview Tracking Booklet

INTERVIEW TRACKING BOOKLET

The Zydeco Project, May 2-6 2011

Group ID: ______________ Interviewer: ____________________________________

Record what time you begin and end each interview, based off the time on the students' device:

	Monday	Tuesday	Wednesday	Thursday	Friday
Start Time					
Stop Time					

Interview Progress

Mark partially completed sections with "/" and finished sections with an "X".

Question Stage 1

- ☐ Working on tutorial
- ☐ Finding data
- ☐ Interpreting grouped data and curating tags
- ☐ Additional Features (Low priority)
- ☐ Ranking Investigation
    - ☐ Making a claim
    - ☐ Making a reasoning
    - ☐ Giving peer critique
    - ☐ Reviewing/revising explanations
- ☐ Survival Investigation
    - ☐ Making a claim
    - ☐ Making a reasoning
    - ☐ Giving peer critique
    - ☐ Reviewing/revising explanations

Question Stage 2

Daily open-ended questions

M	T	W	T	F

Interview Protocol

Retrieve your audio recorder interview tracking booklet for the group you are about to interview. Review the areas they already have been interviewed on. Turn on your audio recorder then approach the students, following the introduction script. Spend ~4 minutes (up to 7 min if you are receiving ample feedback) on the first question stage and proceed to the next stage once the time is up. Mark off questions that are completed as you go with an "X".

Start
Greet the group according to the Introduction Script (page 3) making sure to mark the start time on the front of the booklet.

Question stage 1 (~4 minutes)

1. Ask, "Can you show me what you are doing right now?" Match it to a stage on the tracking sheet, if it is something you haven't asked about before, proceed to the page number in the interview booklet with those interview questions and ask them in order.
2. Ask about previous stages they have completed; use any extra time to go back and ask questions about things they already have done but haven't been questioned on before.
3. Ask questions from the "Additional Features" section if there is still time.

Question stage 2 (~2-3 minutes)
Ask the 4 daily questions.

Finish

1. Note the time on the students' device at the end of the interview and mark it on the front of the booklet.
2. Fill out the Interviewer Records on the appropriate page at the end of this booklet directly after finishing the interview.
3. Update the Interview Progress section.

Introduction Script

Day 1

Hi! You are _______ and ________, right? I'm _______ from the University of Michigan. I want to know what you think about the iPad program you are going to be using this week, so if it is ok I will come around and ask you some questions about it for a couple of minutes each day? Nothing you tell me will affect your grade in class and it is all confidential, so it would only be seen by the people who are creating this program.

No- Ok, that's fine. Have fun using the program this week!
Yes- Thanks! There are no right or wrong answers; everything you tell us will help improve something with the program. We just need your help to figure out what's good and what's not-so-good about it. For our research, we need one of you to wear an audio recorder while you are using the iPad each day. Would this be ok?

No- Ok, that's fine. However, because what you and your classmates say is important, I do need to interview groups who are OK with audio recording. Thank you for being willing to be interviewed though! Have fun with the program this week!
Yes- Ok, great! Thank you so much. I will see you tomorrow, then.

Day 2
Hey there/hello/good morning! Do you have any questions about what we're going to be doing? [questions] Okay, so what are you working on right now?

Following days:
Hey there/hello/good morning! So, what are you working on right now?

Standard triggers for more information

Usability Issue
- What are/were you trying to do?

Content issue
- (Echo probe) Repeat their response and probe for additional information on why

Classroom Issue
- Only probe if it seems to be hindering their ability to work over time (affecting them for more than 20 minutes in a non-trivial way)

Incomplete response – response with no causation
- "Oh? Why is that?" or "How come?"

3

Q: Can you show me what you are doing right now?

Working on tutorial

☐ What did you learn about creating explanations from the tutorial?

☐ What are some important things to include in scientific explanation?

☐ What parts of the tutorial were hard?
- Probe for complete answers ("How come" or "Oh? Why is that?")

☐ How can we make the tutorial better?
- Probe for complete answers ("How come" or "Oh? Why is that?")

