Connecting Visitors to Data: Exploring Tools for Mediating Learning Talk at an Interactive Museum Exhibit

BY

JESSICA A ROBERTS
B.S., Northwestern University, 2003

THESIS
Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Learning Sciences
in the Graduate College of the
University of Illinois at Chicago, 2016

Chicago, Illinois

Defense Committee:

Leilah Lyons, Chair and Advisor
Susan Goldman
Tom Moher
Robb Lindgren, University of Illinois at Urbana Champaign
Michael Horn, Northwestern University

This thesis is dedicated to my kids, Quentin and Violet, whose silliness and joy brighten my every day.

ACKNOWLEDGMENTS

The work presented in this dissertation was made possible by the tremendous support of colleagues, friends, and family. I wish to sincerely thank the many people without whom this work would never have been possible.

I first wish to thank my advisor, Leilah Lyons, whose advice and friendship over the years have helped me grow tremendously. Her creativity and her ability to draw on her extensive breadth of knowledge in the field have pushed our project, and my work in it, far forward. She, with her unmatched wordsmithing skills, could undoubtedly craft the perfect line about my zone of proximal development to put right here.

I also wish to thank my initial advisor, Josh Radinsky, for shepherding me into the field of learning sciences. Though our research interests ultimately diverged, the grounding he gave me in GIS and mediated action were foundational for this work and continue to influence my research. Josh also gave me the exceptional opportunity to see behind the scenes of academic publishing as the editorial assistant for the Journal of the Learning Sciences, an opportunity that was eye opening and gave me the invaluable chance to work with Iris Tabak and the rest of the JLS leadership team.

I am forever grateful to the other members of my committee for their help along the way. On multiple occasions Susan Goldman was able to give me straightforward feedback to help me clarify my work. I feel lucky to have had the opportunity to learn from Tom Moher and Robb Lindgren, the work of both of whom I admire greatly. I would especially like to thank Mike Horn, who made the mistake of hiring me during my dissertation writing year, giving me the opportunity to corner him and coerce feedback on my research way more frequently than he bargained for when he agreed to serve on my committee. His willingness to always take time to talk, and his confidence in me and my work truly helped get me across the finish line with this project.

The project itself would not exist were it not for the efforts of Francesco Cafaro, a great computer scientist and a great friend. Lukasz Stempniewicz built the handheld tablet controller, without which the 2x2 study presented here would not have been possible. The data collection was completed thanks to the efforts of Rebecca Eydt, who rose to every challenge (and there were many) to get the job done. Thanks also to the NYSCI Explainers Shanna Ramsamooj, Kumari Biswas, Jennifer Cumbe, Madiha Naeem, and Celina Muñoz for their participation in data collection. The project was supported by grants from the National Science Foundation, the National Endowment for the Humanities, and the University of Illinois' Institute for Policy and Civic Engagement.

My growth as a researcher throughout this program has been shaped by my fellow students and colleagues. In particular I would like to thank José Melendez and Jackie Popp, Ray Kang, Michael Bolz, Chandan Dasgupta, Katie James, Candace Burkett, Catherine Kaduk, Emma Hospelhorn, and Dane DeSutter. Each of you at some point (or many points) gave me guidance through peer review, sparked ideas through conversation, and made LSRI a friendly space to work. My prelimbo group members Brian Slattery, Priscilla Jimenez, and Joey Shelley have been helpful through every stage of this project, and Natalie Jorion gave me invaluable just-in-time guidance on statistics. As this dissertation neared completion, the editorial assistance from Abigail Goben and Amartya Banerjee clarified ideas as only fresh eyes can. Deana Donzal keeps the LSRI ship afloat and makes sure all our paperwork is in order, for which I am forever grateful, and more importantly she is a kind friend and selfless listener.

ACKNOWLEDGMENTS (continued)

I was taught to write by many great teachers throughout my academic career, but the greatest credit is due to my high school world and AP lit teacher Beth Weber. There are many things she taught me about the mechanics of writing, for example avoiding the passive voice, using descriptive nouns instead of generic "things," and not starting sentences with "there are," but most importantly she taught me how to craft an argument in writing. Her directive was simple: If you are writing a paper about the décor of a room, don't write a paper telling your audience that the carpet is blue; they can see for themselves that the carpet is blue. Write a paper telling them why it is blue, why it matters that it is blue, and how it relates to other blue or non-blue elements of the room. Since then I have examined all my work to ensure I was not writing a "blue carpet" paper, and I owe her a great deal of thanks for providing that guidance just as I was developing as a writer.

Finally, the support of my friends and family has been crucial through this process. Whether listening to me talk about my theoretical framework *again* or talking about anything else so I can remember what the real world is like, I am lucky to have my brothers Matt and Zach; my sisters-in-law Carlee, Ellen, and Trina; my friends Daniel, Steph, Roxy, Davora, and Ray; and my in-laws David and Beth. My ability to make it this far is due to the love and support I have received throughout my life from my parents, Val and Dave.

Most importantly, with the greatest possible appreciation and love, I thank my husband Anthony. His unwavering support through all the ups and downs of this process has pulled me through, and his sense of humor has kept me smiling. He is a wonderful father and a great friend, and he amazes me every day.

JAR

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

FIGURE **PAGE**

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS

APT	Actor Perspective Taking
GIS	Geographic Information Systems
MOC	Means of Control
DOC	Distribution of Control
IU	Idea Unit
OPRI	Operational/Researcher Intervention
FB	Full-Body
HH	Handheld
S	Single-input
M	Multi-input
NYSCI	New York Hall of Science
WIMP	Windows-Icon-Mouse-Pointer

SUMMARY

A study of an interactive data map museum exhibit was carried out using a 2x2 study design varying the *means of control* — whether visitors completed control actions through full-body movements and gestures or using a handheld tablet controller — and the *distribution of control* to one or more active participants. Visitors' dialogue as they interacted in groups of two or more using one of four versions of the interactive exhibit was analyzed to reveal the depth and content of their dialogue. Dialogue from 119 user sessions was coded for five categories of talk: manage, instantiate, evaluate, integrate, and generate. These categories comprised 30 subcodes that were then weighted according to their relevance to the exhibit's learning goals. The weights of each code applied to a session were summed to generate a content score for that session, and content scores were averaged and analyzed using statistical tests (t-tests, one-way ANOVAs, and two-way ANOVAs) to determine the impact of means of control and distribution of control on learning talk.

It was found that the handheld conditions produced the highest content scores, and that score differences could not solely be attributed to differences in group interactions with the exhibit, including the amount of time spent interacting, the amount of data rendered on the display, the amount of dialogue spent learning the controls, and the learning curve for understanding the exhibit's interactive features.

Dialogue was further analyzed for visitors' spontaneous adoption of a first-person actor perspective. Visitors in 54 of the 119 sessions used the actor perspective at least once, for a total of 125 actor perspective taking (APT) statements. These statements were analyzed qualitatively

to identify twelve unique applications of APT in dialogue, forming three self-to-data relationships.

1 INTRODUCTION TO THE PROJECT

1.1 Problem Background

As "designed spaces" for informal learning (NRC, 2009), museums aim to support social learning, inspire curiosity, and engage visitors with new ideas and phenomena that they can connect to their prior knowledge and understandings. Recent decades have seen a shift in the design of museum environments from a traditional transmission model—in which the museum presents content for visitors to absorb—to a dialogic model, encouraging visitors themselves to take a more active role in the learning process (Roberts, 1997). Exhibit designers and researchers have taken a particular interest in the role dialogue among visitors plays in museum learning. It is through dialogue that visitors with varying levels of expertise in an individual topic can interact within what Vygotsky (1978) calls their "zone of proximal development" to support each others' learning and meaning making as they discuss exhibits, providing explanations, inferences, hypotheses, and questions to each other (Zimmerman, Reeve, & Bell, 2008; Crowley & Jacobs, 2002; Atkins, Velez, Goudy, & Dunbar, 2009).

Museums have responded to the trend to engage visitors in dialogue in part by introducing interactivity into exhibits to provide visitors a more active experience. Early "interactives" consisted of low-tech devices like flip cards, levers, and switches, which often asked visitors questions or prompted them to make predictions related to exhibit content. As novel interactive technologies are becoming cheaper, more robust, and more accessible, museums are increasingly turning to technology-based interactives as tools to help engage people and provide memorable and valuable experiences. Touchscreen displays, for example, can augment classic quizzes with animations and video. Multi-touch tables can support collaborative group activities (Block et al., 2012; Lyons et al., 2015). Video and audio recording features can allow visitors to record responses to exhibit content, thereby joining a "conversation" with curators and other visitors (Devine & Bernstein, 2015). While technology-based exhibits possess great potential to augment the visitors' experience, they also carry the risk of disrupting the interpersonal interactions as visitors engage in "heads-down" activities on single-input devices (Hsi, 2003; Heath & vom Lehn, 2008).

Advances in "off the desktop" interfaces involving whole body or full body interaction have gained popularity in museums in recent years (Price, Sakr, & Jewitt, 2015; Williams, Kabisch, & Dourish, 2005; Abrahamson & Lindgren, 2014) as a way to keep visitors' "heads up." These designs are intended to draw upon affordances of embodied cognition theories, which suggest physical involvement with a task potentially supports learning (Abrahamson & Lindgren, 2014; Hornecker, 2011). Recent research has investigated an additional affordance of these full-body controlled systems in that they may—through the physical connection between the user and the system—promote the adoption of a first-person or *actor* perspective which shows promise for assisting learners in reasoning about the content in productive ways (Enyedy, Danish, & DeLiema, 2013; Roberts, Lyons, Cafaro, & Eydt, 2014).

These technology-based interactive exhibits are gaining prominence in a variety of museums, from art museums to cultural heritage sites to history museums, but they are an especially logical fit for science museums: increasingly the product of scientific endeavor is data, and the authentic practice of many scientists involves analyzing, manipulating, and visualizing large digital data sets. In order to convey authentic scientific practice, science museums need to be able to engage and connect visitors to data, potentially through the use of interactive data visualizations.

This confluence of interactive technologies, data visualizations, and informal learning presents a new challenge for the design and study of museum learning environments. Much work has examined how visitors learn in museums, particularly through dialogue (Crowley & Jacobs, 2002; Allen, 2002; Atkins et al., 2009,Gutwill & Allen, 2012; Kisiel, Rowe, Vartabedian, & Kopczak, 2012; Leinhardt, Crowley, & Knutson, 2003; Falk & Dierking, 2000; Ash, 2003; Zimmerman, Reeve, & Bell, 2008) and a growing body of work is addressing informal learning mediated by interactive technologies (Yatani, Sugimoto, & Kusunoki, 2004; Hornecker, 2008; Block et al., 2012; Correia, Mota, Nóbrega, Silva, & Almeida, 2010; Falcão & Price, 2009, Lyons et al., 2015), including with embodied interaction (Lindgren & Moshell, 2011; Lindgren & Johnson-Glenberg, 2013; Williams, Kabisch, & Dourish, 2005). A similarly large body of research has examined how people learn with graphical data representations (Friel, Curcio, & Bright, 2001; Uttal, 2000; Shah & Hoeffner, 2002; Libarkin & Brick, 2002; Fischer, Dewulf, & Hill, 2005; Glazer, 2011). While all these bodies of work speak to components of problem, in a real-world setting these components do not exist in isolation. The dialogue among visitors, the exhibit content, and the design of the both physical space and an exhibit's interactive features are all *mediational means* (Wertsch, 1998) contributing to the exhibit interaction experience, each affecting the learning taking place. In order to move forward with productive designs for engaging visitors with data representations in museums, informal learning researchers and designers need to understand how these mediations means work together—or in conflict—to affect learning outcomes.

The study presented here investigates how visitors' learning talk is mediated during interactions with one kind of technology-based exhibit, an interactive census data map display called *CoCensus*. This exhibit, part of a larger design-based research project described in detail below (Section 1.1.2), seeks to help visitors connect with complex census data. This study takes a sociocultural perspective on learning, focusing on visitors' dialogue with each other as they explore the presented data and attending to the "cultural tools" (Wertsch, 1998) mediating the interactions. This dialogue is assumed to be mediated by many factors, two of which are directly investigated here.

First, a 2x2 quasi-experimental study explores how changes to the control device for the interactive exhibit—that is, the means by which visitors are *physically* connected to the data—impact visitors' learning talk. This control device is what Vygotsky (1978) might call a "technical" tool mediating the interaction. Specifically, this study varies the *means of control* (MoC) for the interactivity—whether visitors use a handheld tablet or a whole-body interactive system to enact changes to the display—and the *distribution of control* (DoC) to a single user or multiple users. The second area of investigation is a *psychological* tool connecting visitors to the data: the perspectives visitors take as they make sense of the map display. Visitors' dialogue during interactions in the 2x2 study described above was analyzed for spontaneous usage of a first-person *actor* perspective (Brunyé et al., 2009). To situate the research questions, the subsections below provide background on data visualizations as tools for learning (Section 1.1.1, covered in greater depth in Section 2.3) and on the exhibit that is the focus of this study (Section 1.1.2). More on cultural tools and the rationale for exploring these particular two factors in this dissertation are discussed further in Section 1.2.

1.1.1 Visualizations for Representing Data

As complex data sets are increasingly products of modern science, technology-based data visualization exhibits are a natural fit for hands-on science centers. Data visualizations, commonly defined as "computer-based, interactive visual representations of data to amplify cognition" (Card, Mackinly, Shneiderman, 1999), have long been used by scientists in order to make sense of their large and complex data sets (Van Dam, 1992). The interactive nature of these visualizations facilitates manipulation in order for users to see patterns and phenomena not otherwise apparent. Visualizations designed for experts, however, are not well designed for collaborative exploration (MacEachren, 2005). Despite some recent work demonstrating that well-designed data visualizations show great promise for helping novice (non-expert) learners make sense of data in informal environments (Pousman, Stasko, & Mateas, 2007), the visualization tools typically used by experts are not well-suited to novice users (Lloyd, 2001; Marsh, Golledge, & Battersby, 2007).

One of the oldest and most commonly used type of visualization is the data map: a visualization that overlays geographically-referenced data onto a map through shading or representative symbols. These geovisualizations, typically created using Geographic Information System (GIS) software, traditionally have a single-user focus: they are employed by professionals at individual workstations in order to address a particular problem. Recently geovisualizations have been explored as an object of collaboration via interaction designs that support dialogue and coordinated activity, an approach dubbed "geocollaboration" (MacEachren, 2005; MacEachren & Brewer, 2004). While there seems to great potential for adapting geocollaboration design strategies for informal learning, little is known about how geocollaboration and other collaborative visualization interpretation can occur among groups of novices in informal, social "free-choice learning" environments like museums (Falk & Dierking, 2000).

All visualizations, and data maps in particular, require decisions to be made about what will be visualized and what won't, and how it will be presented. These decisions fundamentally impact how the visualization will be interpreted (e.g. Monmonier, 1991; McCabe, 2009; Shah & Hoeffner, 2002; Plass, Homer, & Haywood, 2009; Uttal, 2000). Providing a user agency in selecting and manipulating the representation affords the opportunity for her to see and play with data in new ways; rather than a limited representation demonstrating a narrow, curated narrative, the user has the opportunity to explore different representations, each with unique affordances for data interpretation and reasoning. This capacity for open exploration matches well to the nature of free choice learning environments, where learners are supported in choosing their own trajectory and exploring according to their own interests (Falk & Dierking, 2000). Entirely unconstrained interactions, however, can be detrimental for novice users, who can be overwhelmed and get lost in what Marsh, Golledge, & Battersby (2007) call "buttonology" where their focus is on clicking buttons to get things to happen rather than on content reasoning. Particularly in an informal learning environment like a museum, constraints on a visualization's interactivity are key for appropriately scaffolding the learning experience. The *CoCensus* exhibit, the context for this research, seeks to provide these appropriate constraints on both the content and the interactivity to afford productive engagement with census data by novice museum visitors.

1.1.2 CoCensus: An Interactive Data Visualization Exhibit

The research presented here is situated around a digital data map exhibit called *CoCensus* located at a mid-sized urban hands-on science center, the New York Hall of Science (NYSCI). This exhibit is part of a larger study examining the use of geographic information systems (GIS) data maps across three settings: in the museum, in classrooms, and online (NSF INSPIRE 1248052). While the classroom studies in this broader project are focused on more targeted learning goals as set out by the instructors, for example creating a community profile for a geographic area or identifying nuances to "the Black Community," (Radinsky, Melendez, & Roberts, 2012), the museum exhibit does not aim to convey a set narrative or outcome. Instead, the exhibit is designed to help individuals see themselves, or a reflection of themselves as defined by the census, in the data in order to enable them to compare their data with that of their companions and across time; to identify trends and hypothesize about the causes of those trends using outside knowledge; to relate the data to their lived experience in the geographic area; and to question the data itself and what it does and doesn't reflect about them and their identities. A productive learning interaction with this exhibit does not require acquisition of particular facts about the census or New York demographics, it involves productive discussion among companions about facets of the dataset they find relevant and interesting to their own lives and experiences. For this reason the exhibit is designed to be a multi-user system, so two or more visitors viewing their own datasets can work together to make sense of the data by comparing and contrasting datasets with each other (see Figure 1).

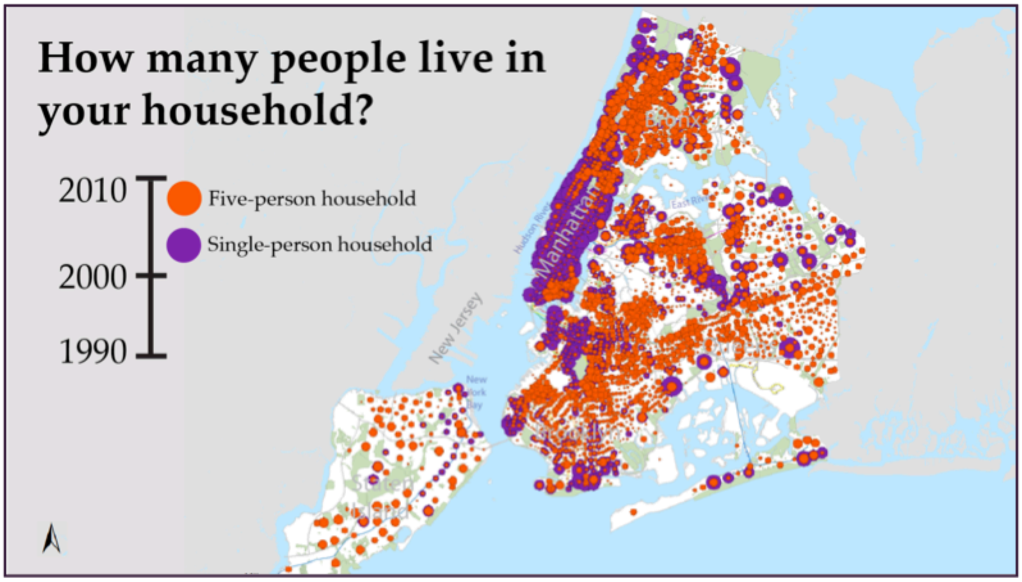


Figure 1. Screenshot of the CoCensus map showing single-person households and five person households in New York

The exhibit uses three techniques to help visitors make sense of what is otherwise a complex and dense corpus of data. First, the census data are displayed as scaled centroids (or "bubbles") overlaid onto a map of the local geography, New York City. This tactic allows visitors to relate the data to familiar places and lived experiences. The second tactic involved slicing down the large, overwhelming corpus of census data into manageable and personal pieces that visitors are able to select for themselves. To use the exhibit, visitors are asked to complete a mock "mini-census" survey at a kiosk outside the interaction area (see Figure 2). This survey consists of four questions adapted from the census and American Community Survey:

1. What is your ancestry or ethnic heritage?
2. How many people live in your household?

3. In what kind of house do you live?

4. In what industry do you work or want to work?

These categories of census data were selected because they are relevant to children and adults, they connect in some way to an individual's identity or lifestyle, and they are relatively self-explanatory. Visitors select answers that represent them or that they find personally interesting, which we hypothesized would help "hook" visitors in to exploring what can otherwise be an overwhelming and abstract dataset.