Finding data

☐ Show me how you search through the data.
  - ☐ Why do you like that way of searching?
  - ☐ How else might you search through the data?

☐ What is hardest about finding the data you want?

☐ How do you determine if the information you and your classmates collected is good?

☐ When you are looking at the data, how do you pick out which data you want to use?
- If they give an answer saying they wanted animal X, Y, and Z and looked at that data, without saying how they chose the specific data
  - How did you pick the specific pieces of data for those animals that you used in your explanation?

☐ When you are looking through the class data, did you find the labels describing individual animals helpful?
- Yes -> How did they help?
- No -> How come?

☐ Did you listen to any audio notes?
- Yes -> Were they useful in helping to understand the evidence?
- No -> Why not?

Interpreting grouped data and curating tags

☐ Was grouping evidence helpful?
- Yes -> How did grouping evidence help?
- No -> Why not?

☐ How did you change the labels after you put the data into a group?
- If didn't change labels
    - How do you change the labels of the group? (if unsure, show them)

☐ Why did you choose the pieces of data that you have in this group (choose a group on screen)?
- Unsure (or other non-descriptive answer) *Pick a piece of data*: How about this piece of data?

Additional Features (Low priority)

Main Argument Page

☐ Has item with audio in explanation: Do you find it useful to playback audio on this page? Why or why not?

Investigation Guide

☐ What was most useful about the investigation guide?

☐ How do you use the investigation guide?
- What does the "Go Here" button do?

☐ Do you think there was anything missing in the investigation guide?

Looking up more information about an animal

Ranking Investigation

☐ Show me how you look up additional information
- *If unsure, show them the extra info tab*
- *If didn't need to show them the feature*: How did this extra information help you?

Survival investigation

☐ Did you try looking up additional information on the animal?
- Yes -> How did the additional information help you?
- No -> Why not?

Ranking Investigation

Making a claim

No reasoning yet

- [ ] How did you choose the ranking of your animals?

Has reasoning in their claim

- [ ] What is the purpose of making a claim?

Making reasoning

- [ ] What do you think the purpose of reasoning is?
- [ ] What were you thinking about while you were writing your reasoning?
    - *They have wrong or nonsensical reasoning*: How does your evidence support your claim?

Giving peer critique

- [ ] What do you think the purpose of peer critique is?
- [ ] Do you like peer critique? Why is that?
- [ ] What did you look at when giving peer critique?
- [ ] Do you think your critique for this investigation will be helpful? How come?

Reviewing/revising explanations (Post peer critique)

- [ ] Did you get useful feedback?
    - Yes -> How was it useful?
    - No -> Why wasn't the peer critique useful?
- [ ] Do you need to change your explanation after the peer critique?
    - Yes -> What do you need to change and why?

• No -> How come?

Survival Investigation

Making a claim

No reasoning yet

- ☐ How did you determine how well it would survive?

Has reasoning in their claim *and didn't ask before*

- ☐ What is the purpose of making a claim?

Making a reasoning

☐ What were you thinking about while you were writing your reasoning?

- *They have wrong or nonsensical reasoning:* How does your evidence support your claim?

Giving peer critique

☐ What did you look at when giving peer critique?

☐ Do you think your critique for this investigation will be helpful? How come?

☐ What would you change about how you give peer critique?

Reviewing/revising explanations (Post peer critique)

☐ Did you get useful feedback?
- Yes -> How was it useful?
- No -> Why wasn't the peer critique useful?

☐ Do you need to change your explanation after the peer critique?
- Yes -> What do you need to change and why?
- No -> How come?

<Page 9 of the interview booklet was left blank to accommodate interviewer notes.>

Daily open-ended questions

M	T	W	T	F	
					What is the most difficult thing you did today? How so?
					What did you feel successful at today?
					What is one thing you really liked that you did today?
					What is one think you would change about the system that you noticed today?

10

Interviewer Records

Record daily notes based on your conversation with the group you interviewed. These should be filled out directly after speaking with each group.

Monday

What was the overall mood of the group (happy, nervous, confused, upset at each other, neutral, etc)?

(Relative to tool?) Was there anything unusual or noteworthy about what came out of the interaction today? For instance, were there a lot of interruptions? How much probing did you have to do?

Overall thoughts on the group's progress (completing the task correctly? stuck? progress relative to rest of class?)

From observing the group through the class period, are they on task? Do any events stand out?

Other:

<Pages 11-15 of the booklet were the same questions, one page for each day of the post-activities during study.>

APPENDIX C: Student Pre-survey

First and Last Name: ________________________________

Circle A, B, C, D, or E in Scale					
A. Strongly disagree B. Disagree C. Neutral D. Agree E. Strongly agree					
Item	**Scale**				
1. I am good at science.	A	B	C	D	E
2. I would like to do more science at school.	A	B	C	D	E
3. I don't like working in groups in science class.	A	B	C	D	E
4. I have experience doing science outside of school.	A	B	C	D	E
5. I don't like science class when it is hard.	A	B	C	D	E
6. Science is boring.	A	B	C	D	E
7. I like it when activities in science class make me think.	A	B	C	D	E
8. I don't care how hard science class is as long as it is interesting.	A	B	C	D	E
9. I do as little work as possible in science class.	A	B	C	D	E
10. Complicated activities in science class are not fun.	A	B	C	D	E

Circle your responses below:

1) Have you ever used a mobile touch device (like an iPod Touch, iPhone or Android phone) before this week?

Yes / No

1a) If yes, how often:

every day / one a week / once a month / only once or twice ever

2) Have you ever used an iPad or other tablet before?

Yes / No

2a) If yes, how often:

every day / one a week / once a month / only once or twice ever

APPENDIX D: Student Post-survey

First and Last Name: ___________________________________

Circle A, B, C, D, or E in Scale					
A. Strongly disagree B. Disagree C. Neutral D. Agree E. Strongly agree					
Item	Scale				
1. I did not like giving peer critique with Zydeco.	A	B	C	D	E
2. I like learning science using Zydeco in the classroom.	A	B	C	D	E
3. The feedback from peer critique helped me revise my work.	A	B	C	D	E
4. I would like to use Zydeco in all of my science classes.	A	B	C	D	E
5. I didn't like creating science explanations this week.	A	B	C	D	E
6. I learn better with Zydeco than in my science class.	A	B	C	D	E
7.	A	B	C	D	E
8.	A	B	C	D	E
9.	A	B	C	D	E
10.	A	B	C	D	E

You worked with a partner building explanations this past week. How much of your group's work did you do? (circle one)

all of the work, most of the work, half of the work, less than half of the work, very little of the work

What are your least favorite things about using Zydeco?

What are your favorite things about using Zydeco?

Would you like to do more activities like this? (circle)

Yes / No / Maybe

Why?

APPENDIX E: Student Pre-Activities Worksheet

How is my animal related to other animals?

Your Name: ______________Your Learning Partner's Name: ____________________

Your group's favorite animal: __________________

Please Check Your Class Time Period: 10:20-11:30 11:30-12:45 12:45-2:35 (1:05-1:35 LUNCH)

You are going to explore more information about internal traits and external traits of your favorite animal and see how it is related to other animals. You can write down information and fill those into the table (on page 2-3) under appropriate category.

Note1: **Internal traits** are characteristics you CANNOT see from how the animal looks **when it is alive**. This can include skeletal traits and types of organs the animal has. **External traits** are some characteristics you CAN see from how the animal looks **when it is alive**. This can include the shape of an animal's ear, whether or not it has fur, and color of its skin.

Note2: Your animal may not have all the traits provided in the table. However, the more information you collect, the more you understand your animal!

Note3: If you think some traits are useful but don't fit to any category in the table, you can write it down in the "other" column.

Note4: You may find other traits information about animals relevant to your favorite animal, but be sure to indicate that animal's name.