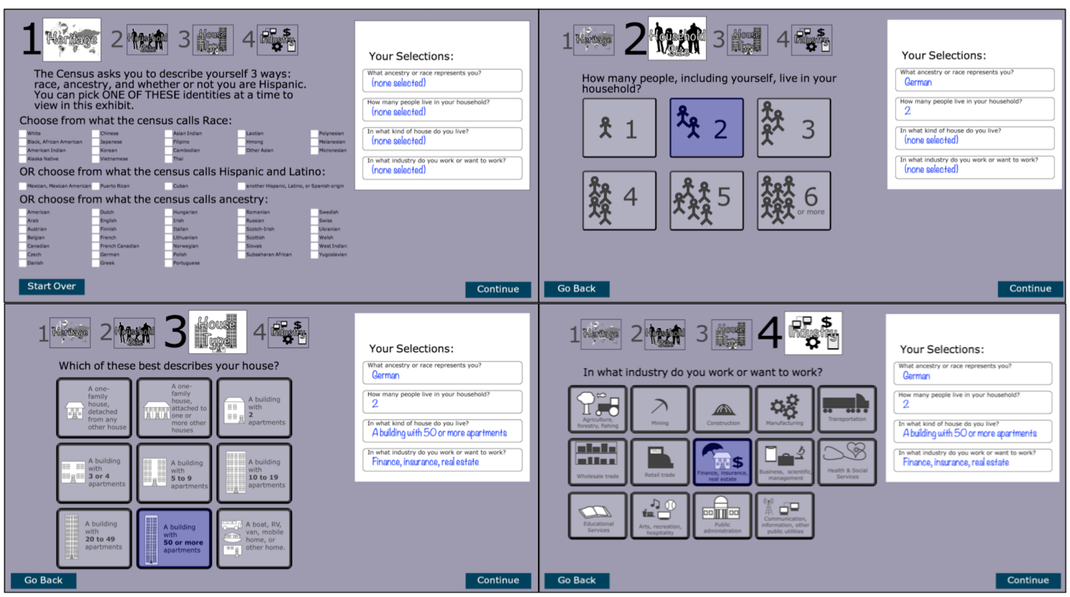


Figure 2. Screenshots of kiosk for creation of visitors' census profile.

The final tactic for connecting users to the data is to allow them control over the visualization. The broader design-based research program within which this study is situated has manipulated multiple aspects of the interactivity (Roberts et al., 2014; Roberts, Radinsky, Lyons, & Cafaro, 2012; Cafaro, Panella, Lyons, Roberts, & Radinsky, 2013). The iteration of the exhibit tested in this study affords three manipulations by visitors:

- Visitors can choose which of four *categories* of census data to view: heritage, household size, housing type, and industry;
- Visitors can choose which *census year* of data to view: 1990, 2000, or 2010;
- Visitors can choose the *aggregation* level at which to view the data: census tract, borough, or city-wide (see Figure 3).

The intent behind giving visitors the ability to manipulate the data representation was twofold. We anticipated that providing hands-on interaction with the data would support visitors' agency in exploring the content, and we wanted to allow the opportunity for different representations of data (i.e. different aggregation levels) to demonstrate how the same data can seemingly "say" different things depending on how it is represented.

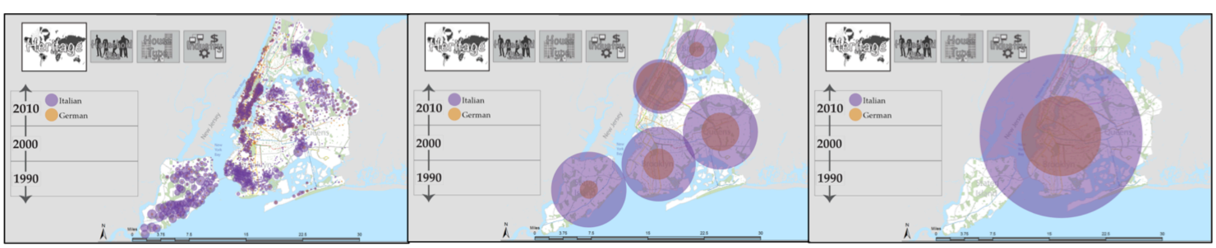


Figure 3. The CoCensus display at 3 aggregation levels: census tract (left), borough (center), city (right).

In order to engage multiple simultaneous visitors in playful interactions, CoCensus was conceived as a full-body multi-input system, in which multiple people could simultaneously change the data representations through movements and gestures in the physical space around the display (see Figure 4). The design team hypothesized that allowing multiple visitors the ability to make changes—that is, *distribution of control* (DoC)—would make the exploration engaging for more than one user, rather than leading to the isolation that can occur around single-input technologies (Heath & vom Lehn, 2008; see Section 2.2). Furthermore, we hypothesized that the use of full-body movements and gestures as the *means of control* (MoC) would further strengthen visitor engagement (see Section 1.3.3 for theoretical grounding): the movements would be fun to do, in keeping with the energetic atmosphere of the informal hands-on science center, and the visibility of control actions (compared to similar control actions on, for example, a handheld controller) would enhance the ability for co-visitors and bystanders to easily attend to changes in the visualization state, which in turn might support collaboration and dialogue. A more thorough review of these design strategies, and their origins in various bodies of literature, can be found in Section 2.2.

Figure 4. Visitors interacting with the full-body, multi-input control version of CoCensus at the New York Hall of Science. Here, two students examine tract-level industry data from the year 2000 while three of their classmates look on from the periphery.

Several iterations of the CoCensus exhibit manipulating different facets of the interaction design have been studied as part of this ongoing design research. The study presented here controls other aspects of the design (e.g. available data, design of the map display) to focus on effects of the MoC and DoC, independently and in interaction with each other, on visitors' learning talk. It also investigates more deeply the usage of *actor perspective taking* (APT) in mediating dialogue throughout the conditions.

1.1.3 Summary

In the exhibit studied here, the controller by which the map display is manipulated is a kind of cultural tool (Wertsch, 1998), what Vygotsky (1978) would call a "technical" tool, mediating visitors' learning experiences. Many interaction design decisions can assist with adjusting the pace of the learning experience, facilitating valuable visitor-visitor dialogue, and exposing properties of the data representations to learners for their investigation and interrogation. In this research, I claim that another cultural tool, *perspective taking*, may be present in CoCensus, and that it may play a role both in mediating learners' interpretations of represented data and in encouraging them to share their emerging interpretations with their companions. Literature demonstrating the affordances of perspective taking for promoting learning in other contexts (see Section 2.4) and pilot work on CoCensus (see Section 1.2; Roberts et al., 2014; Roberts et al., 2013) suggest that perspective taking may serve as a significant scaffold for assisting learners as they interpret new information. This research examines how these two kinds of tools mediate visitors' dialogue in their exhibit interaction experience.

The remainder of this chapter describes pilot work conducted on CoCensus to show the emergence of perspective taking as a phenomenon of interest, and establishes the theoretical framework on which this dissertation is based. Chapter 2 surveys existing literature from the three bodies of work converging in this research. The first two sections review work in museums to assess learning through dialogue and design strategies for engaging museum visitors through technology. These sections are followed by a discussion of the role of visualizations in learning and how other analyses have assessed map and graph interpretation. The chapter closes with a review of the ways perspective taking has been examined in other fields in order to situate this work within the broader, interdisciplinary context. Chapter 3 presents the methods of the quasi-experimental study conducted for this dissertation. Chapters 4 and 5 present analyses and findings; Chapter 4 is dedicated to the physical tool—the controller—and the impacts on variations of the *means of control* (MoC) and *distribution of control* (DoC) on visitors' productive learning talk, while Chapter 5 examines the role of the psychological tool of *actor perspective taking* (APT). The dissertation closes in Chapter 6 with a discussion of the implications of the presented findings and directions for future work.

1.2 Problem Statement

Pilot work on CoCensus has investigated the role of embodied control on visitor dialogue and perspective taking (Roberts et al., 2013) and the relationship between perspective-taking and data interpretation remarks in visitor conversations (Roberts et al., 2014). Effects of variations in the interaction design of CoCensus were explored in a December 2013 study examining affordances of two competing designs for the timeline control (Roberts et al., 2014). That study analyzed dialogues of pairs of users interacting naturalistically (i.e. in an unstructured session, without any mediation by a researcher) with one of two versions of the timeline floor control. In the Horizontal (H) configuration, small timeline "buttons" were placed parallel to the display, with the past (1990) on the left and current data (2010) on the right, in alignment with standard graphical conventions where time is often represented on the horizontal axis (see Figure 5, right). Two visitors could participate simultaneously, but a control action from one visitor (to change the decade or category) would change both visitors' data sets. This design—left to right timeline representations and mutually-exclusive timeline control—was meant to be *externally consistent* with common timeline representations with which visitors would be familiar. The alternative Vertical (V) configuration afforded separate simultaneous control of individual data sets, where stepping back (away from the display) moved back in time and stepping forward moved forward (to 2010), in an *ego-moving* metaphor expected to help visitors feel personally connected to the data. The two designs yielded differences in the amount of data talk engaged in by visitors, with the ego-moving V condition supporting more productive talk than the H condition (Roberts et al., 2014).

The importance of perspective taking in the CoCensus visualization environment emerged earlier in the design-based research project, upon initial incorporation of distributed, full-body control of the exhibit in early prototype testing. Compared to preliminary testing of a non-interactive prototype (Roberts et al., 2012), in the full-body interactive version some visitors were found to spontaneously use first-person pronouns when discussing the data displayed in the exhibit (Roberts et al., 2013). In pilot testing when the embodied interaction components were not in place (and all manipulations of the visualization were performed by the researcher), visitors spoke of the data exclusively from a third-person perspective (e.g. "The Germans are all over the North Side."). However, once visitors were individually controlling their selected data sets through physical movements, some spoke from the perspective of someone in the map, for example, "I'm along the Lake." Temporal analysis of these sessions, during which pairs of visitors engaged in semi-structured interviews with the researchers while interacting with the display, revealed that the use of this *actor* perspective (Brunyé et al., 2009) occurred in conjunction with visitors' body-based control movements within the interaction area. Moreover, the pair of visitors demonstrating the highest frequency of *actor perspective taking* (APT) were highly engaged in the data interpretation, making inferences about populations and posing questions about relationships between their data sets over time (Roberts et al., 2013).

Visitors' dialogue in that study was coded for *actor* (first-person) and *onlooker* (third-person) perspective taking statements and three categories of data talk: reasoning about time, reasoning about data sets, and spatial reasoning. Although in this pilot work the *actor* perspective taking was correlated (across all cases in both conditions, $N = 28$ individuals, 14 sessions) with a greater frequency of data talk statements (Roberts et al., 2014), that analysis did not take into consideration differences in the quality and complexity of those statements: a surface-level, descriptive statement was treated as equal to a rich inference. Furthermore, because the two designs were based on overarching metaphors, they varied multiple aspects of the interaction simultaneously, i.e. individualized versus mutually-exclusive control and configuration of the interaction area. Therefore, differences in visitor conversations cannot be directly attributed to a specific design element. These differences could have been affected by individuals' ability to control their own data sets independently (the *distribution of control*), or they could have been related to the more physically engaging floor configuration involving larger front-to-back movements (the *means of control*). This dissertation builds off that work to isolate the *means of control* (MoC) and *distribution of control* (DoC) in a 2x2 design in order to speak to the contributions of each for mediating visitors' learning talk. It further explores the function of *actor perspective taking* (APT) across all sections to understand whether and how visitors' spontaneous adoption of a first person perspective productively mediates dialogue.

Figure 5. Two timeline configurations tested in December 2013. The vertical configuration (left) allowed each user to select her own decade for which to view her data. The horizontal configuration (right) allowed either visitor to change the decade, but any changes affected both data sets.

1.2.1 <u>Research Questions</u>

This study investigates the following research questions:

RQ1: How do the *distribution of control (DoC)* – single-input or multi-input – and the *means of control (MoC)* – full-body or handheld – impact visitors' learning talk while interacting with an exhibit? Specifically:

RQ1A: How do the *means of control* (MoC) and *distribution of control* (DoC), separately and in interaction with each other, affect the content of visitors' dialogue?

Interaction design greatly impacts visitors' experiences with a museum exhibit. Specifically, prior work suggests that in general, giving users individual access to control features in an interactive experience or in particular, giving that control to them via highly embodied means, can impact both their interactions and their learning. However, the impact of each design decision on how visitors use and learn from the exhibit is unknown. This analysis will identify five categories of learning talk drawn from literature on informal learning and graph interpretation: *management* of the interaction with companions, *instantiation* of exhibit elements, *evaluation* of presented information, *integration* of multiple pieces of information, and *generation* of new ideas based on exhibit information and prior knowledge. It will then relate those categories of talk to the learning goals of the CoCensus exhibit in order to investigate variations among conditions in the content of visitors' dialogue.

RQ1B: What differences in the content of visitors' learning talk can be attributed to differences in interactions with the exhibit due to variations in MoC and DoC?

It is expected that variations in the controller design will affect multiple facets of the visitor experience in addition to dialogue, such as how much effort users spend learning the interactive system and how many control actions they make. These differences may in turn affect the learning talk; for example, time spent discussing the operation of the interactive features is time visitors cannot focus on making sense of and talking about the data. Differences in learning talk among conditions will therefore be analyzed in light of interaction differences to assess the extent to which variations in learning talk can be attributed to the MoC and DoC directly (i.e. embodiment and physical interactions with data) rather than indirect factors.

RQ2: What role does *actor perspective taking* (APT) play in mediating visitors' learning talk in this context? Specifically:

RQ2A: What is the relationship between APT and learning talk?

Pilot work for this project and other work investigating perspective taking suggest that first-person *actor perspective taking* (APT) might be related to visitors' productive reasoning about data. This analysis begins by comparing learning talk by visitors who used APT with those who did not to identify whether this pattern persists in this dataset.

RQ2B: What are the applications of APT in visitor learning talk?

In open-ended, spontaneous dialogue, visitors have been found to use APT in a variety of ways. This analysis examines those functions to understand productive uses of APT in this kind of data interpretation task.

1.3 Theoretical Framework

Like many studies of museum learning, this study takes a sociocultural perspective, meaning that the analysis views mental functioning as being situated within a cultural, historical, and institutional context, a stance that assumes that "analytic efforts that seek to account for human action by focusing on the individual agent are severely limited"(Wertsch, 1998). This perspective follows the Vygotskian view that humans virtually always learn through dialogues, asking questions, and negotiating meaning. Falk & Dierking (2000) point out that museums are ideal environments for capitalizing on the sociocultural context for learning because unlike formal environments like classrooms, "in the real world, … if you do not know the answer to something you want to know about, you ask for help, read about it, or in some way seek out ways to maximize your zone of proximal development. Free-choice learning in general and museum learning in particular are commonly marked by some sort of socially facilitated learning." Atkins (2009) agrees, saying, "we view one of the richest forms of learning in a museum to be evident in the patterns of discourse and activities that groups engage in - such as labeling, theorizing, predicting, recognizing patterns, testing ideas, and explaining observations. These patterned activities provide a structure through which visitors construct scientific ideas and, in doing so, learn what it means to participate in scientific activities." Falk & Dierking (2000) summarize, saying, "learning, particularly learning in museums, is a fundamentally social experience."

Research taking a sociocultural perspective on learning focuses on “meaning-making in the broad sense, which emphasizes social interaction and cultural symbols and tools as crucibles for appropriating and adapting forms of knowledge, values, and expression” (Schauble et al., 2002). Ash (2003) notes that “museums are ideal locations to test out socioculturally framed research questions,” providing the following definition of “sociocultural”: a social group or ensemble is engaged in an activity; this activity is collaborative and is “informed by the individuals who comprise it [and] yet the activity reciprocally informs the individuals/group; the social activity is mediated by tools, signs, people, symbols, language, and actions.”

I follow this definition and look, therefore, at visitors’ meaning making as a form of mediated action. The agents, in this case the museum visitors, are constructing meaning not individually, but through the interactions between themselves and their mediational means: other agents in the space, the exhibit and its control mechanisms, and the perspectives they employ during the interaction. The visitors are constructing the meaning, but they rely on the exhibit and their particular perspective(s) to mediate the meaning-making. That is, it is not the visitors, the exhibit, or the perspective that is constructing the meaning; it is the combination that brings it forth.

Wertsch (1998) gives the example of a pole vault to explain the principle of mediated action, saying that “it is futile, if not ridiculous, to try to understand the action of pole vaulting in terms of the mediational means (the pole) or the agent in isolation. The pole by itself does not magically propel vaulters over a cross bar; it must be used skillfully by the agent. At the same time, an agent without a pole or with an inappropriate pole is incapable of participating in the event.” Likewise, we cannot understand the learning of one museum visitor in isolation from the context of the visit, which consists of a variety of mediational means or cultural tools (I follow

Wertsch in using the two terms somewhat interchangeably) shaping the knowledge construction. All cultural tools are not physical, however. Wertsch describes psychological tools—for example the syntax for arranging mathematical problems as described in section 1.3.2 below—that mediate an experience as much as physical ones.

One such cultural tool that has been found to mediate learning in other contexts is *perspective taking* (see Section 1.3.1). By looking at a problem or situation from a particular viewpoint, for example by envisioning oneself as a first person "actor" in a scenario, learners can potentially draw upon a unique set of resources for reasoning (Enyedy, Danish, & DeLiema, 2013). For example, when children envision themselves as a ball moving across a surface, they can reason about the velocity and forces of friction of the ball in a different way than they would by just watching a ball. They are able to transform their physical bodies into "components in the microworld that structure students' inferences." (Enyedy, Danish, & DeLiema, 2013).

1.3.1 Dialogue for Mediating Learning

Collaboration and social interactions are key components of the museum learning experience, where pedagogical structures like formal lessons and assessments are replaced by dialogue, questioning, and explaining (Allen, 2002; Atkins et al., 2009; Ash, 2003; McManus, 1994), and the learning goals tend to be fluid and are rarely explicitly defined. The exhibit, people, language, and actions mediate the collaborative meaning making occurring during the interaction as "both individual and collective understandings are enhanced through the successive contributions of individuals that are both responsive to the contributions of others and oriented to their further responses" (Wells, 2000). Examples of such meaning making can include an explanation by one visitor to another, perhaps in response to a question; an argument between two or more visitors about the meaning of represented data; or a shared inquiry experience in which multiple visitors explore a particular phenomenon of interest together. This dissertation follows prior research on collaboration and collaborative knowledge building in assuming that learners are building new knowledge on top of and in relation to prior conceptions. These conceptions are not assumed to be innately shared among visitors, but the act of articulating their preconceptions and prior understandings into public statements is an important phase of collaborative knowledge building (Stahl, 2000). It is through the interactions between individual and shared understandings that meaning is collaboratively made. Therefore the measure of learning considered by this research is visitors' "learning talk" (Allen, 2002) produced during their interaction. Individual cognitive gains are not measured, nor are visitors subjected to content tests before or after the interaction.