Note5: These websites might be useful for you to explore your favorite animal:

1. Digital Morphology: http://digimorph.org
2. Animal Diversity Web: http://animaldiversity.ummz.umich.edu
3. Michigan Department of Nature Resources: http://ppt.cc/xzBt
4. The Hiker's Notebook: http://ppt.cc/y9uI
5. Environmental Education for Kids: http://ppt.cc/NPUp
6. Encyclopedia of Life: http://www.eol.org/
7. Discover Life: http://www.discoverlife.org

Traits of My favorite Animal- Internal Traits

Skull	Size of brain	Number of Holes	Number of Bones	Location of Nostril Hole
	(compared with its body size)			
Shape of Types of Bone (You may draw if it is easier to describe)	Jaw	Hip	Vertebra	Describe the bone's structure (i.e. are the solid or hollow)
Breathing Organ	Describe how your animal breathes (i.e. gill, skin, lung, or something else)			
Heart Structure	Describe how your animal's heart looks like (i.e. how many chambers..)			
Other				

Traits of My favorite Animal- External Traits

<table>
<tr><td rowspan="2">Body Covering</td><td colspan="5">Describe what covers on your animal’s body (i.e. fur, scale, feather, shell, or others)</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td rowspan="2">Teeth</td><td colspan="2">Number of Teeth</td><td colspan="2">Shape of Teeth
(You may draw if it is easier to describe)</td><td>Other</td></tr>
<tr><td colspan="2"></td><td colspan="2"></td><td></td></tr>
<tr><td rowspan="2">Limb</td><td>Number of limbs</td><td>Number of fingers/ toes</td><td colspan="3">Describe your animal’s limb structure
(i.e. paw, claw, web, or others)</td></tr>
<tr><td></td><td></td><td colspan="3"></td></tr>
<tr><td rowspan="2">Eye</td><td colspan="2">Describe your animal’s eye location
(i.e. on the side, on the front)</td><td colspan="3">Other</td></tr>
<tr><td colspan="2"></td><td colspan="3"></td></tr>
<tr><td rowspan="2">Average Body Length</td><td colspan="2">Use inches/feet</td><td rowspan="2">Average Body Weight</td><td colspan="2">Use pounds</td></tr>
<tr><td colspan="2"></td><td colspan="2"></td></tr>
<tr><td>Other</td><td colspan="5"></td></tr>
</table>

MAKE SURE you write down your animal's internal and external traits on Page 2 & 3.
This page allows you to write some other traits that are not provided in the table (just in case).

Other Internal Traits	Other External Traits

REFERENCES

Abbas, J., Norris, C., & Soloway, E. (2002). *Middle school children's use of the ARTEMIS digital library*. Proceedings of the Second ACM/IEEE-CS joint conference on Digital libraries, Portland, OR.

Altman, D. G. (1991). *Practical statistics for medical research.* Chapman and Hall.

American Association for the Advancement of Science. (2009). *Benchmarks for Science Literacy*. New York: Oxford University Press.

Ames, M., & Naaman, M. (2007). *Why we tag: motivations for annotation in mobile and online media*. Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA.

Balling, J. D., & Falk, J. H. (1980). A Perspective on Field Trips: Environmental Effects on Learning. *Curator, 23*(4), 229-240.

Bamberger, Y., & Tali, T. (2007). Learning in a Personal Context: Levels of Choice in a Free Choice Learning Environment in Science and Natural History Museums. *Science Education, 91*(1), 75-95.

Basu, S., Sengupta, P., & Biswas, G. (Under review). A scaffolding framework to support learning in multi-agent based simulation environments. *Computers & Education*.

Bell, P. (1997). *Using Argument Representations to Make Thinking Visible for Individuals and Groups*. 2nd International Conference on Computer Support For Collaborative Learning, Toronto, Ontario, Canada.

Bergman, E. (2000). *Information Appliances and Beyond: Interaction Design for Consumer Products* (1 ed.): Morgan Kaufmann.

Berland, L. K., & McNeill, K. L. (2009). *Using a learning progression to inform scientific argumentation in talk and writing*. Learning Progression in Science, Iowa City, IA.

Bilal, D. (2001). Children's use of the Yahooligans! Web search engine: II. Cognitive and physical behaviors on research tasks. *Journal of the American Society for Information Science and Technology, 52*(2), 118-136.