1.3.2 Perspective Taking as a Cultural Tool

In order to study learning as a mediated activity, attention must be paid to the cultural tools mediating the interactions. Wertsch (1998) provides a range of examples of what can be considered cultural tools, from the pole in pole vaulting as described above (what Vygotsky (1978) would call a "technical tool"), to the vertical syntax of multiplication problems (a "psychological tool"). Wertsch explains that the math problem 343 x 822 is for most people solvable when oriented with one number on top of the other; the familiar syntax affords the calculations required to find the solution. This affordance is removed when the numbers are placed horizontally, making the calculation much more difficult. Though it is the "same" problem, the removal of the cultural tool of the vertical array removes our ability to easily solve the problem. It can be argued, then, that the syntax is itself doing some of the "thinking," and is mediating the problem-solving activity (Wertsch, 1998).

In this research I view *perspective taking*—the act of positioning oneself in a particular viewpoint relative to the data being interpreted—as a cultural tool similar to that of the mathematical syntax. Much like we learn to arrange multiplication problems in a particular way, we learn to take on different perspectives, "the capacity to imagine what another's thoughts and feelings might be" (Falk & Dierking, 2000). For example, it is a tool frequently drawn upon by teachers trying to help their students understand a particularly challenging problem: "Try to think of this another way, from the viewpoint of …" Utilizing perspective in this way allows it to mediate the reasoning engaged in by the student, and it affords a different means of interpreting the data.

The perspective taking phenomenon has been explored in a wide variety of contexts to contribute to an understanding of learning in cognition across disciplines (see section 2.4). Ochs, Gonzales, and Jacoby (1996) explored linguistic constructions seeming to blend perspectives of physicists trying to understand particle behavior, and Enyedy, Danish, & DeLiema (2013) have investigated perspective taking by students learning science. The tactic of shifting a learner to a first-person perspective through the use of technology has been shown to augment performance in procedural tasks (Lozano, Hard, & Tversky, 2006; Lindgren, 2012). This dissertation follows Filipi and Wales (2004), who performed conversation analysis of perspective taking to study direction-giving, by examining spontaneous perspective taking during dialogue. Viewing *actor perspective taking* (APT) as a cultural tool is in line with conclusions drawn by Ochs and colleagues, who determined the blending of identities "created through gesture, graphic representation, and talk, appears to be a valuable discursive and psychological resource as scientists work through their interpretations and come to consensus regarding research findings."

1.3.3 Embodied Learning

As advances in technologies allow computing to move "off the desktop," researchers and designers are exploring the interactional and educational potential of "systems relying on embodied interaction, body movement, tangible manipulation, and physical embodiment of data, being embedded in real space and digitally augmenting physical space" (Hornecker, 2011). These systems utilize sensors and cameras to track body movements and gestures, allowing physical involvement not possible with traditional WIMP (windows-icon-mouse-pointer) interfaces. Hornecker explains embodied interaction thusly:

> "Dourish highlights how embodied interaction is grounded and situated in everyday practice, constituting a direct and engaged participation in a world that we interact with. Through this engaged interaction, meaning is created, discovered, and shared. Embodied interaction is thus socially and culturally situated. But phenomenologies and ecological psychologies would argue that being situated also means being situated in a body: Your body affects your experience of the world, changing your viewpoint quite literally as well as your experience of the world in terms of what it allows you to do (Husserl's "I can's"). In this sense the physicality of our bodies is tightly linked with our experience of the physicality of our surroundings."(Hornecker, 2011)

Importantly, theories of embodied cognition and embodied learning suggest that this physical involvement may fundamentally impact how we experience a situation (Abrahamson & Lindgren, 2014; Hornecker, 2011). For example, studies suggest that learners utilizing a full-body simulation of kinematics were less likely to focus on "surface features" of the simulation compared to those using a desktop version of the same simulation (Lindgren & Moshell, 2011; see Section 2.2.4), and when the underlying concepts to be learned were logically mapped to the physical movements, learners were able to retain generative physics knowledge at a 1-week follow-up better after a high-embodiment condition compared to a low-embodiment condition (Lindgren & Johnson-Glenberg, 2013, citing Johnson-Glenberg, et al., 2013).

While it may seem intuitive that physically experiencing a concept from the scientific domain of physics would augment learning gains, it has also been argued that all cognition, even conceptual understandings of more abstract domains such as math, are grounded in bodily experience (Abrahamson & Lindgren, 2014). This dissertation explores the role of physicality in another abstract domain, that of represented data. This research tests two forms of embodied interaction: a *full-body* control condition in which visitors manipulate the representation through movements and gestures, and a *handheld control* condition in which the same manipulations are performed using a handheld tablet controller. Both handheld and full-body controllers are forms of "embodied interaction" (Dourish, 2001) and both are utilized in the same learning environment: CoCensus (see Section 1.1.2). Many studies have looked at the affordances of full-body, immersive simulations in comparison to more traditional instructional methods (e.g. Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014; Richards, 2012), but these studies include other immersive sensory stimulants besides the whole-body interactivity in the experimental conditions, such as sounds at different pitches corresponding to actions with acids (high pitches) and bases (low pitches) in a titration lesson (Johnson-Glenberg et al., 2014). This research keeps the environment and available interactions constant in order to isolate the role of physicality in mediating the experience.

A system utilizing full-body control allows participants to manipulate input through large movements and gestures (see Section 2.2.4), making the user's body itself a controller. An alternative and more traditional input device is a handheld controller such as a joystick or touchscreen tablet. These handheld controllers are still "embodied," but rely on smaller movements. These two means of control conditions both utilize congruous actions for control, but at different scales, e.g. users change the decade by pressing a button on the handheld tablet controller or by stepping into a floor "button" in the full-body condition (see Section 3.1.1 for complete descriptions of controller designs). The control actions were designed to map to each other so that they both mapped equally well (or poorly) to the data manipulations they produce. By comparing group interactions and dialogue (Lindgren & Johnson-Glenberg, 2013) within these two conditions, this research can further contribute to an understanding of the role of physicality in reasoning with data visualizations.

1.4 Importance of the Study

1.4.1 Intellectual Merit

Embodied, whole-body, and tangible interaction techniques are of great interest to researchers and designers, not only because they are fun and engaging, but because embodied cognition theories suggest they may positively impact learning in a variety of settings (e.g. Lindgren & Johnson-Glenberg, 2013; Johnson-Glenberg et al., 2013; Hornecker, 2011; Abrahamson & Lindgren, 2014). We know that to be educationally effective, physical actions should be logically mapped to the underlying learning goals (Lindgren & Johnson-Glenberg, 2013), but we don't know the extent to which the degree of embodiment impacts shared meaning making and whether congruent gestures in two embodied interfaces — *full-body* and *handheld* controllers — afford different kinds of dialogue and interactions by museum visitors (Malinverni & Parés Burgués, 2015). The work presented here will contribute to theories of embodied learning and to the discipline of tangible, embedded, and embodied interactions by addressing this knowledge gap.

Furthermore, as the gathering, analysis, and visualization of large data sets becomes increasingly central to modern scientific activity, museums seeking to model authentic scientific practices to their visitors need to attend to the specific challenges learners face with graphical data representations. Novel interactive technologies are potentially beneficial for visitors' use in interpreting data, but how design choices impact learning specific to data interpretation is not yet known, as most work understanding learning with data representations is situated in formal classroom environments. This work examines the confluence of embodied interaction, informal learning, and data reasoning to inform future design work in these areas.

Finally, while first-person *actor perspective taking* (APT) is known to be productive in some learning contexts, we know very little about its spontaneous use in conversation in an informal setting, and whether and how it is associated with productive learning talk. Understanding the affordances of perspective taking in this context has the potential to inform future attempts to foster human-data connections through museum exhibits.

1.4.2 Broader Impacts

As technologies afford greater and greater flexibility in how systems can receive input, designers of museum exhibits are increasingly exploring options for allowing multiple people to simultaneously control an exhibit through full-body novel interactive controls, rather than a more traditional single-input windows-icon-mouse-pointer (WIMP) interface. Observational studies of these single-input interactive technologies in museums have long shown that restricting input to a single user can limit the engagement of peripheral participants (e.g. Heath & vom Lehn, 2008; vom Lehn, Heath, & Hindemarsh, 2001), but we have little empirical evidence demonstrating precisely what effects this has on collaborative reasoning. Full-body interactive systems show great potential for supporting greater learning gains according to embodied cognition theories, but we do not yet fully understand how to best leverage them, particularly for the increasingly important task of facilitating interpretation of data visualizations. Due to the logistical, technical, and financial challenges involved in implementing multi-user full-body interactive systems, their direct impacts on learning interactions are of great interest to museum educators. This research presents an analysis comparing visitors' spontaneous dialogue while using single-input or multi-input versions of either the handheld or full-body condition in order to empirically discuss the benefits and pitfalls of each for mediating collaborative interactions.

The findings on visitors' use of perspective taking will help museum exhibit designers incorporate scaffolds into an exhibit to facilitate the types of interactions the exhibit is intended to support. If we understand how perspective can mediate shared data reasoning, we can better aid learners in interpreting and reasoning about complex data sets in both formal and informal environments. Docents may use these findings to structure their framing of an exhibit interaction; for example, if the findings show that the Actor perspective leads to the type of reasoning in which the docent is trying to engage visitors, the docent can better understand how and when to model that perspective for the visitors. Teachers can similarly benefit from this deeper understanding of the role of perspective taking in meaning making, both in the museum on field trips and in the classroom when engaging students in inquiry activities.

2 REVIEW OF THE LITERATURE

This dissertation is situated in an emerging problem space for informal learning research, as the convergence of increased use of interactive technologies for collaborative exhibits and the need to present data as the "stuff" of modern science creates new challenges for both the design of museum exhibits and the study of the learning occurring around them. This chapter grounds the analysis to be presented in this dissertation by describing how prior work has addressed the study of dialogue at museums (section 2.1), how interactive technologies have been designed for museums to facilitate learning and dialogue (section 2.2), and what is known about how visitors make sense of data visualizations in general and geographic information systems (GIS) maps specifically (section 2.3). The chapter closes with a review of how *perspective taking*, a cultural tool thought to play a role in visitors' exhibit interactions, has been studied in relation to learning at the individual and group levels (section 2.4).

2.1 Designed Spaces for Informal Learning

Museums are one form of "designed spaces" for informal learning (NRC, 2009), meaning they are created to engage learners with content and advance disciplinary understanding, but unlike formal environments with a designed curriculum and prolonged engagement, they "tend to be more fluid and sporadic" and are "typically experienced episodically, not continuously" (NRC, 2009). Museums have been called "free choice" environments (Falk & Dierking, 2000) because visitors are allowed to navigate freely according to their own interests.

Nevertheless, as an educational environment, museum visits are still expected to lead to learning outcomes, and museums are committed to enhancing their role as educational environments (Heath & vom Lehn, 2008). As such, more and more museums are progressing from the “cabinet of curiosity” model in which museums present privileged curatorial knowledge to the public to engaging, interactive spaces where visitors can shape their own narratives of the experience (Roberts, 1997). Instead of encountering room after room of silent objects in glass cases, visitors are immersed in halls filled with elaborate scenery and ambient noise, videos showing real-world research and interviews with practitioners, and displays asking questions and inviting button-pushing and other interactions to find the answers. Informal educators seek to “encourage visitors to step beyond the ‘browse mode’ that they typically adopt in museums” (Schauble et al., 2002) and engage deeply in inquiry and dialogue with each other and the exhibits. Designers of informal environments, then, are faced with the task of shaping exhibits to promote these activities while simultaneously focusing learners’ attention on the relevant content.

Complicating the designer’s task is the fact that they have to manage the attention of more than one learner. For many visitors, the social component of the visit is a key part of the experience, above the desire for strict knowledge acquisition (Ash, 2003). Whether accompanied by friends, family members, or an organized group such as a school field trip, visitors share their experiences with each other through dialogue and co-construct knowledge based on individual and shared prior knowledge and experiences. Paris (1997, as cited by Packer & Ballantyne, 2005) suggests five ways in which social interaction facilitates visitor learning:

1) People stimulate each other's imaginations and negotiate meaning from different perspectives;

2) The shared goal of learning together enhances motivation;

3) There are social supports for the learning process;

4) People learn through observation and modeling; and

5) Companions provide benchmarks for monitoring accomplishment.

Given that dialogue is important for supporting learning in museums, it follows that studying visitors' learning requires analysis of their dialogue as they interact with exhibits. Examples of frameworks for analyzing visitor dialogue are presented next.

2.1.1 Learning Talk

The analysis of dialogue among visitors can give greater insight to the nature of their collaborative meaning making experiences. The manner in which this type of conversation analysis is conducted varies greatly by researcher and research focus (Allen, 2002), but these studies are generally based in a sociocultural perspective, assuming that the deepest learning and engagement occurs within the dialogue among learners (Vygotsky, 1978).

During a museum visit, people engage in "performance indicators" and "significant behaviors" (Borun, Chambers, & Cleghorn, 1996) or "engagement behaviors" (Packer & Ballantyne, 2005) such as looking at displays, reading text, asking and answering questions, and engaging actively with exhibits. Such behaviors give a picture of what visitors are doing during their trip, and some of these behaviors have been linked to learning. For instance, Borun et al. (1996) triangulated video and audio recordings and post-visit interviews to link observable behaviors of family groups on a museum visit with the groups' learning about the exhibit content. They found that the frequency of certain "significant behaviors" to be a distinguishing factor between family groups in successive learning levels. These significant behaviors were:

- Ask a question
- Answer a question

- Comment on the exhibit, including explaining how to use the exhibit (for interactives)
- Read the label aloud
- Read the label silently

These categories, along with movement tracking and counting heads, are common in the museum learning literature, with some adaptations. Diamond, Luke, & Uttal (2009) caution, for example, that "you may not be able to verify whether a visitor read the exhibit label unless he or she read it aloud" and therefore advise the use of 'look at label' instead of 'read the label silently.' Regardless, these types of observations can give some insight to how the groups are interacting with each other and the exhibit and may be indicators of the learning occurring in the groups.

In looking for "learning talk," Allen (2002) analyzed groups rather than individuals, and utilized a framework drawing upon categories for affective, cognitive, and psychomotor learning. The coding scheme divided 16 subcategories of codes into five categories: Perceptual, Conceptual, Connecting, Strategic, and Affective. These codes and subcategories were emergent based on the dialogue in order to capture the nature of the conversations, and were iteratively refined until all researchers agreed they captured the evidence of visitors' learning. The coding scheme utilized by Atkins et al. (2009) bears resemblance to Allen's codes, but includes the categories "Navigation," "Creating and noticing data," "Experimenting," and "Affect," each with two to eight subcategories.

Together these coding schemes illustrate threads common to the study of visitors' dialogue in museums, specifically the need to attend to how visitors are managing use of the exhibit ("Strategic" codes for Allen, "Navigation" for Atkins) and how they are connecting and engaging with the content. This dissertation adopts the term "learning talk" from Allen (2002) to encompass all productive dialogue produced by visitors during their interactive sessions. The five categories of codes utilized in analyzing visitors' dialogue for this study (see section 4.3) are derived from the categories described here and informed by work on graph interpretation (section 2.3, below).

The needs of museum exhibits to present focused content in accessible, engaging ways for individuals and groups of learners presents a significant challenge to designers of informal learning spaces who need to balance fun, social interactions with specific learning outcomes. The remainder of this chapter will illustrate how some technology-based exhibits and activities have attempted to foster visitors' connections to presented content (Section 2.2). It will then discuss challenges specific to data visualizations and data reasoning (Section 2.3) and conclude with a discussion of the potential affordances of a particular kind of connection, *actor perspective taking,* for engaging visitors (Section 2.4).

2.2 Design Strategies for Facilitating Engagement

Many attempts have been made to promote visitors' engagement with exhibit content. Arguably as important as the individual reasoning visitors are doing around the exhibit, however, is the social nature of the experience. Museums are unique learning environments because they rely so heavily on interpersonal interactions; in a museum environment, "the content may be less important than practice at engaging all group members in a shared scientific activity. A key issue here is whether or not visitors create and investigate their own questions, engage the entire group, and express delight"(Atkins, 2009). Exhibit designers want to promote active, prolonged engagement by visitors, more than just pushing a button and moving on (Allen & Gutwill, 2004;

Horn et al., 2012). We know from decades of research on the design of learning environments and from studies specifically situated in museums that seemingly small design decisions can have large impacts on the dialogue and interactions taking place around an exhibit. For example, Atkins et al. (2009) recorded visitor interactions around two versions of a heat camera exhibit at a science center. The functionality of the exhibit was the same in both conditions — a screen showed how objects look through an infrared camera — but one version provided mittens for visitors to try on and included brief explanatory text inviting them to observe differences in how each trapped heat. The alternate version had no such framing activity and just allowed for open explanation. Though in each case the exhibit did the same thing, what the visitors tried to do with it varied greatly. The "mittens" version of the exhibit yielded more classroom-like talk, in which one visitor, typically a parent, tried to instruct others about the lesson and what you were "supposed" to do and learn with the exhibit; by contrast, the "no mittens" version prompted exploration and "creating data" by spontaneous investigations of other objects in the space (Atkins et al., 2009). The little bit of framing led to qualitatively very different interactions. Neither type of dialogue is inherently good or bad, but the differences need to be kept in mind depending on the kinds of thinking and talk an exhibit is intended to promote.

The intricacies of preserving desired social interactions become even more pronounced when attempting to incorporate interactive technologies, which can result in "heads down" experiences (Hsi, 2003) that are too constrained and inhibit social engagement (Heath & vom Lehn, 2008). This section highlights some of the strategies that have been used for engaging visitors with content, including screen-based interactives (Section 2.2.1), handheld games for groups (Section 2.2.2), multi-user touchscreen devices such as tabletops (Section 2.2.3), and spatially indexed interactive experiences to provide an overview of how these technologies have been integrated into museum experiences and their affordances in supporting dialogue.

2.2.1 Screen-based Interactives

Traditionally exhibit labels have shouldered the burden of conveying relevant information to visitors, taking care to strategically provide sufficient information within reasonable length limits. As these labels need to be useful to a wide range of visitors, this is no easy task. However, with the emergence of affordable desktop and tablet computers that can be mounted in exhibit halls, supplemental information can easily be made available to visitors on demand. Many exhibits incorporate computer-based displays to provide extra information via text or video, to quiz visitors on domain knowledge, and to present simulations illustrating the concepts. One such example is the Sickle-Cell Counselor (Bell, Bareiss, & Beckwith, 1993), a computer program placing visitors in the position of a genetic counselor advising couples on the risk of genetic disease in their offspring. Visitors can perform simulated lab tests and "ask questions" of experts by viewing pre-recorded videos in order to make their recommendations. This Goal-Based Scenario design provides users with an easily recognizable goal (to advise the couple) and provides opportunity for deep engagement with the data through authentic practices (lab tests and consultations with experts).

Heath & vom Lehn (2008) note that the single-user input design exemplified by Sickle Cell Counselor is prevalent in many computer based museum interactives, which subscribe to a model that "prioritizes the single user and disregards the socially organized interaction that underpins the use of technologies… it is as if the design of exhibits presupposes a neutral domain that consists of a series of isolated individuals and individual actions, who at best are prepared to wait their turn and if necessary become a passive audience" (Heath & vom Lehn, 2008). Hornecker (2010) noted a similar sense of isolation among visitors using *Juroscopes,* media-augmented telescopes at a natural history museum. When pointed at dinosaur skeletons, these Juroscopes played 30-second animations in which the skeletons "grew" organs and skin to look like live dinosaurs in their natural environments, moving, feeding, and hunting (Hornecker, 2010). These individual devices were popular and described as "exciting" as they provided an immersive experience, causing some visitors to cringe and pull back from the viewfinder when the dinosaur "noticed" them. But the isolated experience made conversations among visitors disconnected. As Hornecker notes, "Being virtually alone adds to the feeling of a direct and personal experience. It also means it is difficult to share." Dialogue around these devices was compared with a "barrier-free version" of the same animations that utilizes a large, angled screen to show the animations. These screens were intended for wheelchair users and children too short to reach the Juroscopes but have come to attract larger crowds of visitors of typically 6 to 15 people who were "not just waiting, but actively watching and commenting, scaffolding and negotiating use of the lever mechanism (for selecting the dinosaur to be animated)" (Hornecker, 2010). The barrier-free versions lost some of the excitement of full immersion but gained social interactivity.