Bilal, D. (2002). Children's Use of the Yahooligans! Web Search Engine. III. Cognitive and Physical Behaviors on Fully Self- Generated Search Tasks. *Journal of the American Society for Information Science and Technology, 53*(13), 1170-1183.

Bischoff, K., Firan, C. S., Nejdl, W., & Paiu, R. (2008). *Can all tags be used for search?* Proceeding of the 17th ACM conference on Information and knowledge management, Napa Valley, California, USA.

Blumenfeld, P. C., Soloway, E., Marx, R., Krajcik, J., Guzdial, M., & Palinscar, A. (1991). Motivating Project-Based Learning: Sustaining the Doing, Supporting the Learning. *Educational Psychologist, 26*(3 & 4), 369-398.

Bollier, D. (2010). The Promise and Peril of Big Data: The Aspen Institute.

Borgman, C., Hirsh, S., Walter, A., & Galagher, A. (1995). Children's searching behavior on browsing and keyword online catalogs: The Science Library Catalog Project. *Journal of the American Society for Information Science and Technology, 46*(9), 663-684.

Bradsher, M., & Hagan, L. (1995). The Kids Network: Student-Scientists Pool Resources. *Educational Leadership, 53*(2), 38-43.

Cahill, C., Kuhn, A., Schmoll, S., Lo, W.-T., McNally, B., & Quintana, C. (2011). *Mobile learning in museums: how mobile supports for learning influence student behavior*. Proceedings of the 10th International Conference on Interaction Design and Children, Ann Arbor, Michigan.

Cahill, C., Lo, W.-T., Kuhn, A., Quintana, C., McNally, B., Schmoll, S., & Krajcik, J. (2011). *Student use of multimodal data and metadata tools during nomadic inquiry.* Proceedings of the 10th World Conference on Mobile and Contextual Learning (mLearn ’11), Beijing, China.

Cho, K.-L., & Jonassen, D. (2002). The Effects of Argumentation Scaffolds on Argumentation and Problem Solving. *Educational Technology Research and Development, 50*(3), 5-22.

Clegg, T., Bonsignore, E., Yip, J., Gelderblom, H., Kuhn, A., Valenstein, T., & Druin, A. (2012). *Technology for promoting scientific practice and personal meaning in life-relevant learning*. Proceedings of the 11th International Conference on Interaction Design and Children, Bremen, Germany.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Cosley, D., Baxter, J., Lee, S., Alson, B., Nomura, S., Adams, P., & Gay, G. (2009). *A tag in the hand: supporting semantic, social, and spatial navigation in museums*. Proceedings of the 27th international conference on Human factors in computing systems, Boston, MA, USA.

Druin, A., Foss, E., Hatley, L., Golub, E., Guha, M. L., Fails, J., & Hutchinson, H. (2009). *How children search the internet with keyword interfaces*. Proceedings of the 8th International Conference on Interaction Design and Children, Como, Italy.

Druin, A., Foss, E., Hutchinson, H., Golub, E., & Hatley, L. (2010). *Children's roles using keyword search interfaces at home*. Proceedings of the 28th international conference on Human factors in computing systems, Atlanta, Georgia, USA.

Falk, J. H., Moussouri, T., & Coulson, D. (1998). The effect of visitors' agendas on museum learning. *Curator, 41*, 107-120.

Finarelli, M. G. (1998). GLOBE: A Worldwide Environmental Science and Education Partnership. *Journal of Science Education and Technology, 7*(1), 77-84.

Fisher, D., DeLine, R., Czerwinski, M., & Drucker, S. (2012). Interactions with big data analytics. *Interactions, 19,* 50-59.

Fraser, D. S., Smith, H., Tallyn, E., Kirk, D., Benford, S., Rowland, D., & Fitzpatrick, G. (2005). *The SENSE project: a context-inclusive approach to studying environmental science within and across schools*. Proceedings of the 2005 conference on Computer support for collaborative learning, Taipei, Taiwan.

Gong, J., & Tarasewich, P. (2004). *Guidelines for handheld mobile device interface design.* In Proceedings of the 2004 DSI Annual Meeting.