Other research supports the idea that simply enlarging the display for some technology-based exhibits removes one of the isolating components described above: a large display provides a shared output, letting multiple visitors engage more easily (DiPaola & Akai, 2006). For example, Bao & Gergle (2009) manipulated visual display sizes (large wall-sized vs. small desktop) and tasks (object identification and narrative description) and found significant differences in user references. Specifically, they found that presenting visual information on large screens was associated with significantly increased usage of deictic pronouns (e.g., "this," "that") when users produced open-ended narrative descriptions of the visual data. From these findings, they posited that producing narratives about information presented in large displays may "yield a more immersive experience, or greater sense of presence, that translates into measurable differences in language use" (Bao & Gergle, 2009).

While screen-based displays can provide unprecedented amounts and forms of interactivity and information, designers must attend to the inherent tradeoffs between highly immersive but exclusive individual experiences and group engagement in a less personalized interaction. Informed by this work, the *CoCensus* exhibit utilizes a large, shared screen visible to multiple participants and bystanders and attempts to maintain personalization through interactivity and customization.

2.2.2 Structured Games for Groups of Visitors

In addition to the physical components of the exhibit itself, designers can impact interpersonal interactions by structuring a designed activity. One tactic for engaging learners with content that is particularly focused on maintaining the social component of the visit is that of proving a structured game or task for the group to engage in together. For example, *Inquiry Games* for family groups (Allen & Gutwill, 2009) and student groups (Gutwill & Allen, 2012) gave visitors a "crash course" of inquiry methods on the museum floor by providing simple guides for conducting inquiry-based activities. In pilot work families were taught a sequence of skills at an exhibit before asking them to apply those skills at a novel exhibit (Allen & Gutwill, 2009). The propensity of families to forget the sequence and abandon some skills prompted the researchers to refine the inquiry skills to *Proposing Actions* and *Interpreting Results* (Gutwill &

Allen, 2012). It was determined that "these two skills complement students' natural exploration activity at exhibits, are intellectually accessible to diverse groups of students…, and are simple enough for students to understand quickly and remember easily." These skills became the foundation of two "games": *Juicy Questions* and *Hands Off*. While these games, particularly *Juicy Questions*, were found to positively impact visitors' engagement and learning, the required investment of time and resources for the researchers to teach the games to visitors makes widespread use of the tactics impractical.

Some work has been done to incorporate technologies that encourage visitors to collaborate with each other while actively engaging with the presented content. For example, the Musex system (Yatani et al, 2004) challenged pairs of students to answer questions about exhibits, where each correct answer revealed part of an image serving as a clue to answer a final question. The designers anticipated this strategy would increase students' motivation to correctly answer as many questions as possible. While user testing showed this system *did* lead to an increase in interactions with non-interactive exhibits, interactions between students was largely limited to task-completion talk. Similarly, the Donation system (Yiannoutsou et al., 2009) provided pairs of students with hints related to pieces at an art museum with the task of helping an imaginary benefactor make a donation. The system was intended to encourage collaboration among teams by limiting the number of hints each pair received. Findings demonstrated task-oriented behavior dominated interactions in this study as well, though participants did seem to come away with concrete and clear ideas about the exhibits featured in the game.

These structured games did encourage visitors to interact with each other, but designers must take care not to over-script the games at the expense of visitors' own inquiry. The technology-based games described here by necessity draw visitors' attention to particular exhibits and particular features, which may have the unintended consequence of restricting the "free-choice" nature of the space. While they have the benefit of making otherwise static exhibits (such as artifacts and paintings) "interactive," the inability to help visitors personally connect to the exhibits may mitigate one of the key benefits of informal learning environments. More open-ended activities such as Inquiry Games demonstrate that design can be used to promote inquiry and exploration, but the reliance on instruction cards and special training suggests that designers should find a more integral way to encourage questioning and shared inquiry via exhibit design.

2.2.3 Multi-user Inputs and Outputs for Shared Interaction

Multi-user input presents an open challenge: traditional technologies are designed with one input, for example a mouse pointer, for control. In social museum environments where groups of visitors want to engage simultaneously, technologies must be adapted. Interactive multi-touch tabletops are gaining popularity in museums (e.g. Hornecker, 2008; Horn et al., 2012; Correia et al., 2010) and show potential for engaging multiple users with complex phenomena, though not without significant challenges (Marshall, Morris, Rogers, Kreitmayer, & Davies, 2011; DiPaola & Akai, 2006). Their large screens provide access to multiple users, but input gestures must be carefully calibrated to accommodate multiple simultaneous touches. For example, Block et al. (2012) describe a tabletop display designed to illustrate the Tree of Life, the phylogenic relationship of all life on earth. The iterative design sought to allow museum visitors — assumed to be novices with respect to phylogenetic content — to interact with the tree in order to discover, among other things, relationships among organisms. In configuring the design, researchers drew upon familiar touchscreen devices such as the iPad to create a suite of interactive actions, but these gestures had to be adapted to compensate for the display's inability to distinguish between touches of different users, and at the same time to be discoverable and useable for participants.

Another consideration to address in the design of multi-user systems is the extent to and manner in which visitors could disrupt each other's interactions. Some such "interference" has been shown to create a productive, collaborative environment (Falcao & Price, 2009), but productivity breaks down when one user "hogs" control of an exhibit, excluding others (vom Lehn, Heath, & Hindemarsh, 2001).

2.2.4 Indexing Control to Physical Space

The final design strategy to be covered here is the tactic of employing physical space as part of the interaction. As motion tracking technologies from location sensors to cameras like the Kinect become cheaper and more readily available, educational designers are utilizing them in order to incorporate space and physical movements into the interactive experience, treating space "not simply as a container for our actions, but as a setting in which we act" (Williams, Kabisch, & Dourish, 2005). This trend is in line with theories of mediated action, which consider the physical context (Falk & Dierking, 2000), what Wertsch (1998, citing Burke, 1989) calls the "scene," as an integral component of the mediated action. By situating the control of an interactive museum exhibit in visitors' own bodies, allowing them to control the exhibit through movement and gestures, exhibits utilizing these technologies may gain multiple affordances related to embodied learning (see Section 1.3.2), visibility and awareness of the presence and actions of co-visitors, and distribution of control among multiple users in the space (Section 2.2.3).

One example of such a system is SignalPlay (Williams, Kabisch, & Dourish, 2005), an installation inviting users to move large physical objects around an interaction area in order to manipulate the sound-scape created by the system. Each object resembles a familiar, everyday artifact, such as a compass, blocks, or chess pieces, and manipulation of each produces an individual discrete sound and a system-wide audio effect on the collective sound-scape, such as a change in tonal harmony, tempo, or timbre, but users have to experiment with the objects in order to understand precisely how. The controller objects are large enough that they cannot all be used by a single user at once, so that multiple people have to coordinate actions if they want to achieve a desired effect. Users were found to begin with an "iconic" mode of use, in which they used the pieces the way their iconic objects would be used, for example by walking around the room with the compass as if way-finding. As they gained familiarity with the interaction, they progressed to "instrumental operation" in which objects were used as instruments to produce sound effects, "people … reach 'through' the objects, focusing on using them as controllers of a digital system" (Williams et al., 2005). As people interacted with the system, Williams et al. report bystanders watching participants to learn how to instrumentally operate the pieces — the visibility of the interaction helped others successfully interact. Moreover, the pieces invited groups of visitors to collaborate and experiment in achieving sounds, such as playing a game of "speed chess" in order to hear the effects of rapid (non-legal by chess rules) movements of the pieces.

Another exhibit utilizing space to augment the social learning experience is MEteor (Lindgren & Moshell, 2011), an immersive simulation game in which learners enact movements of an asteroid traveling through space. Participants use their whole bodies to predict the movements of asteroids as they approach planets and other objects with gravitational forces.

Through increasingly difficult levels of play, learners are subjected to complex configurations "designed to disrupt pre-existing misconceptions and give rise to new perspectives that have the potential to be built upon with formal instruction" (Abrahamson & Lindgren, 2014). While the use of physical space helps immerse learners in the experience, this intervention is designed to be augmented with formal instruction and relies on docent facilitation to teach learners the underlying laws of physics.

Such spatially-indexed interactions are well suited for museum exhibits, where "Encounters… are collective experiences, and individual actions around them are organized with regard to the presence, orientation, activities, and gaze of others" (Williams et al., 2005). Moving the locus of interaction out of a screen or individual controller and into the shared exhibit space affords joint attention and collaborative exploration, which can contribute to the shared learning experiences. This mode of interaction, though promising for its ability to maintain the social, interactive component of an exhibit, presents challenges: the lack of visible controls threaten the discoverability of the system, and designers must implement a suite of gestures that are engaging and fun without making visitors feel silly.

2.2.5 Summary

Promoting deep engagement with an exhibit is a known challenge, and many strategies have been employed with varying degrees of success. Structured games on handheld devices (Yatani, 2004; Yiannoutsou et al., 2009) can focus learners on important content, but can also limit meaningful engagement if users are too focused on completing the task and finding the right answers. This is especially true if those answers are based primarily on label information, allowing the learners to ignore the artifacts completely. While small, individually controlled computers can engage single users with deeper content (e.g. Bell, Bareiss, & Beckwith, 2009), such displays can limit the social interactions among visitors (Heath & vom Lehn, 2008). Using large, shared displays can mitigate the isolating effects, but the mode of input is important: single-user inputs such as a mouse similarly limit interactions. Multi-user inputs such as multi-touch tabletops (Block et al., 2012) and embodied interaction displays (Lindgren & Moshell, 2011) are gaining popularity as advances in technology allow for such devices to distribute control among a larger number of simultaneous users, but designing interactives that support visitors in engaging with each other and the content is still an open problem in informal environments. This research makes steps toward exploring this problem space.

2.3 Reasoning about Data through Visualizations

This section shifts from the form factors supporting learning in museums to the content that is to be learned. Understanding data is an important 21st century skill. When dealing with large data sets, for example the United States census, the sheer size and complexity can be overwhelming to people not familiar with this type of data. Organizing the data into visualizations is one way of scaffolding interpretations and reasoning, but making sense of visualizations is a non-trivial task for learners. This section discusses the processes known to impact learners' interpretation of and reasoning about a variety of data visualizations in order to explicate design challenges specific to data visualizations and outline a framework for analysis of visitors' interpretations of data visualizations.

In order to conceptualize the way learners make sense of data maps, it is useful to first examine literature on other types of data visualizations, especially graphs. While some differences in interpretation schema are known to exist between maps and graphs (Schiano &

Tversky, 1992), the frameworks for conceptualizing how learners decode graphs is highly relevant to decoding data maps, as illustrated below.

2.3.1 Graph Interpretation

Much work has been done to understand how people understand and interpret scientific graphs. Shah & Hoeffner (2002) provide a review of empirical studies on the matter, describing three major component processes of graph interpretation: 1) encoding visual information and identifying important visual features; 2) relating visual features to the conceptual relations they represent; 3) determining the referent of the concepts being quantified and associating those referents to the encoded functions. The three main factors affecting these interpretive processes, the authors continue, are 1) the characteristics of the visual display; 2) the viewer's knowledge of graphical schemas and conventions; 3) the content of the graph and the viewer's prior knowledge and expectations about the content.

Regardless of the content being depicted, it is widely acknowledged that the purpose and audience must be kept in mind when designing a graph or visualization (Libarkin, 2014; Shah & Hoeffner, 2002; Glazer, 2011; McCabe, 2009, Few, 2009, Fischer, Dewulf, and Hill, 2005). Graphs designed for novices and children must take into account cognitive load and their available working memory resources by reducing the difficulty in keeping track of graphic references (Shah & Hoeffner, 2002; Friel, Curcio, & Bright, 2001; Plass et al., 2009). Furthermore, Libarkin and Brick (2002) point out that "cognitive skills must be acquired to assist in interpreting visual representations of actual phenomena. These skills are not necessarily a natural consequence of exposure to visual communication, and scaffolding between verbal and visual modalities may be an integral component of effective communication." Essentially, we cannot assume that simply providing the graphic to learners will ensure that they can and will interpret it, much less interpret it accurately and critically.

Neither can we assume that learners will bring a common set of schemas or interpretive tools to the display. As Shah & Hoeffner point out, “Peoples’ knowledge of the content in graphs has an influence on their interpretations of, and memory for, data … this is especially true for novice graph viewers who often do not have the graph schemas necessary to overcome the strong influence of their own content knowledge.” Friel (2001) agrees, noting, “The graph reader’s situational knowledge may interrupt her work on the cognitive, information-processing tasks performed in interpreting the graph.” Even in a classroom where students share a curriculum, we cannot assume shared schemas and prior knowledge when approaching visualizations. This is especially true in a museum where a wide variety of visitors with varied education and experience mingle.

The use of schemas for interpretation of information is by no means limited to graph interpretation. Wertsch (1998) refers to the cultural tools employed by learners in producing narratives. He describes prevalent use of what he calls the “quest for freedom” theme employed by students in a variety of ways as they attempt to compose an essay about the origins of the United States. This theme is one organizational tool, or schema, that students use to make sense of the information they recall about U.S. History. Similarly, learners interpreting data maps and graphs draw upon pre-existing themes when decoding the visualization (Roberts et al., 2012). Some of these themes are generally shared across a culture, such as Wertsch’s “quest for freedom” theme, which is ubiquitously part of the history curriculum in the United States. However, even these common themes are by no means universal, and they are picked up and utilized by people quite differently (Wertsch, 1998). When examining the learning occurring during the act of data interpretation, focusing on the intersection of the data and these pre-existing schemas or themes may provide insight to the conceptual changes occurring — or being resisted — as visitors encounter new information.

2.3.2 Maps as Reasoning Tools

Geographic Information Systems (GIS) maps are a particular category of data visualizations that overlay spatially-referenced data onto a geographic area in order to afford analysis of the data's spatial relationships. GIS maps visualize relationships and contrasts in striking ways in order to allow users to reason about the presented data. When used properly, GIS can assist with the recall of information (Rittschof & Kulhavy, 1998) and facilitate spatial reasoning and higher levels of thinking (Lloyd, 2001; Marsh, Golledge, & Battersby, 2007). Though data maps are one of the oldest and most popular forms of data visualizations and proliferate online, in textbooks, in the news, and in museums, people still have a hard time interpreting them correctly, and with a critical eye (Monmonier, 1991). Maps can be particularly powerful tools for decoding data because "maps affect how we think about spatial information; maps may lead people to think about space in more abstract and relational ways than they would otherwise. For these reasons, maps can be construed as tools for thought in the domain of spatial cognition"(Uttal, 2000). However, novice map readers need to be taught how to effectively utilize these tools, and especially to understand how maps "lie" and can easily distort data in order to convey a particular message (Monmonier, 1991; Krygier & Wood, 2011).

Much literature related to the interpretation of maps focuses on wayfinding (e.g. Beheshti, Van Devender, & Horn, 2012), but some researchers are concerning themselves with the map as a spatial cognition tool. Uttal (2000) reviews the utilities of maps for helping us conceive of the world beyond our immediate experiences, for making different kinds of information perceptually available, and for highlighting abstract spaces. The research presented here explores how data maps can serve as a tool for collaborative data exploration in a social museum context.

2.3.3 GIS Maps for Collaborative Activities

Traditional geospatial applications are designed for single users at individual workstations (MacEachren & Brewer, 2004; Forlines & Shen, 2005; Forlines, Esenther, Shen, Wigdor, & Ryall, 2006), despite the reality that most work with geospatial data (such as scientific analysis and urban planning) is carried out by groups (MacEachren & Brewer, 2004). Geovisualizations, if designed in ways that afford collaborative investigations, have great potential to support dialogue (MacEachren, 2005). MacEachren & Brewer (2004) coin the term geocollaboration as "visually enabled collaboration *with* geospatial information *through* geospatial technologies." This dissertation research will explore a geocollaboration activity in a museum; here I discuss efforts toward collaborative work around a GIS in classrooms.

Some work has addressed the incorporation of geospatial data into classroom inquiry activities. Groups of students in both middle school and undergraduate contexts have used a web-based GIS for inquiry into migrations and neighborhood change (Radinsky, Melendez, & Roberts, 2012; Melendez, Roberts, Radinsky, 2010). While students were able to use the GIS to support their narratives with some success, these interventions consisted of multi-week units in which students' inquiry skills and content knowledge were built through prolonged engagement. Discussion and reflection were guided by teachers who were able to appropriately scaffold the students in supporting claims with evidence and considering multiple sources of information.

Still, the undergraduate students who were given access to the full corpus of data and a wide range of manipulations of the maps were found in some cases to misrepresent or misinterpret the data in their presentations, for example conflating census variables (Radinsky et al., 2012). Middle school students were given a pared-down version of the interface presenting only a small set of census variables with limited data manipulations available in order to scaffold their explorations. In short, novice users (i.e. students) required a great degree of support in order to engage with the data.

Similarly, Edelson and colleagues (1999) discuss efforts to implement technology supported inquiry learning in classrooms through a multi-week inquiry unit around global warming utilizing custom-built scientific visualization software. In their work designing this interface to engage students with weather data, Edelson and colleagues discovered that merely giving students data maps did not immediately engage them with the content. Instead, having students create their own color-coded map representations of their predictions about weather patterns helped the students later interpret similarly visualized representations of actual data. The activity served as a bridge between something familiar and appealing (coloring) to the task at hand (interpreting weather visualizations). Such activities "can address motivational, accessibility, and background knowledge issues"(Edelson et al., 1999). Through multiple iterations of the software design, they discovered the importance of fostering student motivation, carefully sequencing activities to bridge students' prior knowledge, and developing supports for documenting new information in order to sustain prolonged inquiry.

These studies suggest the power of GIS as a tool for supporting learning but highlight the challenges of engaging learners even in a formal classroom environment with a teacher and extended curriculum present to facilitate and guide the learning. In an informal, unstructured museum exhibit experience, these challenges are amplified. Visitors support each others' learning through dialogue, so the exhibit must provide not only the content but also supports for engaging visitors with the content and with each other.

2.3.4 Summary

While data maps and graphs are unique ways of presenting complex information to users in a more easily comprehensible ways, they require active interpretation by users in order to make sense of the data. Many factors can influence comprehension of these visualizations, and designers need to take into account users' cognitive abilities, the nature of the data being presented, and the schemas inherent in different representations.

2.4 Perspective Taking

As outlined in Section 1.3.2, this research views *perspective taking* — the act of positioning oneself in a particular way with respect to the data — as a *cultural tool* that visitors can draw upon to help make sense of complex data for themselves and their co-visitors. This section summarizes how perspective taking has been examined by scholars in a variety of fields in order to extrapolate how it might relate to reasoning in a museum environment.

2.4.1 Perspective Taking in Individual Cognition

Some research has examined the role perspective taking plays in individual cognitive tasks. The field of cognitive psychology, for example, has taken interest in perspective taking in a variety of studies, for instance a study conducted by Brunyé and colleagues (2009) in which they examined readers' perspectives as they read an excerpt of fiction. In their exploration of narrative connections, Brunyé and colleagues provide a framework for examining the perspectives taken by readers in their mental simulations of a narrative. These simulations, measured by response times and accuracy of participants describing whether a sentence containing a first-, second-, or third-person pronoun matched a pictured event, examined the effects of pronouns on readers' embodiment; that is, whether the readers take an *actor* or *onlooker* perspective. Those readers taking an *actor* perspective position themselves in the scene as if they are carrying out the action themselves, which was associated with encountering first-person pronouns such as "I" and "we." However, as Brunyé et al. conclude, "embodying an actor's perspective is the exception, rather than the rule, in discourse comprehension." More often – and more easily – subjects take on the *onlooker* perspective, marked by third-person pronouns such as "he" and "they." While the Brunyé et al. study focuses on readers' positioning while engaging with pre-written narratives, this concept of the *actor*/*onlooker* framework as evidenced by pronoun usage is pervasive in the perspective taking literature.

In the context of procedural tasks and skill acquisition, Lindgren (2009) found that users were better able to reassemble a toaster when they viewed a tutorial of an expert filmed from the expert's perspective using a head camera than from the same tutorial filmed over the expert's shoulder, distancing the viewer from the task. In another study, participants who described actions in a segmented demonstration video from the actor's perspective later performed the assembly task better than those who had described the video segments from their own perspective (Lozano, Hard, & Tversky, 2006). A similar phenomenon was noted in a study of visitors engaging with a safety simulation. Those whose simulation positioned them as first-person actors required less help during the simulation and performed better on post-assessments than those who controlled avatars with a third-person view (Lindgren, 2012). In these cases, the technologies — video and simulation — were manipulated to encourage, or even force, users into taking a particular perspective, and in these cases the *actor* perspective, which provided users access to the information from a competent, expert point of view, was associated with improved learning gains.

2.4.2 <u>Perspective Taking in Social and Collaborative Learning</u>

While many studies have demonstrated how perspective taking can play a role in individual cognition, of interest to this research is how it can mediate group learning through dialogue. Research in a variety of contexts has shown that not only do learners mentally adopt a perspective, they use this perspective as a tool to explain a difficult concept or to collaborate with other learners. Because of the social, dialogic nature of museum learning, this use of perspective taking is of particular interest to this study.

Filipi & Wales (2004) describe how many studies examining perspective taking in spatial reasoning environments rely heavily on individual statements produced in lab settings around memory and representation tests, rather than the spontaneous use of perspective in dialogue. Because of the lack of work analyzing spontaneously occurring speech, they note that few claims can be made about the use of perspective taking in interactive settings. To address this concern, they set up a map task in which pairs of participants jointly constructed talk around route directions. Participants engaged in a structured task in which one participant attempted to describe a known route (printed on a map) to his co-participant who was viewing a blank map. They were particularly interested in participants' usage of two perspectives identified in prior geospatial reasoning work: the route perspective in which the speaker gives a "walking tour" of the route (essentially becoming an *actor* in the environment), and a gaze perspective where the speaker takes "a fixed, external vantage point" (Filipi & Wales, 2004), which can be considered analogous to the *onlooker* perspective. They found that in this particular task, the route perspective was utilized overwhelmingly during the activities. Overall, all dyads participating in the study naturally found the first person *actor* perspective to be most effective for conveying information to their partners. Participants temporarily shifted to gaze perspective (or perspective-free talk) to "repair" miscommunications, however, and these perspective shifts were therefore beneficial to the successful completion of the collaborative exercise.

Outside of geospatial reasoning, Ochs, Gonzales, and Jacoby (1996) investigated the role of perspective (and shifting perspectives) when trying to understand how expert physicists reason about their data during group meetings. In observing a group of expert physicists, Ochs et al. noted that the lead physicist on the project did not limit himself to a single perspective when discussing a phenomenon of interest. Instead, this expert used language consisting of both first- and third-person pronouns when discussing the data. Rather than referring only to a particle as a separate, external entity by utilizing third-person pronouns, the physicist constructed statements such as "When I come down, I'm in the domain state" to describe the particle's actions, in what is called an *indeterminate construction*. Prior work on scientific discourse had examined physicist-centered talk, in which the scientists talk about their own role in the process, e.g. "That's not what we observe really experimentally," in which "we" refers to the scientists, and physics-centered talk about the objects of the experiments, (e.g. "So many the domains need to grow."). Here Ochs and colleagues argue that a third perspective exists in which the physicist blends his identity with that of the particle when trying to understand that particle's actions. Ochs et al. explain:

> Indeterminate constructions are thus a resource which enables physicists to routinely manifest an extreme form of subjectivity by stepping into the universe of physical processes to take the perspective of physical constructs (i.e., to symbolically live their

> experiences). Like actors playing characters or reporters quoting others, however, while both voices are heard, the voice of the physicist is backgrounded, and that of the physical construct is foregrounded."

Essentially, by taking on the perspective of the particle using a linguistic construction that blends the particle with the physicist, this expert scientist is able to more easily and thoroughly comprehend the complex phenomenon observed in the data. When the unique construction was pointed out to the lead physicist in the lab, he acknowledged that this kind of talk was very common and used unproblematically by many physicists in their interpretation of their work. Prevalence of this linguistic construction by experts supports the notion that perspective taking can serve as a cultural tool that helps people in engaging in complex reasoning. Ochs and colleagues in particular noted that these productive, blended identities were inextricably tied to physicists' use of graphical data representations and physical movements during their discussion, making the findings particularly relevant to the interactive data map display at the center of this dissertation.

Enyedy, Danish, & deLiema (2013) report students' use of perspective in a classroom activity — Learning Physics through Play, or LPP — designed to model the effects of forces of friction. This activity utilizes whole-body interactions (Section 2.2.4) to allow students to act as a ball traveling along different surfaces. By encouraging students to think about friction from the perspective of the ball (the *actor* in the environment), the simulation facilitated students' ability to talk about the phenomenon and make predictions about the physical forces at play. The researchers note, however, that not all of these predictions based on blending the students' own physical experiences with the simulated object were accurate, such as when a student used her own experience of slipping on linoleum as the basis for predicting that a ball reaching a linoleum tile would also "slip" and therefore speed up (Enyedy, Danish, & DeLiema, 2013). The *actor* perspective in this environment afforded new ways of reasoning about the phenomena but carried the potential for encouraging misconceptions.

Similar misconceptions, or at least misconstruing of data, resulting from the use of the *actor* perspective have been identified in pilot work for this study (Roberts et al., 2013). Two adult museum visitors interacting with an early iteration of *CoCensus* eagerly adopted the *actor* perspective when viewing ancestry data in Chicago. The visitors, Belle and Peg (pseudonyms), observed differences in the distributions of the British (controlled by Belle) and Polish (Peg) populations and used the patterns as springboards for playful storytelling about the populations (Lyons & Roberts, 2014). Upon noticing that the British were more densely populated closer to Lake Michigan while the Polish population is distributed further inland, they bantered:

> BELLE: It looks like I'm along the lake.
> PEG: And I'm not. ((laughter)). Polish are inland. We're farming folk!
> BELLE: We're sailors!
> PEG: You're right, you're sailors!... So right, we have no idea... that's... if you're in the, if you're in the upscale? ((laughs))
> BELLE: I would think so…
> PEG: I think so too. You're by the water. It's more expensive.
> …
> PEG: So you're partying and I'm there in the fields.
> BELLE: Right, I'm having (all the fun).

The inferences made by Belle and Peg are not factually accurate or supportable by sociological evidence. However, we view this type of dialogue as productive for reasoning. They used the available data to notice patterns, draw upon prior knowledge (e.g. that farmland would be inland and that lakefront property would be more expensive), and make inferences (e.g. that living in the more expensive property probably means they are “upscale” and party more than field workers) (Lyons & Roberts, 2014). While these two visitors did not leave the exhibit interaction with an in-depth understanding of the reasons behind the spatial patterns, they did engage deeply with the data and likely formed meaningful connections that would make it more likely that they could reactivate this knowledge later (Falk & Dierking, 2000). Most importantly, these visitors engaged in this productive talk while using the *actor* perspective, with many inferences they made about the population referring to what “we” or “you” do. This connection between the use of *actor* perspective and engaged reasoning became the foundation for further investigations into the relationship, forming the pilot work for this research (see Section 1.2).

2.4.3 Summary

Perspective taking has been studied in a variety of contexts at both the individual cognitive level and social collaborative level. While the contexts and analyses varied, each study analyzed differences in usage of the first person *actor* perspective and the third person, external *onlooker* perspective. Findings suggest that perspective taking may serve as a cultural tool (Wertsch, 1998) to help people make sense about new information and share this knowledge with other learners through dialogue. Importantly, with the exception of pilot studies for this work, most studies that have looked at social use of perspective taking in meaning making situated one learner in a privileged expert or *actor perspective*, for example when one student acts as the ball in a physics simulation (Enyedy et al., 2013). Little work has been done to examine spontaneous use of perspective taking among learners in a shared social activity where both learners are assuming the same role. This work seeks to address that gap to further understand the affordances of perspective taking in a shared data exploration activity at a museum.

3 METHODS

3.1 Research Design

This research examines the role whole body interaction might play in mediating museum visitors' learning talk around a complex data display by looking at group dialogue during visitors' naturalistic interactions with one of four versions of an interactive data map exhibit. It then looks more deeply at one affordance thought to be associated with whole body interaction, *actor perspective taking* (APT). Two aspects of control that have high potential for affecting the degree of embodiment—single-input versus multi-input *distribution of control* (DoC), and full-body versus handheld *means of control* (MoC)—are studied here in a 2x2 quasi-experimental design (see Figure 6). Analysis of differences in visitors' learning talk and interactions among the four conditions tested here advance our understanding of the role of whole body interaction in mediating a group's learning dialogue around an interactive museum exhibit. This chapter details the functionality of the four controller designs tested (Sections 3.1.1 and 3.1.2) and describes the participants (3.2), instrumentation (3.3) and procedure (3.4) for this study. Descriptions of the codings applied to the dialogue and the categories of statements comprising "learning talk" are given in the following chapter.

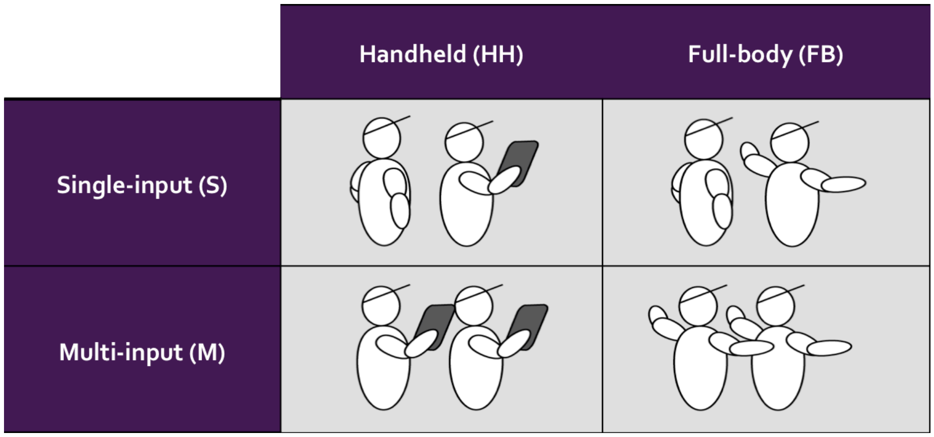


Figure 6. 2x2 Design of quasi-experimental study.

3.1.1 <u>Testing the Means of Control</u>

The exhibit analyzed in this study allows visitors to control three variables of the visualization: decade of data, category of data, and aggregation level of data. Two form factors for enacting these control actions were tested. The handheld controller is a tablet-based interface that utilizes familiar touch controllers — buttons and sliders — to manipulate the visualization (see Figure 7). The full-body controller condition places the locus of interaction and input in the user's body instead of an external device.

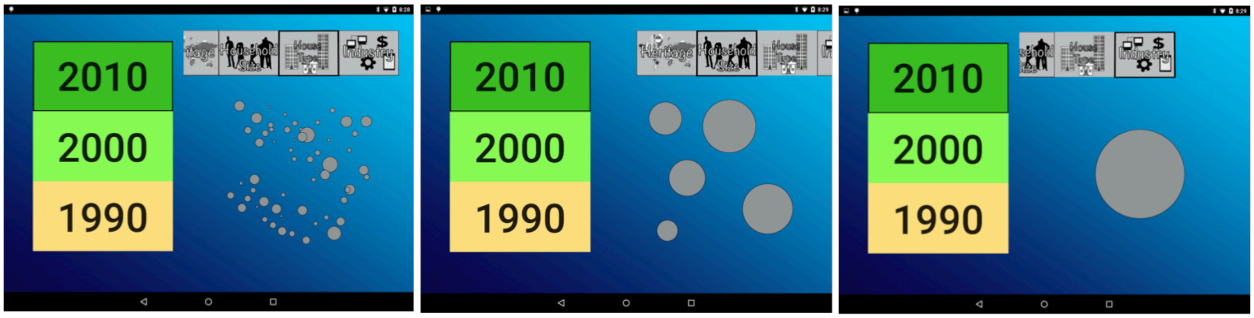


Figure 7. Screen shots of handheld tablet interface.

In order to select the census year of their data (1990, 2000, or 2010), visitors in the full-body condition step into marked areas on the floor, and visitors using the handheld controller push a button. In both *means of control* conditions visitors are able to view the decades sequentially to view change over time, toggle between two decades to highlight a phenomenon of interest, or remain in one decade to examine distribution of data sets at a particular point in time.

The other two controls to change the data category (e.g. ancestry, household size) and the aggregation level (tract, borough, or city) utilize arm gestures in the full-body condition and comparable hand movements in the handheld condition. The category change is accomplished with a sideways swipe that in the handheld condition mimics the flick smart phone users utilize to browse pictures in a camera photo gallery, and in the full-body condition is the same gesture but bigger, as if the user is flipping through the pages of a very large book. Changing the aggregation level is accomplished by a pinch motion on the handheld tablet and by pulling both hands apart or pushing them together in the full body condition, in a movement similar to that of playing the accordion.

The instructions for enacting these controls were given to participants via a 30-second slideshow at the beginning of their interactive session. As this research was not intended to be a usability study of the controllers themselves, if visitors had trouble controlling or missed (or misunderstood) the instructions, researchers intervened to explain and sometimes demonstrate the control actions.

3.1.2 Testing the Distribution of Control

The second variable tested is the number of users simultaneously operating the system. In all cases two participants completed the kiosk survey in order to view their own self-selected data. Participants in the single-input conditions were given a single global controller that manipulated both data sets simultaneously, i.e. in the handheld condition, a single tablet was provided that controlled both data sets, and in the full-body condition, one participants' movements and gestures controlled the data manipulations. Multi-input conditions allowed two participating visitors to independently control their own datasets. In the handheld condition each visitor was given a tablet, and in the full-body condition both participants wore an RFID tag coded with their own responses and could move independently to select the decade and aggregation level of the data. Because of the differences in affordances—e.g. that multi-input users had the ability to look at different decades or different aggregation levels than their companions, while single-input users' data were tied together—some differences in the interactions were anticipated and are accounted for in the analysis (section 4.7). However, it was not known whether the independence of the multi-input condition would cause confusion and hinder the interaction or would increase the connection each individual felt, therefore enhancing the experience.

3.1.3 Assigning Participants to Conditions

Due to the significant alterations to the exhibit required for each condition (i.e. modifying the floor covering and resetting the mode on the exhibit computer and tablets), random assignment to conditions was not practical. A pilot study testing an earlier version of the exhibit indicated that within-subjects testing was not productive, because once visitors had explored their data in one version, they had little or no productive dialogue in the second interaction. Therefore, conditions were tested individually in a between-subjects design on comparable days. Data collection occurred primarily on weekends and school holidays when family groups most often visit the museum.

3.2 Participants

This study was conducted *in situ* at the New York Hall of Science (NYSCI) between May and October, 2015, using members of the general visitor population. All visitors who were interested in using the exhibit were invited to do so, and all visitor groups who orally consented to participate in the research and wear lapel microphones during their interactions were assigned a group ID number. A total of 258 groups consented to participate in the research, but only a subset of these sessions were analyzed for this study. Visitor groups were excluded if they did not speak English during the interaction (N=17), if technical issues with the system interfered significantly with their experience (N=62), or if neither visitor understood the exhibit, either because they were too young to have developed map sense or they simply did not have the map fluency to understand, as evidenced by a total lack of substantive talk beyond reading aloud labels on the display (N=40). The latter exclusion was justified because the analysis framework presupposes that at least one of the visitors understands that the map they are viewing is a data map, and the data being shown are from the United States census. Future work investigating how such an exhibit could scaffold non-map-readers into map-reading practice would be of interest but is outside the scope of this study.

Consent was gathered in two ways. Lollipop stands posted outside the exhibit as well an explanation on the kiosk used for profile selection notified visitors that they were being video recorded in the space (Gutwill, 2003). Prior to the beginning of the interaction, a researcher asked visitors if they were willing to wear lapel microphones during their session. Visitors who chose not to wear the lapel microphones were not assigned an ID number and were excluded from the analysis. This research is IRB exempt (UIC IRB protocol 2012-0673). No personally identifiable data was collected about participants; they were not asked for their name, address, or any other sensitive information. Video recordings were taken from the back of the room so as not to deliberately capture their faces.

In total, 119 visitor groups were analyzed in this study. Of those, 41 were adult pairs and three more were groups of adults (in which two adults completed the kiosk and were viewing their self-selected data, and other adults were in the interaction area participating in the dialogue). Nine groups were child-child pairs and another nine were teen-teen pairs. There was one teen-child pair and 11 adult-teen pairs. Thirty-four groups were adult-child pairs, and 12 were mixed age groups of more than two people (e.g. a family with two parents and two children, together viewing two data profiles they had selected together).

3.3 Instrumentation

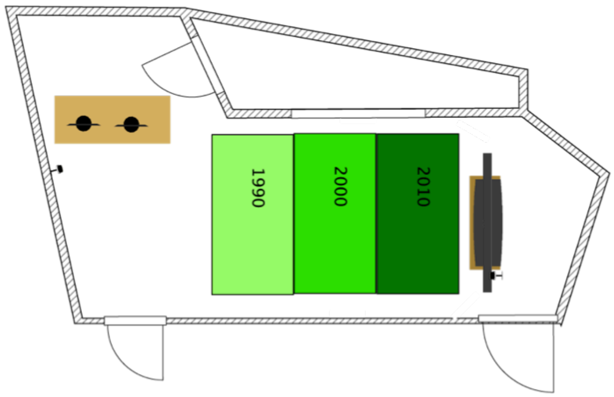
The kiosk running the mini-census survey is made from an HP Slate 21 Pro All-in-One touchscreen device running a custom application built in Processing 2.0 (see Figure 2). Once a visitor completes the kiosk survey, her profile is sent via the scripting language PHP over a private wireless network to a database housed on the computer running the exhibit and entered into two MySQL databases hosted on the exhibit's main computer. The "logDB" database creates a new record for each user completing the kiosk in order to keep a log of each participants' selections. A separate "kioskDB" database contains one record for each available input device (either RFID tags or tablet controllers, depending on the condition) and updates these records each time new information is sent from the kiosk in order to tell the system who is currently in the space, which then allows the display—a 90" LCD television located in a small room just off the main museum floor—to react to the individuals present in the space. This display consists of a map of the local geography, New York City, upon which are drawn scaled centroid "bubbles" that indicate the number of people in each geographic region who answered the census question the same way as that visitor over three decades (1990, 2000, and 2010). Visitors are able to control the visualization in three ways: selecting the decade of data to view at any given time, changing the category of data being viewed (i.e. ancestry, household size, housing type, or industry), and specifying the aggregation level (census tract, borough, or city).