Google inc. Google. 2012, from http://www.google.com

Griffin, J. (1998). *School-Museum Integrated Learning Experiences in Science.* PhD Doctoral dissertation, University of Technology, Sydney, Australia.

Guy, M., & Tonkin, E. (2006). Folksonomies: Tidying up Tags? *D-Lib Magazine, 12*(1).

Guzdial, M. (1993). *Emile: Software-Realized Scaffolding for Science Learners Programming Multiple Media.* Ph.D., University of Michigan, Ann Arbor.

Hearst, M. A., & Rosner, D. (2008). *Tag Clouds: Data Analysis Tool or Social Signaller?* Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences.

Hutchinson, H. B., Bederson, B. B., & Druin, A. (2006). *The evolution of the international children's digital library searching and browsing interface*. Proceedings of the 2006 conference on Interaction design and children, Tampere, Finland.

Hutchinson, H. B., Druin, A., & Bederson, B. B. (2007). Supporting elementary-age children's searching and browsing: Design and evaluation using the international children's digital library. *Journal of the American Society for Information Science and Technology, 58*(11), 1618-1630.

Jensen, J. J., & Skov, M. B. (2005). *A review of research methods in children's technology design*. Proceedings of the 2005 conference on Interaction design and children, Boulder, Colorado.

Jochmann-Mannak, H. E., Huibers, T. W. C., & Sanders, T. J. M. (2008). *Children's information retrieval: beyond examining search strategies and interfaces*. The 2nd BCS-IRSG Symposium: Future Directions in Information Access, London. http://doc.utwente.nl/65025/

John, B., & Kieras, D. (1996). Using GOMS for User Interface Design and Evaluation: Which Technique? *Transactions on Computer-Human Interaction (TOCHI), 3*(4), 287-319.

Kali, Y., & Linn, M. C. (2007). Technology-enhanced support strategies for inquiry learning. In J. M. Spector, M. D. Merrill, J. J. G. Van Merriënboer & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 445-490): Erlbaum.

Kali, Y., & Linn, M. C. (2008). Designing Effective Visualizations for Elementary School Science. *Elementary School Journal, 109*(2), 181-198.

Kang, R., Fu, W.-T., & Kannampallil, T. G. (2010). *Exploiting knowledge-in-the-head and knowledge-in-the-social-web: effects of domain expertise on exploratory search in individual and social search environments*. Proceedings of the 28th international conference on Human factors in computing systems, Atlanta, Georgia, USA.

Krajcik, J., & Blumenfeld, P. C. (2006). Project-based learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. New York.

Kuhn, A., Cahill, C., Quintana, C., & Schmoll, S. (2011). *Using tags to encourage reflection and annotation on data during nomadic inquiry*. Proceedings of the 2011 annual conference on Human factors in computing systems, Vancouver, BC, Canada.

Kuhn, A., Cahill, C., Quintana, C., & Soloway, E. (2010). *Scaffolding science inquiry in museums with Zydeco*. Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems, Atlanta, Georgia, USA.

Laru, J., Sanna, J., & Clariana, R. B. (2010). Supporting collaborative inquiry during a biology field trip with mobile peer-to-peer tools for learning: a case study with K-12 learners. *Interactive Learning Environments, 1*(15).

Linn, M. C., Clark, D. B., & Slotta, J. D. (2003). WISE Design for Knowledge Integration. *Science Education, 87*(4), 517-538.

Lizotte, D. J., Harris, C. J., McNeill, K., L., Marx, R. W., & Krajcik, J. (2003). *Usable assessments aligned with curriculum materials: Measuring expalantion as a scientific way of knowing*. annual meeting of the American Educational Research Association, Chicago, IL.

Luchini, K., Quintana, C., & Soloway, E. (2004). *Design guidelines for learner-centered handheld tools*. Proceedings of the SIGCHI conference on Human factors in computing systems, Vienna, Austria.

MacGregor, G., & McCulloch, E. (2006). Collaborative tagging as a knowledge organisation and resource discovery tool. *Library Review, 55*(5), 291-300.

Maldonado, H., & Pea, R. (2010). *LET's GO! to the Creek: Co-design of Water Quality Inquiry Using Mobile Science Collaboratories*. IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education.