The handheld conditions are run using one or two (depending on condition) Android Nexus 9 tablets running a custom built application that communicates with the exhibit's main computer via Bluetooth. The full-body exhibit utilizes antenna located in the interaction space that read a radio frequency identification (RFID) card coded with an individual's kiosk responses (see Section 3.3) in order for the system to recognize who is in the room. To capture fine motor movements and more precise position of each user, the RFID system is augmented with a motion-sensing Kinect camera. The two sensing technologies together utilize a probabilistic Bayesian model to distinguish the location and identity of each individual in the interaction space (Cafaro et al., 2013). Limitations in the accuracy and stability of this system, however, impacted participants' ability to interact with the exhibit and forced the researchers to employ a "Wizard of Oz" technique for most of the full-body condition sessions, in which one or two members of the research team used the handheld tablet controllers to control the display according to visitors' movements and gestures. None of the participants mentioned realizing this manipulation during the experience. See section 3.1.1 for descriptions of the controller equipment and functionality.

The map itself was created in ArcMap 10 using United States Census TIGER shapefiles. The formatting, layout, and labels were completed in Adobe Illustrator. The census data were retrieved from factfinder.census.gov and the National Historical Geographic Information System databases (nhgis.org).

3.3.1 Interaction Area

The exhibit is situated in a small, irregularly shaped room separated from the main exhibit floor at the New York Hall of Science (NYSCI) by a partial wall. The main interaction space measures approximately 9.5 ft x 20 ft. A 90" LCD flatscreen television serves as the main display at the far end of the room, away from the main entry door. A second screen powered by a rear-throw projector is mounted in the wall to the left of the main display but was not used in this study. In the full-body condition a rectangular "interaction area" is marked on the floor by a green floor mat labeled with the decades of data available for exploration. In the handheld condition the interaction area is demarcated only by a line of tape on the carpet. See Figure 8.



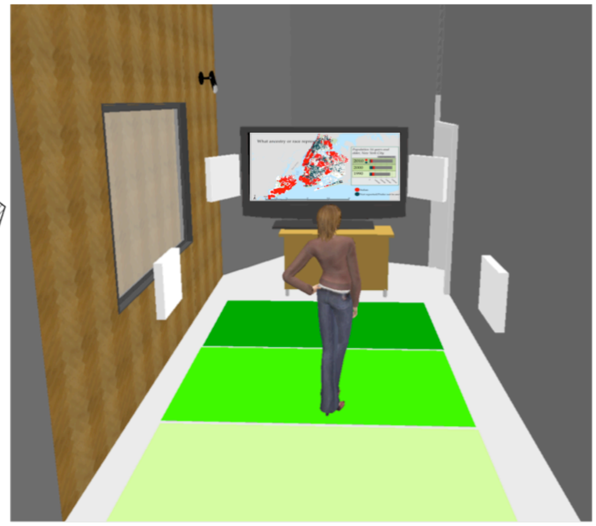


Figure 8. Configuration of experimental space at the museum. Floor plan (left) and 3D view of space (right)

In the back of the room a table holds two HP Slate 21 Pro devices serving as the kiosks for completion of the mini-census survey. During data collection, members of the research team or museum staff explainers remained by the kiosk displays in order to facilitate visitors' use of the kiosk, to assign consenting users ID numbers, and to record basic demographic information (See Appendix A: Visitor Tracking Sheet).

3.3.2 Data Sources

Two video feeds served as the primary data sources for this analysis. One camera mounted at the back of the room captured visitors' movements and large gestures in the interaction area. A second camera recorded a screen capture of the exhibit display. A third camera mounted at the front of the interaction area captured visitors' faces but was not used in this analysis. A shotgun microphone was mounted above the space in order to capture visitor dialogue. To augment the audio recording, clip lapel microphones were placed on visitors' shirts or on lanyards given to the participants. All video files and the audio file from one lapel microphone per session were synchronized using Chronoviz, and the two primary feeds (rear camera and screen capture) were exported with the microphone audio for analysis.

3.4 Procedure

Participants were recruited for this study in pairs off the museums floor. A member of the research team or museum facilitation staff ("Explainers") asked passing visitors if they would like to try a new exhibit to explore census data. Recruitment targeted visitors in groups of two or more who appeared to be at least ten years of age. This age threshold was used because pilot work showed that users younger than age ten generally lacked the map sense required to understand the complex representations. Younger participants were allowed to participate if they wanted, however, and they were included in the study if they and their guardians consented/assented to be recorded and their conversations indicated they understood the content. After participants were recruited, they completed the mock mini-census using the kiosk application (Figure 2) to choose a heritage, household size, house type, and industry that represented them, as well as a color selection out of three options per kiosk for their bubbles to appear on the main display. Members of the research team assisted and answered questions about the interface or exhibit content as necessary. After the survey was completed and the participants were satisfied with their answers, the responses were sent to the kioskDB and logDB databases in order to tell the exhibit who is currently interacting (kioskDB) and maintain a record of participants (logDB). Visitors assigned to the full-body control conditions were given lanyards with RFID cards and a clip microphone. Visitors assigned to the handheld control conditions were given an Android tablet running the controller application (Figure 7). In the single-input conditions one visitor was given a tablet or RFID tag that controlled both data sets, and in the multi-input conditions both visitors who completed the kiosk received a tablet or RFID tag that controlled their own data individually. If more than two visitors were present in the group, spectators (those who did not complete the kiosk survey) were invited to watch.

Because this research focuses on the collaboration of museum visitors in naturalistic interactions with the exhibit, museum staff explainers and members of the research team remained at the back of the room and only intervened in the interaction when requested by the participants, for example in the case of a technical glitch with the system. Participants were allowed to remain in the interaction area as long as they wished. When they were finished, they returned the lanyard(s) or tablets(s) to the research team as they left the exhibit area. For exhibit development purposes, some visitors were asked after their interaction brief open-ended or Lickert questions about the exhibit's functionality and responsiveness, but these responses were not included in this analysis.

4 MEDIATIONAL EFFECTS OF A PHYSICAL TOOL — ANALYSIS AND FINDINGS

This chapter examines the quasi-experimental 2x2 study exploring the effects of the *means of control* (MoC) and *distribution of control* (DoC) on visitors' learning and interactions. The chapter opens with a detailed account of the coding process for dialogue and visitor interactions with the exhibit, followed by a brief review of the four controller conditions. Findings are presented starting in section 4.6, in which differences among conditions on learning talk are described. Section 4.7 examines differences among conditions on interaction measures in order to contextualize and attempt to account for the observed differences in learning talk.

4.1 <u>Overview of the Coding Process</u>

Sessions were coded from video using the qualitative data analysis software MaxQDA 12. Complete transcripts were not created. Instead, videos were segmented into periods of substantive dialogue, non-substantive dialogue, silence, and researcher intervention, as described below. The decision to code this way rather than more traditional transcripts was twofold. First, earlier studies in this research project had used transcripts to analyze dialogue, and even with rigorous transcription conventions the loss of the context of the talk—i.e. what visitors were seeing and doing as they made a particular statement—removed or changed the meaning of the statement. It was necessary to consult the video while coding to ensure correct interpretation of dialogue, rendering the separate transcripts superfluous. The second reason is that one of the key affordances of traditional transcripts is the delimitation of talk into *turns*—when a new speaker begins talking, or someone resumes speaking after a significant pause—as a unit of analysis. As will be shown in the following sections, the turn is a meaningless delimitation in this context due to the spontaneous nature of the dialogue. Visitors regularly interrupt each other and themselves as they interact, and the fluid, nonlinear dialogue requires a more flexible system for segmentation of talk content. This study is more interested in documenting what kinds of content surfaced during talk, not in exploring the collaborative mechanisms by which visitors surfaced the content (although this data is certainly ripe for that type of analysis, it is out of scope here).

Therefore in this analysis substantive dialogue—that is, talk related to the behaviors of interest to this study, as discussed in detail below—was delimited by "idea units" (Jacobs, Yoshida, & Stigler, 1997, see section 4.2). Dialogue unrelated to those behaviors of interest was marked "non-substantive" and was disregarded. Examples of non-substantive talk include ambiguous utterances like "oh" and "hmm" and off-topic talk like parents telling their children not to climb stanchions. Periods of silence were marked, with a pause in between idea units or non-substantive talk lasting more than 1.5 seconds coded as silence. Pauses within idea units were not coded, and although idea units were coded during the playing of the instruction slides to capture any pre-interaction interpretation or scaffolding that was occurring, silence during the interaction slides (when no data was available to view or discuss) was not counted. This measure of silence allows for an understanding of how much of the active session time each visitor group spent in conversation. I opted to include pauses within an idea unit as part of the active session time, and not silence, because the listening participant's attention is likely to still be devoted to the speaking partner to some extent – insofar as an observer can tell, both partners treat these interrupted idea units as contiguous dialogue in the majority of cases. Periods of time where a member of the research team explained controls, reset the system to fix an error, or otherwise intervened were marked as "researcher intervention." Audio files from both participants' lapel microphones were consulted during coding, and any comments that were still inaudible even using all available audio files were coded as "inaudible" and excluded from analysis.

Coding was conducted using an "open coding" approach (Corbin & Strauss, 2008). Five categories of codes, described below, were gathered from museum studies and graph interpretation literature. Dialogue pertaining to those categories was coded with descriptive subcodes that emerged as unique instances of productive talk occurred. Sessions were coded in a random order, sampling from all four conditions, until all sessions had been coded and a final coding dictionary of 30 subcodes had been produced. All the sessions were then re-coded using the final 30-code dictionary. A subset of dialogue was coded by a second coder to compute inter-rater reliability, the results of which are discussed in Section 4.5.2 below.

4.2 Idea Units

Dialogue in this context is viewed as a group activity, and the analysis examines ideas spoken aloud by group members during their interaction with the exhibit. Some ideas are spoken by only one visitor and are contiguous and completed in a single conversational turn. Due to the spontaneous nature of the joint exploration, however, many are split among two or more visitors as they work together to make sense of the data. Visitors interrupt each other and in some cases interrupt themselves mid-idea as they notice new information. The fragmented nature of museum dialogue is a known challenge for assessing learning in this context (Allen, 2002). The spontaneous and non-linear nature of the dialogue makes coding conversational turns impractical, and larger delineations, such as segmenting by topic or referenced data, would obscure the intricacies of the productive dialogue. To reach the appropriate level of granularity, this analysis adopts a unit of analysis introduced by Jacobs et al. (1997), the *idea unit*.

Jacobs et al. (1997) defined *idea units* as being "marked by a distinct shift in focus or change in topic." While this definition is conceptually useful, the informal context of this study requires a slight addendum to the definition to reveal the phenomena of interest. Thus I have appended the definition by Jacobs et al. for this study to define an idea unit as marked by a distinct shift in focus or change in topic *or purpose.* This adjustment segments visitor conversation into chunks that delimit units of speech produced according to what that speech is doing in the group interaction. Idea units can range in length from a single word, for example reading aloud a category name (see section 4.3.2, Instantiations) to a multi-sentence utterance.

To illustrate the concept of idea units in this study, below are two excerpts of dialogue from two visitor sessions. The first shows somewhat straightforward linear idea units.

[1] A: I want to see how it changes.
[2] A: Like that area over there changed a lot in regards to... demographics, you see it?
[3] B: Yeah.
[4] A: And up there.
[5] B: More spread out.
[6] A: But you see the greatest change here on this side.

This excerpt was divided into four idea units. In [1] visitor A is stating his intention for exploration. Lines [2-4] he is pointing out areas of notable change, with visitor B agreeing. In [5] visitor B characterizes the difference. In [6] visitor A points out an area of particular interest. These idea units vary in length and in one case span multiple turns and speakers, but they are fairly straightforward, as annotated below:

[1]	A:	I want to see how it changes.	[states intention]
[2]	A:	Like that area over there changed a lot in regards to... demographics, you see it?	[draw joint attention to areas that changed over time]
[3]	B:	Yeah.	
[4]	A:	And up there.	
[5]	B:	More spread out.	[characterize data]
[6]	A:	But you see the greatest change here on this side.	[identify area of particular interest]

Some idea units are less obvious, because they are detached and inter-spliced. Take this segment from another pair:

> A: So whatever's, I'm assuming there must be railway or, oh wait, isn't that a road? That goes across, across the water. So there's-
>
> B: It's a bridge.
>
> A: So my guess is, oh it's a waterway or a roadway or whatever. Waterway maybe. But that area's most likely industry.

Visitor A's main goal is to pose his theory about the area being industrial but he keeps interrupting himself trying to correctly describe the roadway. This segment is counted as two overlapping idea units, as the participants are doing two meaning making moves in these three turns: decoding the map representation, represented with a dashed underline below, and posing an inference about the area based on the data ("So whatever's…So there's..So my guess is…But that area's most likely industry," double-underlined below).

> A: So whatever's, I'm assuming there must be railway or, oh wait, isn't that a road? That goes across, across the water. So there's-
>
> B: It's a bridge.
>
> A: So my guess is, oh it's a waterway or a roadway or whatever. Waterway maybe. But that area's most likely industry.

In this particular example the repeated starts to the inference ("So whatever's", "So my guess is…") would not have impacted the coding of the session. However, this same type of stopping and starting sometimes did involve coded elements, especially instantiations. The segmentation into idea units prevents stutters and echoing from unfairly weighting a statement beyond its contribution to the dialogue, which can occur in a speaking-turn-based quantification of talk (Chi, 1997). The idea unit coding used here is particularly useful when one is interested in characterizing the overall educational quality of a group's conversation, rather than trying to draw attention to the individual contributions or cognitive acts of each speaker in a social setting. Given the sociocultural perspective this work is taking (namely, learning is evidenced in the group's talk, and benefits the group as a whole), idea units seemed more appropriate than a turn-based approach.

4.3 Coding Substantive Talk

Talk was deemed substantive when it related to the product or the process of meaning making; that is, to the connection between ideas or creation of new ideas, or the productive talk that helped visitors get to those ideas. To identify instances of substantive talk, idea units were coded according to categories informed by two bodies of literature: informal learning (section 2.1) and graph interpretation (section 2.3). While these disciplines tend to have disparate research foci, they converge on key ideas that can illustrate the sense making done by groups in a data visualization exhibit. Five categories of substantive talk were identified based on these literatures, using pilot work for this project to inform the final selection of categories: instantiations, management, evaluations, integrations, and generations. Within these five categories multiple sub-codes were identified "up from" the data, using a descriptive coding approach (Miles & Huberman, 1994). This section describes the five key coding categories and then identifies sub-codes for each category. The full coding dictionary can be found in Appendix B: Learning Talk Coding Dictionary.

4.3.1 Manage Codes

It is to be expected that when multiple people are interacting with an exhibit, particularly when that exhibit is based on a novel technology, some amount of talk will directly address the interaction with the exhibit. This dialogue may be purely operational, for example explanations about how to perform a control action (e.g. "You have to swipe to change the category."), or it may concern how the group is managing its meaning making ("Let's look at household size."). This coding scheme thus differentiated between those two types: *operational* and *manage*.

Talk that was purely functional (i.e., pertaining to how one can operate the exhibit) was coded as *operational,* and was not considered for the purposes of content analysis. Operational statements were distinguished from management statements because they do not necessarily represent any attempt at coordinating the group's shared experience, and thus do not align with the behaviors flagged in informal learning literature as being beneficial for learning. Operational statements sometimes involved instantiations (e.g. "You can step here to change to 1990") but did not fall into any of the other management or reasoning categories of interest. While not part of Research Question 1's analysis of the learning talk that emerged under this study's different conditions (Section 4.6), I will revisit *operational* talk in Section 4.8.3, while exploring how visitors' interactions affect their learning talk for Research Question 1b.

Talk that related to the establishment of joint attention, negotiation of action, or scaffolding was coded as *management.* These kinds of behaviors are of interest to researchers of museum learning because they speak to how visitors are working together and mediating each others' experiences. For example, Allen (2002) categorized these kinds of actions as "strategic" with only two sub-codes: "use" and "metaperformance." Borun et al. (1996) attended to observable coordination behaviors like "call over." Multiple studies have attended to facilitative behaviors such as explaining, asking and answering questions, and suggesting actions (Ash, 2003; Eberbach & Crowley, 2005, Diamond et al., 1986; Atkins et al., 2009). Researchers of technology-based multi-user interactives are similarly concerned with interpersonal interactions like interference (Falcão & Price, 2009), negotiation of exploration (Davis et al., 2013), and collaboration (Williams, Kabisch, & Dourish, 2005).

The nine *management* sub codes defined here blend relevant components of coding schemes in each of these fields to capture the behaviors mediating visitors' interactions with this exhibit and each other. Three sub-codes—*ask interpretive question, purpose of exhibit*, and *clarify*—identified attempts among co-visitors to coordinate knowledge in order to establish an understanding of the exhibit's content. Four codes—*suggest action, narrate intentionality, direct co-visitor's movements,* and *negotiation of control*—mark different speech events related to coordinating movements and control actions. Some visitors, particularly parents trying to help their young children make sense of the display, would attempt to manage the shared experience by engaging in facilitation roles like a teacher might in a classroom setting by *asking guiding questions*. Visitors' attempts to explicitly establish joint attention to an aspect of the exhibit the speaker deems interesting were coded *direct co-visitor's attention*. Joint attention can provide a shared grounding to support further group meaning making, as the next section will describe in more detail.

4.3.2 Instantiate Codes

Here the term "instantiation" indicates when a user makes information part of the shared learning space via their conversation by saying it aloud. The instantiation of information (and new ideas related to that information) provides opportunities for the *individual* visitors to internalize that information (i.e., learn from the exhibit) and can be thought of as a step toward data collection to lay the foundation for further reasoning among learners on a museum visit (Kisiel, et al., 2012).

Verbal instantiation is also an important part of the *social* learning process: saying something aloud puts it into the shared social space, helping to establish joint attention (also referred to as “grounding”). Per sociocultural learning theory (see sections 1.3 and 2.1), learners must articulate ideas via communication before learning can take place. Processes of noticing and establishing joint attention among visitor group members have been found to be productive in facilitating learning talk in museums (Povis & Crowley, 2015; Leinhardt & Crowley, 1998), and reading labels aloud was identified as a “significant behavior” linked to increased group learning by Borun et al. (1996). The graph interpretation literature is typically concerned with how solo learners make sense of graphs, but there is an alignment between this kind of social grounding talk and a graph interpretation behavior identified by Friel, Curcio, and Bright (2001) as *translation*, in which a learner changes the form of communication of a map or graph. When a visitor engages in grounding talk by verbally highlighting aspects of the exhibit, the learner is essentially translating the written or visual information into verbal form.

Therefore analysis of visitor dialogue looked at instantiations of available exhibit content, that is, when they explicitly verbalized an aspect of the data displayed in the CoCensus exhibit. Six types of information in the exhibit could be instantiated, any number of times each: the *dataset* (e.g. “German,” or “educational services”); the *decade*; the census *category* (heritage, household size, house type, or industry); the *geography* (whether labeled, e.g. “Manhattan” or unlabeled, e.g. “the Park”); the visitor’s *self* (e.g. “I’m blue.”); or the *representation* (either the data bubbles or their color). Visitors could also instantiate *outside knowledge* by adding a piece of information relevant to but not directly represented in the display. These instantiations bring the information from the representation into the shared social space through dialogue. Drawing attention to content, particularly in this visually rich display, can focus visitors' attention jointly on a single aspect and make it available for internalization.

4.3.3 Evaluate Codes

The remaining three categories of codes attend to visitors' responses to data, and thus rely less on concepts drawn from the informal learning literature. Early attempts at coding this dataset drew on the three-tiered framework for behaviors involved in graph interpretation as codified by Friel, Curcio, & Bright (2001): "translation*,*" "interpretation*,*" and "extrapolation or interpolation," which roughly map to the three levels of "read the data," "read between the data," and "read beyond the data" succinctly described by Curcio (1987). In practice these designations, which were developed to apply to the kinds of graph interpretation tasks performed by students while engaged in schoolwork, proved to be too bound to the types of graph interpretation tasks found in classrooms to capture the nature of visitors' talk in a social, free choice learning experience. In particular, the idea of a tiered taxonomy of graph interpretation skills does not necessarily apply to a free-choice setting, where visitors might, for example, jump right into "read beyond the data" talk by incorporating their own personal knowledge into the discussion. The sorts of informal talk visitors engage in do not allow observers to reliably discern which "read beyond" comments are reflective of deep graph comprehension and which are not, because visitors are not required to perform specific, measurable tasks as they are engaging with the data. Thus, preserving the existing tiers would not be valid for this context, although the categories themselves were a useful starting place for characterizing visitors' talk about data. The lower-tier "read the data" concept of "translation" became subsumed by the *instantiation* code, described previously in Section 4.3.2, and the higher-tier concepts were replaced by several new codes developed to stay closer to the dialogue, delimited by the scope of data to which statements refer. Statements that interpret only one dataset were deemed *evaluation*, a code I describe in more detail here, while statements combining or connecting multiple pieces of data were considered *integration* or *generation,* as described in more detail in sections 4.3.4 and 4.3.5 below.

Evaluation statements go beyond merely instantiating data and make some kind of judgment or assessment about a piece of information by assigning some kind of value, whether qualitative or quantitative. In the graph interpretation literature, the "read the data" categories like "translation" are strongly associated with inspecting graphs and correctly reciting specific quantitative information – for example, that a particular bar in a graph reflects five units. This sort of behavior is less applicable to free-choice informal settings – visitors are much more likely to engage in qualitative evaluations, as when they note that a graph shows "a lot" of something. In fact, personal qualitative evaluations are arguably very important in informal learning settings, where developing one's identity is seen as just as much of a goal of the meaning making process as absorbing content (Rounds, 2006). Interestingly, the idea of making personal judgments or evaluations is largely absent from the graph interpretation literature, possibly because of the strong task focus present in classrooms (where learner actions are either correct or incorrect). Sometimes evaluation occurs in the literature's higher-tier categories like "interpretation," as when a learner is asked to judge the direction of a trend in data, or "extrapolation," as when a learner is asked to judge if a data set's variability is small or large. In this informal context, however, evaluations can be simple standalone comments or part of a more complex statement. The most common kind of evaluative statement was *characterize*. Examples of *characterize* evaluation statements are those remarking that there are "a lot" or "not very many" of something, or describing a population as being "everywhere." The characterizations could be spatial or quantitative in nature.

The evaluation category contains two additional sub-codes. *Win* statements mark when people make a competitive evaluation of the data, making statements about winning or dominating or destroying the other person. These statements generally refer to quantitative comparisons (and thus are dual-coded with Integrate-Comparison code, see Section 4.3.4) but are also coded as evaluate because of the value judgment being made about the comparison itself. These statements are of interest because they were often associated with *actor perspective taking*, as will be discussed in Chapter 5. The final sub-code in this category is *question census categories*. Occasionally visitors explicitly call into question how a dataset is calculated by the census and what the category means or doesn't mean. This kind of questioning—which is rooted in visitors' readiness to make judgments about data sources—is one of the behaviors we hope the exhibit will produce as visitors question how and whether the census represents them (Roberts et al., 2012).

Statements were only coded as evaluations if they were unambiguously tied to an aspect of the exhibit. Statements that could be considered evaluative but were vague, such as "Wow! That's cool," were treated as non-substantive in this analysis, because they did not make explicit what aspect of the data was being evaluated. Thus, given the sociocultural perspective that ideas need to be articulated for them to contribute to learners' shared meaning-making, I could not reliably infer how such statements were directly contributing to meaning-making (beyond indicating the learners' affect). Of course, affect is widely noted as being indirectly important for supporting informal meaning making, but given the current lack of reliable ways of documenting affect, I deemed an analysis of affect outside of the scope of this current work.

4.3.4 Integrate Codes

While evaluation statements refer to a single dataset, the next two categories connect multiple pieces of information in some way. Friel, Curcio, and Bright (2001) refer to the act of looking for relationships in data or identifications of trends as "interpretation." This dissertation adopts the more precise term *integration* from Murray, Kirsch, & Jenkins (1997) to describe the act of pulling together multiple pieces of information. Statements that integrate are those that make explicit connections or comparisons between multiple pieces of information presented in the display, such as between two different datasets, between a dataset and the geography, between a dataset and itself over time, etc. Two types of *integration* statements were found. Some were straightforward *connections*, linking two types of data such as "There are no single family houses in Manhattan," (connecting the dataset to the geography), and some were *comparisons* between two or more pieces of the same type of data, e.g. geographies, decades, or datasets (e.g. "There are so many more of me than there are of you").

Connections are another step toward data collection, similar to *instantiations*, making explicit the relationship between two or more components in conversation. Early coding in this category attempted to specify the exact components being connected, e.g. "connect dataset to geography" or "connect dataset to decade." However, the additional level of detail was found to add ambiguity to the analysis without returning better understanding of the phenomena of interest. Therefore the original detailed codes were replaced with *connect simple* and *connect multiple*. A simple connection would be e.g. "A lot of them are in like, Queens," (connecting the dataset or representation meant by "them" to the geography "Queens") while a multiple would be "This is showing all the Puerto Rican people in New York in 1990" (connecting the dataset to the geography to the decade). Note that while *instantiations* require the visitor to say the name of the dataset (as this seems a logical requirement for establishing a common grounding in conversation), *connections* can be coded even when the statement uses a pronoun (e.g. "them") with an implied antecedent.

Whereas *connections* relate *across* types of data (dataset to geography, etc.), *comparisons* relate two facets of the same component, for example visitor A's heritage dataset to visitor B's heritage dataset, or visitor A's industry dataset in 1990 to visitor A's industry dataset in 2000 and 2010. To use a familiar analogy, *comparisons* can be thought of as relating "apples to apples" while *connections* relate the fruit to the tree or to the grove.

4.3.5 Generate Codes

The final category of dialogue coding is another form of combining two pieces of information, but unlike *integrate* statements, *generate* statements "[go] beyond the data" (Curcio, 1987) to combine information from the exhibit with visitors' own prior knowledge and experiences. In generating tasks, "one must not only process information in the document but also make document-based inferences or draw on personal background knowledge" (Murray et al., 1997). Drawing on information beyond what is directly presented in the exhibit suggests that visitors are connecting to the new information in ways that are more likely to be meaningful beyond the immediate exhibit interaction. These kinds of connections are widely acknowledged to be important to museum learning. Falk & Dierking's (2000) Contextual Model of Learning posits that what learners gain during a learning experience is inextricably tied to what the personal context they brought into the experience—prior knowledge, experiences, motivations, identities, etc.—as well as where they go next, and that rather than asking the common question of what a person learns in a museum, the question should be "How does this museum, exhibition, or lecture contribute to what someone knows, believes, feels, or is capable of doing?" Allen (2002) incorporates what she calls "connecting talk" into her framework for analyzing visitor conversations at an exhibit, but unlike the *connections* described above as an integrate code, the type of connections she is referencing are making use of outside information, by connecting an exhibit to life, prior knowledge, or other exhibits. She describes this stitching-together of information from different sources as "powerful and ubiquitous means of learning in informal settings."

Multiple kinds of *generate* statements were identified in this analysis. *Contextualize* statements connected outside knowledge or experience to make sense of data, for example when a parent reminded her child of what all the buildings looked like in Manhattan to explain why there were no bubbles representing single-family detached homes there. Two *generate* sub-codes dealt with visitors' reactions to data in relation their prior knowledge: *confirm* marked when data matched the visitors' expectations, and *notice surprising pattern* indicated when the presented data conflicted with expectations. When two visitors negotiated an interpretation of the data using prior knowledge, it was coded *negotiate meaning*, and when they noted that they would require additional outside knowledge to complete their understanding, it was coded *identify knowledge gap*. Some visitors would *make predictions* about data patterns, either discoverable in the available data (e.g. "I bet if I go to 1990 there will be less.") or not discoverable ("I bet in the next census it will change."). Finally, visitors would *pose inferences* about the data, for example suggesting an explanation of why a particular pattern exists ("I think [house type] is more restricted by the space around here").

4.4 Coding Overlap

These five categories of codes each addressed a particular function of any given idea unit, and any utterances that did not match any of the above categories and sub-codes were marked *non-substantive* and were disregarded in the analysis. Some idea units were coded with a single code. Many idea units, however, were coded with multiple codes: though the statements were one logical idea, they were deep and complex enough to warrant multiple codes. For example:

"Okay, there's a lot of White in 2000, rather than Mexicans."

This statement as an overall idea unit describes the two heritage groups. Within that broad goal, it does multiple things. It *instantiates* the datasets ("White" and "Mexicans") and decade ("2000"), it *connects* the dataset to the decade "White in 2000", it *compares* the datasets ("White" and "Mexican") using "rather than," and it *characterizes* "White" as being "a lot." Another example:

"But if we go back to 1990, I'm sure there were less apartments."

This statement contains a *management* code, *suggesting the action* of going back to 1990, *instantiates* the decade and datasets ("apartments," though the category name was longer than that, abbreviations were counted), and *makes a prediction* about what they will find when they arrive at that decade. One final example illustrates the segmentation of idea units and the subsequent analysis reported in sections below:

"Oh yes, lots of West Indians in Brooklyn, that is true."

This statement *instantiates* the dataset and geography, *characterizes* the population of West Indians as being "a lot," *connects* the West Indian dataset to the geography Brooklyn, and *confirms* the data against her own prior knowledge. This statement was coded as a single idea unit (IU). It could be argued that it is two embedded idea units, with "Oh yes… that is true" being one IU confirming outside knowledge and "lots of West Indians in Brooklyn" being its own IU, separately *characterizing* and *connecting* the datasets. However, that delimitation would not change the scoring of the reasoning talk. In either case, this dialogue would receive two instantiation codes (dataset and geography) + one characterization + one connect + one confirmation. The analysis presented here does not attend to the way codings were distributed among idea units, only to the total number of codings applied to the dialogue throughout the course of the session. The assumption being made by this approach is that the overall richness of codes corresponds to the overall richness of the shared learning experience. So, for example, this approach would allow a relatively taciturn group that made fewer statements overall but whose dialogue contained a rich amount of content discussion to be favorably compared to a highly talkative group whose dialogue demonstrated less rich engagement with the content.

4.5 Tying Codes to Exhibit Learning Goals

The coding framework described above stays close to the data in identifying how people are talking by flagging conversational acts that are likely to contribute to shared meaning-making at an interactive data exhibit, according to informal learning and graph interpretation literature as described above. The graph interpretation literature had its own way of establishing the relative value of categories, but as the preceding discussion of code categories explained, this tiered valuation makes the most sense in a formal classroom, and does not map cleanly to a social, informal setting. To address Research Question 1, which examines the effect of MoC and DoC on visitors' learning talk, it was necessary to devise a tiering that better matches the exhibit's learning goals. That is, even though *generate* statements may be considered the top of the hierarchy of graph reasoning codes, the exhibit's purpose is not just to foster generate-type statements. Because the CoCensus exhibit is intended to foster group exploration of census data and facilitate multiple kinds of meaning making, multiple kinds of talk about data are considered to be highly relevant to the intended learning. Other kinds of data talk are important but less directly aligned with the learning goals. Therefore, this analysis undertakes a form of magnitude coding (Saldaña, 2009; Miles & Huberman, 1994) by sorting the above codes into high, medium, or low categories according to their relation to the goals of the exhibit, as determined by the research team (see Table 1). These categories are assigned weights in order to quantify and compare the substance of visitor talk across conditions. Using magnitude coding as a way of "quantizing" a phenomenon (Tashakkori & Teddie, 2010) permits the use of inferential statistics (Bernard, 2006; Saldaña, 2009), which is useful for comparing the experimental conditions.

TABLE 1. CONTENT CODES, SORTED INTO CATEGORIES RELEVANT TO EXHIBIT GOALS

Low Relevance (1)	Mid Relevance (2)	High Relevance (3)
INSTANTIATE category	EVALUATE characterize	EVALUATE question census categories
INSTANTIATE dataset	EVALUATE win	GENERATE confirm
INSTANTIATE decade	GENERATE contextualize	GENERATE make prediction
INSTANTIATE geography	GENERATE identify knowledge gap	GENERATE negotiate meaning
INSTANTIATE representation	INTEGRATE challenge interpretation	GENERATE notice surprising pattern
INSTANTIATE self	INSTANTIATE outside knowledge	GENERATE pose inference
INTEGRATE connect multiple	MANAGE ask guiding question	INTEGRATE compare
INTEGRATE connect simple	MANAGE clarify	
MANAGE ask interpretive question	MANAGE direct co-visitor's attention	
MANAGE direct co-visitor's movements		
MANAGE narrate intentionality		
MANAGE negotiation of control		
MANAGE purpose of exhibit		
MANAGE suggest action		

The codes assigned to the "low relevance" category are most of the instantiation codes, the *integrate*: *connections* codes, and many of the *management* codes (see Table 1). These are all activities that are useful for grounding and coordinating the group learning experience and may serve as springboards for future interpretive statements, but are in and of themselves not strongly related to the learning goals of the exhibit. These statements were assigned a weight value of one. "Mid relevance" codes took steps to more directly make sense of the presented data by characterizing and contextualizing it (including instantiating outside knowledge to help with sense-making), clarifying the representational forms, and directing co-visitor's attention to an interesting element of the exhibit (which rises above a simple *instantiate* code because it conveys to the listener that the targeted element is worthy of joint discussion). These codes were given a weight score of two. "High relevance" talk included statements that related presented data to prior knowledge or expectations, predicted or inferred information, compared datasets with each other or over time, and questioned the source of the data (such as how the census counts a particular category). This kind of talk is exactly the kind of exploration and meaning making the exhibit hopes to support, and these comments were assigned a weight of three. Looking back at the earlier example of a multi-coded idea unit:

"Oh yes, lots of West Indians in Brooklyn, that is true."

This statement *instantiates dataset* (1) + *instantiates geography* (1) + *characterizes* (2) + *connect:simple (dataset to geography)* (1) + *confirms* (3) = content score of 8. As noted above, whether this was delimited as one 8-point idea unit versus a 3-point idea unit ("Oh yes… that is true) plus a 5-point idea unit ("lots of West Indians in Brooklyn"), the impact on the session score is the same. The study does not analyze scores of individual idea units, only on the total dialogue in a session.

4.5.1 Validity and Limitations of Magnitude Coding

Any attempt to assign quantitative values to qualitative data will inevitably strip some nuance from the analysis, as a single numerical value cannot possibly capture the full depth of an interaction. On the other hand, rich qualitative analyses do not lend themselves to the kind of cross-condition comparisons sought by this study. In making the decision of how to quantify the dialogue in these visitor sessions, therefore, it is key to bear in mind the focus and purpose of this analysis. The session scores are intended to illuminate differences among the conditions in their ability to support visitors in productive exploratory talk about data. Even the codes identified as "low relevance" are still productive learning talk. In this context, high-value codes often (but not always) build on low and mid-value talk, and a good statement often contains all three.

For example, consider the statement, "Wow, there are so many more Puerto Ricans in Brooklyn than Queens." Saying aloud "Puerto Ricans" and "Queens" are both instantiations; "Puerto Ricans in Brooklyn" is a connection; "so many" is a characterization, and "more in Brooklyn than Queens" is a comparison. That's 3 low-level, 1 mid-level, and 1 high-level code applied to a single statement. In the presented weighting system, this is an 8-point statement (1+1+1+2+3). An alternative method for evaluating visitors' dialogue might be to focus on the proportion of codes at each relevance level (high, mid, and low) applied in each session. If you look at proportions, that example becomes 3/5 (or 60%) low-relevance, and 20% mid-relevance and 20% high-relevance. So by that system, a much simpler statement like "It looks like there are more of them" is a 100% high-level statement, because it is a comparison without any elaboration. But in terms of what it's adding to the conversation—keeping in mind the focus of this study on group dialogue—it is nowhere near the same level of substance. Perhaps most importantly, the simpler statement gives the speaker's companions fewer "hooks" from which to build on the statement: all they can respond to is the comparison, whereas the more complex statement gives companions a number of different directions to take the conversation (e.g., adding information pertaining to Queens, noting the representation of another group in Brooklyn, making a further observation about Puerto Ricans in other boroughs). Given the sociocultural perspective on learning that this work is taking, the more complex statement supports a richer array of potential group learning opportunities.

Productive dialogue about data in this exhibit thus doesn't necessarily need to be efficient (with all or most of the talk being high-level). In fact, we want people to be doing all kinds of data talk, and more low-relevance statements might be required for meaning making, for example in the case of a parent scaffolding the interpretation for a child. An appropriate coding system, therefore, needs to acknowledge the value of all productive talk without losing the qualitative differences between, for example, more substantive comparisons and the cognitively easier instantiations. The magnitude (or “weighted”) coding system presented here provides this affordance.

To illustrate the utility and limitations of this weighting system at the session level, I first will introduce a visualization to be used later in this dissertation: a session timeline. In Figures 9 through 11 below, time is represented on the x-axis. Each coded idea unit is placed chronologically, with color-coded bars representing the codes applied. Purple bars indicate codes categorized as low relevance to the exhibit’s learning goals. These codes mark productive talk for establishing joint attention and gathering data, including instantiating and connecting elements. They are visualized as descending from the x-axis, illustrating their function of laying the foundation for deeper reasoning. Each purple code has a value of one; longer bars indicate multiple low-relevance codes applied to a single idea unit. The orange bars indicate the mid-relevance codings, and the blue represent high-relevance codings. The blue and orange bars are both drawn with their weighted values, with a single mid-level orange statement given a value of 2 and a blue statement 3. Red short bars indicate Actor perspective taking (APT), to be discussed further in Chapter 5. This timeline representation allows for exploration of different ways productive visitor talk in a session can unfold, as the three example cases below will illustrate.

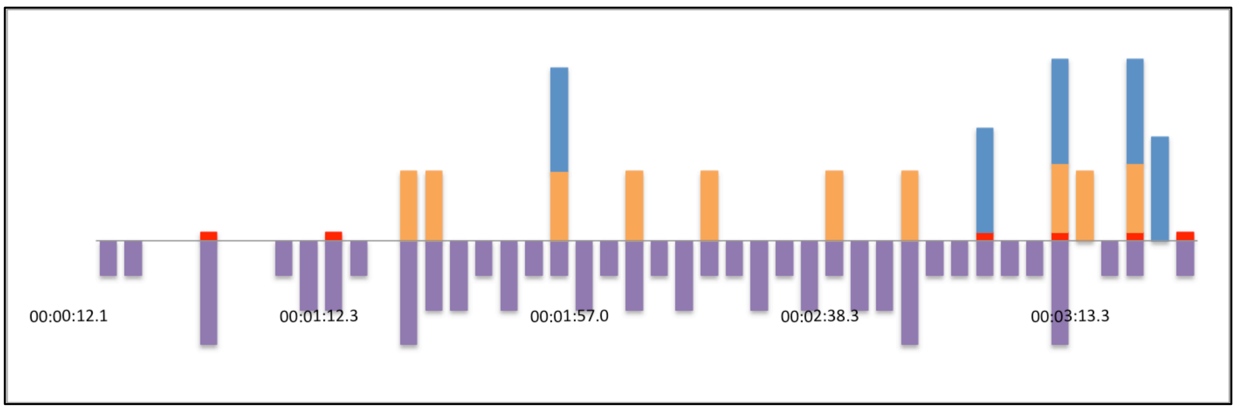


Figure 9. Session timeline for HHM-S60. Purple bars descending from the x-axis represent low-relevance (but still productive) codings. Orange bars represent mid-relevance, with a weighted value of two, and blue bars represent high relevance codings with a value of 3. Red markers indicate *Actor* perspective taking, discussed in Chapter 5.