Martin, S., Stanton Fraser, D., Fraser, M., Woodgate, D., & Crellin, D. (2010). *The impact of hand held mobile technologies upon children's motivation and learning*. 9th World Conference on Mobile and Contextual Learning, Malta. http://opus.bath.ac.uk/21478/

McNeill, K., & Krajcik, J. (2008). Inquiry and Scientific Explanations- Helping Students Use Evidence and Reasoning. In J. Luft, R. Bell & J. Gess-Newsome (Eds.), *Science as inquiry in the secondary setting* (pp. 121-134). Arlington, VA: National Science Teachers Association Press.

McNeill, K., L., Lizotte, D. J., Harris, C. J., Scott, L. A., Krajcik, J., & Marx, R. W. (2003). *Using backward design to create standards-based middle-school inquiry-oriented chemistry curriculum and assessment materials*. annual

meeting of the National Association for Research in Science Teaching, Philadelphia, PA.
McNeill, K., Lizotte, D., Krajcik, J., & Marx, R. (2006). Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *Journal of the Learning Sciences, 15*(2), 153-191.
McNeill, K. L., & Krajcik, J. (2011). *Supporting grade 5-8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing*. New York, NY: Pearson Allyn & Bacon.
Means, B. (1998). Melding Authentic Science, Technology, and Inquiry-Based Teaching: Experiences of the GLOBE Program. *Journal of science Education and Technology, 7*(1), 97-105.
Metz, K. E. (2000). Young children's inquiry in biology: Building the knowledge bases to empower independent inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into Inquiry Learning and Teaching in Science* (pp. 371-404). Washington, D.C.: American Association for the Advancement of Science.
Mohageg, M., & Wagner, A. (2000). Design considerations for information appliances. In E. Bergman (Ed.), *Information Appliances and Beyond: Interaction Design for Consumer Products* (pp. 27-51). San Francisco, CA: Morgan Kaufmann.
National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: National Academy Press.
Nuance communication. Dragon Naturally Speaking. 2012, from http://www.nuance.com/naturallyspeaking
Passani, L. (2002). *Building Usable Wireless Applications for Mobile Phones*. Proceedings of the 4th International Symposium on Human Computer Interaction with Mobile Devices, Pisa, Italy.
Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167-199.
Puntambekar, S., & Kolodner, J. (2005). Distributed Scaffolding Helping Students Learn Science from Design. *Journal of Research in Science Teaching, 42*(2), 185-217.
Puntambekar, S., & Kolodner, J. L. (1998). *The design diary: A tool to support students in learning science by design*. Proceedings of the International Conference of the Learning Sciences, Charlottesville, VA. http://www.cc.gatech.edu/projects/lbd/htmlpubs/ddtoolsupport.html
Quintana, C. (2001). *Symphony: A case study for exploring and describing design methods and guidelines for learner-centered design.* University of Michigan, Ann Arbor.
Quintana, C., Krajcik, J., & Soloway, E. (2002). *A Case Study to Distill Structural Scaffolding Guidelines for Scaffolded Software Environments*. Proceedings of the SIGCHI conference on Human factors in computing

systems: Changing our world, changing ourselves, Minneapolis, Minnesota, USA.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., & Soloway, E. (2004). A Scaffolding Design Framework for Software to Support Science Inquiry. *Journal of the Learning Sciences, 13*(3), 337 - 386.

Quintana, C., Soloway, E., & Krajcik, J. (2003). Issues and approaches for developing learner-centered technology. In M. Zelkowitz (Ed.), *Advances in computers* (Vol. 57, pp. 272-321). San Diego, CA: Academic Press.

Quintana, C., & Zhang, M. (2004). *The Digital IdeaKeeper: Extending Digital Library Services to Scaffold Online Inquiry*. American Educational Reserch Association, San Diego, CA.

Razikin, K., Goh, D. H.-L., Teng, Y. L., Nguyen, Q. M., Kim, T. N. Q., Lim, E. P., Sun, A. (2009). *Sharing Mobile Multimedia Annotations to Support Inquiry-Based Learning Using MobiTOP*. Active Media Technology, Berlin, Heidelberg.

Rivadeneira, A. W., Gruen, D. M., Muller, M. J., & Millen, D. R. (2007). *Getting our head in the clouds: toward evaluation studies of tagclouds*. Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA.

Rogers, Y., Price, S., Fitzpatrick, G., Fleck, R., Harris, E., Smith, H., & Weal, M. (2004). *Ambient wood: designing new forms of digital augmentation for learning outdoors*. Proceedings of the 2004 conference on Interaction design and children: building a community, Maryland.

Roschelle, J. (1995). Learning in interactive environments: Prior knowledge and new experience. In J. H. Falk & L. D. Dierking (Eds.), *Public Institutions for Personal Learning*: American Association of Museums.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*, 513-536.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-Driven Inquiry: Integrating Conceptual and Epistemic Scaffolds for Scientific Inquiry. *Science Education, 88*(3), 345-372.

Sedig, K., Klawe, M., & Westrom, M. (2001). Role of Interface Manipulation Style and Scaffolding on Cognition and Concept Learning in Learnware. *Transactions on Computer-Human Interaction (TOCHI), 8*(1), 34-59.

Shirky, C. (2005a). Folksonomies & Tags: The rise of user-developed classification. *IMC Expo*. New York.

Shirky, C. (2005b). Ontology Is Overrated. *O'Reilly Emerging Technology Conference*. San Diego.

Solomon, P. (1993). Children's Information Retrieval Behavior: A Case Analysis of an OPAC. *Journal of the American Society for Information Science, 44*(5), 245-264.

Songer, N. B. (2006). BioKIDS: An animated conversation on the development of complex reasoning in science. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 355-370). New York: Cambridge University Press.
Suthers, D. (1998). *Representations for Scaffolding Collaborative Inquiry on Ill-Structured Problems.* American Educational Research Association, San Diego, CA, USA.
Suthers, D., & Jones, D. (1997). *An Architecture for Intelligent Collaborative Educational Systems* AI-Ed 97, the 8th World Conference on Artificial Intelligence in Education, Kobe, Japan.
Suthers, D., Toth, E., & Weiner, A. (1997). *An integrated approach to implementing collaborative inquiry in the classroom.* 2nd international Conference on Computer Support for Collaborative Learning, Toronto, Ontario, Canada.
Tabak, I. (2004). Synergy: A Complement to Emerging Patterns of Distributed Scaffolding. *The Journal of the Learning Sciences, 13*(3), 305-335.
Tossell, C. C., Kortum, P. T., Shepard, C. W., Rahmati, A., & Zhong, L. (2010). *Assessing the Performance of Common Tasks on Handheld Mobile Computers*. Human Factors and Ergonomics Society Annual Meeting.
Tsandilas, T., & Mackay, W. E. (2010). *Knotty gestures: subtle traces to support interactive use of paper*. Proceedings of the International Conference on Advanced Visual Interfaces, Roma, Italy.
Vavoula, G., Sharples, M., Rudman, P., Meek, J., & Lonsdale, P. (2009). Myartspace: Design and evaluation of support for learning with multimedia phones between classrooms and museums. *Computers & Education, 53*(2), 286-299.
Wal, T. V. (2005). Explaining and Showing Broad and Narrow Folksonomies. 2012, from http://www.vanderwal.net/random/entrysel.php?blog=1635
Wal, T. V. (2007). Folksonomy. 2012, from http://vanderwal.net/folksonomy.html
Weiss, S. (2002). *Handheld Usability*: John Wiley & Sons.
Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17*, 89-100.
Woodgate, D., Stanton Fraser, D., Gower, A., Glancy, M., Gower, A. P., Chamberlain, A., & Crellin, D. (2010). Using mobile and pervasive technologies to engage formal and informal learners in scientific debate. In T. Goh (Ed.), *Multiplatform E-Learning Systems and Technologies: Mobile Devices for Ubiquitous ICT-Based Education* (pp. 196-214): Information Science Reference.
Yahoo inc. Flickr. 2012, from http://www.flickr.com