In the first example, an adult pair talked throughout most of their three and a half minute session: we can see in Figure 9 that there are very few blank spaces along the x-axis. However, much of this talk was low-relevance talk: they spent a lot of time instantiating and connecting elements and managing the interaction. Approximately 80 seconds into the session they moved into mid-relevance talk, primarily characterizing datasets, and most of the high-relevance talk occurred in the last 30 seconds of the interaction. This session contained 72 codes: 5 high-relevance, 10 mid-relevance, and 57 low-relevance. The session's content score according to the weighting system described above (values of 1-2-3 for low-mid-high) is 92.

Contrast that case with another adult pair example, FBM-S60, shown in Figure 10. During their 4 and a half minute interaction, this pair had more periods of silence (or non-substantive talk – both would appear as blank spaces in this representation) than the previous pair, but they jumped in to mid-relevance dialogue much more quickly and more frequently. Like the earlier pair, this pair also did their high-relevance talk at the end of their interaction. Their 73 codes were broken down as 2 high-relevance, 16 mid-relevance, and 55 low-relevance, for a total content score of 94.

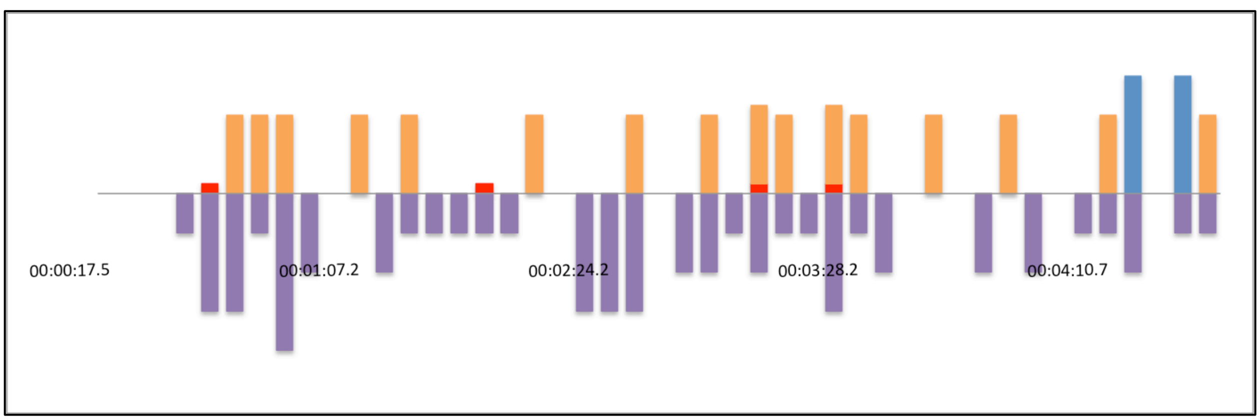


Figure 10. Session timeline for FBM-S60

The final example to be presented here is a third adult pair, this one in the full-body single input condition, FBS-S45 (Figure 11). This pair had the fewest codes applied to their dialogue of these three example cases, with only 48, even though their interaction was similar in length to the previous pair. However, most of their dialogue consisted of high-relevance talk comparing datasets, exactly the kind of talk the exhibit hopes to support. With 11 high-relevance codes, 16 mid-relevance, and 21 low-relevance, this group had the highest proportion of high-relevance talk of any of our examples, resulting in a session score of 86.

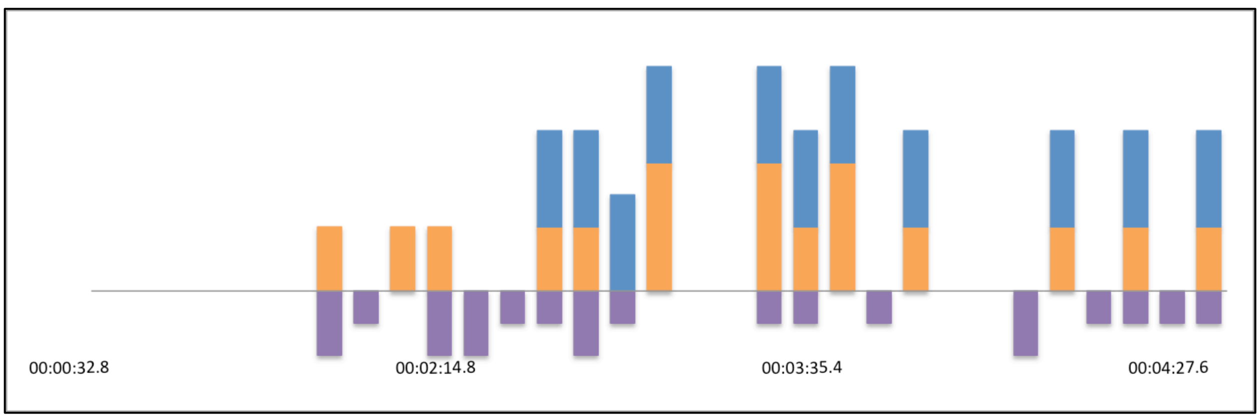


Figure 11. Session timeline of FBS-S45

Given the open-ended nature of the interactions and the underlying assumption that each group will be approaching the exhibit from a unique background and with unique goals, it is to be expected that productive interactions will not be the same for each group. The numerical differences in the presented examples are trivial in the context of the overall analysis, in which the average content score for the 119 sessions was 69.4, with a standard deviation of 42.6. Each of the three examples presented here demonstrates a different kind of interaction that fell roughly one half a standard deviation above the mean in productive talk. The many low-weighted remarks of a group like FBM-S60 provide ample opportunities for companions to build on each others' statements and thus engage in shared meaning-making, whereas the fewer high-weighted remarks of FBS-S45, while containing richer content, provide fewer opportunities for companions to contribute, resulting in a slightly lower score than FBM-S60. Both learning experiences *are* valuable, though, and the utility of this weighted coding system is that it allows different kinds of engagement like these to be acknowledged as productive while still distinguishing low and high performing groups.

Figure 12 illustrates the differences in the three examples in numbers of codings at each relevancy level in comparison to a low-performing group (FBS-S26, score of 27) and very high performing group (HHS-S38, score of 185). The low-performing group involved a largely one-sided dialogue, with the active participant in the full-body single-input interaction narrating her activities and making some interpretive statements but receiving very little input from her companion. By comparison, both members of the high-performing group were actively engaged in data interpretation and discussion, building off each others' comments and their own observations. The richness of their discussion is evidenced by the high numbers of codes applied in all three categories.

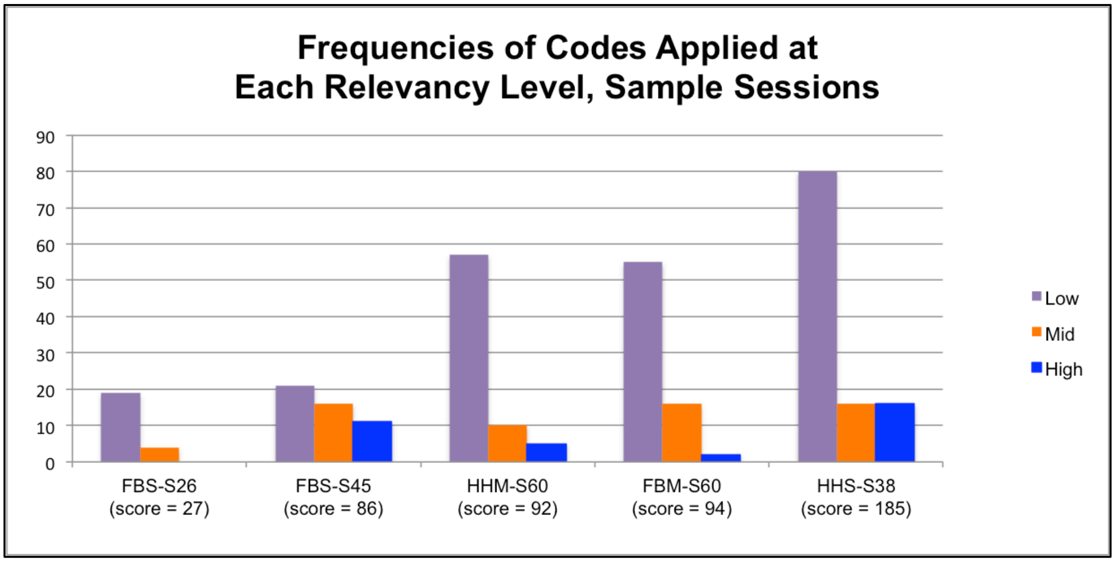


Figure 12. Despite marked differences in the number and types of codings applied, the three example sessions have similar content scores using the 1-2-3 weighting system. They are clearly distinguishable, however, from extremely low-performing (FBS-S26) and high-performing (HHS-S38) groups.

The final point to consider in applying a magnitude coding scheme is the numerical values appropriate to be assigned to each code level. This analysis assigns values of 1-2-3 for low-mid-high codes respectively, but could easily have chosen another relationship, for example 1-3-5 or 1-5-10, to represent the differences. Figure 13 demonstrates the score differentials of these example sessions presented above in the current system ("score") and these two alternatives. Alternative 1 (1-3-5) slightly boosts FBS-S45 above the other two slightly above average cases presented above, and—not surprisingly—alternative 2 (1-5-10) significantly elevates FBS-S45, having given high-relevance statements ten times the significance of low-relevance. However, the three are all still distinguishable from their low and high-performing counterparts.

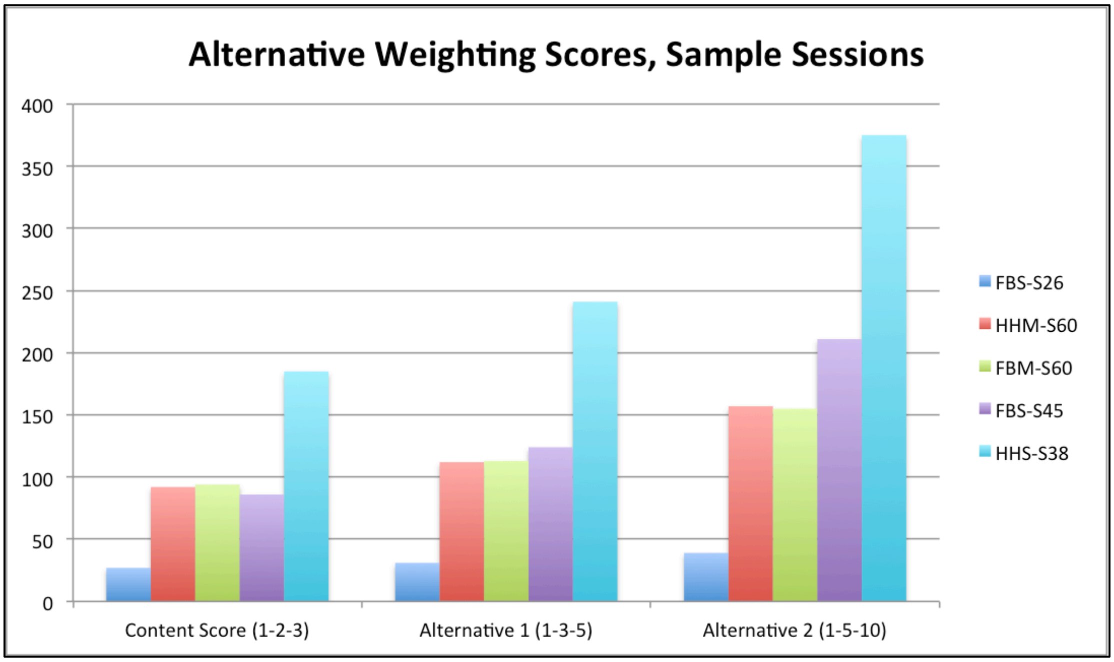


Figure 13. Relationship of example sessions with proposed scoring (1:2:3 ratio of low-mid-high relevance codings) and two alternatives. Alternative 1 uses a 1:3:5 ratio and Alternative 2 uses 1:5:10.

While changes to these values impact individual session scores, analysis presented later in this chapter will show that each of these weighing systems yields comparable results when comparing conditions overall. Therefore unless otherwise noted content scores presented in this analysis are computed with the 1-2-3 weight values for low-mid-high relevance codes, and these content scores will be used as the main metric for comparing visitors' learning talk across conditions. Analysis of the frequencies of codings applied at each relevancy level are discussed in Section 4.7.2, but the bulk of the analysis rests on content scores of sessions based on the weighted code values. Section 4.7.3 analyzes main and interaction effects of the *means of control* and *distribution of control* on content scores using this weight system, and in the interest of rigor also compares the effects with the alternative weight systems presented above (1:3:5 and 1:5:10), finding comparable results regardless of the weighting system.

4.5.2 Inter-coder reliability

In order to test the reliability of the coding scheme, a total of 151 idea units from seven sessions were coded by a second coder. In total, 467 codes were applied with an overall agreement of 85.65% (400 of 467 codes). Cohen's kappa was run to determine if there was agreement between both coders on application of the 29 substantive talk codes, and there was strong agreement, $\kappa = .825$, $p < .001$.

4.6 Addressing Research Question 1: Impacts of Means of Control and Distribution of Control on Learning Talk

The first research question explores how the distribution of control (DoC)—single-input or multi-input—and the means of control (MoC)—full-body or handheld—impact visitors' learning talk while interacting with an exhibit. In this section, I briefly review the experimental conditions (section 4.6.1) and document that the overall conversation that occurred at the exhibit during each condition was equivalent (4.6.2) to establish that subsequent analyses of content differences cannot be attributable to variance in overall talk.

Subsequent sections take on Research Question 1a (examining the content of visitor talk across conditions, section 4.7) and Research Question 1b (examining how visitor interactions relate to their learning talk, section 4.8).

4.6.1 Review of Study Conditions

This study examined the role of whole body interaction and individualized control of an interactive data map display. Two different *means of control* (MoC) were tested. A full-body (FB) interaction design used body movements on the floor space in front of the exhibit display and two distinct arm movements to enact the three manipulations of the visualization – selecting a decade of data, selecting a category of data, and changing the aggregation level at which the data were displayed. This design was contrasted with a handheld (HH) tablet design in which the visitor could perform the same manipulations using similar gestures on a 9-inch tablet (see Figure 7).

The *distribution of control* (DoC) was also tested in this study. In the FB condition, this meant users could separately step through decades and alter the aggregation level for their own datasets without disrupting their companion's data (see Figure 14), and in the HH version, both visitors in the pair were given a tablet to manipulate their data separately. In both versions the category changes were linked: if either visitor changed the category, the whole display changed to that category. In the single-input (S) conditions, only one controller was made available. One tablet was given to the pair in the HH condition, and one person was asked to step into the interaction area in the FB condition.

Figure 14. Users in the Full-Body Multi-Input (FBM) condition viewing their data at different decades and aggregation levels.

The conditions were tested in a 2x2 factorial design (see Figure 6). Of the 119 sessions included in this analysis, 57 utilized the full-body (FB) condition and 62 used the handheld (HH). Approximately half (57) of the visitor pairs were each given the opportunity to control their own data (multi-input, or M condition) and the remainder (62) used the single-input (S) condition.

The mechanics of switching among the conditions (adding or removing floor covering, resetting the system, and connecting the tablets as applicable) made random assignment impractical. Instead conditions were alternated on comparable days to keep samples as close as possible.

4.6.2 <u>Active and Conversational Time</u>

Prior to examining specifically what participants said during their conversations around the exhibit, it is important to look overall at the amount of time spent and dialogue occurring in each condition. Were there any inherent differences among conditions in how much time was spent and how much talking was done, it would impact later analyses quantifying that dialogue.

Active time in the session was defined as the number of seconds that both participants were in the interaction area and data was being displayed on the screen, after the instruction slides ended. Conversational time was calculated by subtracting the amount of time in the session coded as silence (any gap between idea units lasting longer than 1.5 seconds, see section 4.1) from the duration of active time in the session. A one-way analysis of variance (ANOVA) was conducted to compare the four conditions (full-body multiple, full-body single, handheld multiple, handheld single, or FBM, FBS, HHM, HHS) on the duration of visitors' sessions (active session time). The ANOVA revealed that the effect of condition on active time was significant, $F(3,115) = 2.782$, $p = .044$. The Handheld Single-Input (HHS) condition had the shortest average session duration at 145.8s, compared to 191.9s, 197.4s, and 201.6s for FBM, FBS, and HHM, respectively, but a Tukey HSD post-hoc comparison showed that no conditions were significantly different against any one other condition. No significant differences in conversation time were found across groups, indicating that the four conditions are comparable in the amount of talk people did. Therefore differences on other metrics described below were not likely due to chance differences in inclination of visitors to talk to each other.

4.7 Addressing Research Question 1a: Characterizing the Effect of Means and Distribution of Control on Learning Talk by Category

The first research question seeks to understand effects of the *means of control* (MoC) and *distribution of* control (DoC) on learning talk. As described in Section 4.3, all substantive talk was coded for dialogue determined to be useful for the group's learning process. Codes were applied from five categories—instantiate, manage, evaluate, integrate, and generate—and multiple codes could be applied to a single idea unit. This section examines effects of MoC and DoC on visitors' learning talk by first examining if the number of codes applied overall and by category differ by experimental condition (section 4.7.1), as a way of characterizing the content talk occurring in each condition. After this characterization, section 4.7.2 attempts to compare the *quality* of that content talk across conditions by examining the frequency of codes applied at each relevance level: highly relevant to learning goals, mid-relevance, and low relevance. Section 4.7.3 takes the comparison of content talk quality across conditions to a higher level of abstraction by computing and comparing overall session content scores. Section 4.7.4 uses these content session scores to probe the "efficiency" of the learning talk. Section 4.7.5 summarizes the findings with respect to the content of learning talk across conditions.

4.7.1 Impacts of Means and Distribution of Control on the Amount and Types of Learning Talk

A two-way analysis of variance yielded a main effect for the means of control, $F(1,115) = 5.841$, $p < .02$, indicating that the average number of codes applied to a session was higher for handheld sessions ($M = 56.61$, $SD = 29.23$) than full-body ($M = 44.16$, $SD = 27.09$). The main effect for distribution of control was non-significant, $F(1,115) = .668$, $p > .05$. The interaction effect was non-significant, $F(1,115) = .047$, $p > .05$.

In breaking down the number of codes by category, a main effect was found for the means of control on the number of *instantiate* and *integrate* codes applied, with the handheld conditions receiving more of each type. In both of these categories, no main effect was found for the distribution of control, nor was an interaction effect found. No significant main or interaction effects for the MoC or DoC were found for *manage*, *evaluate*, or *generate* codes. See Table 2 for full results.

TABLE 2. FREQUENCY OF CODES OF EACH CATEGORY APPLIED WITHIN EACH EXPERIMENTAL CONDITION

			Manage		Instantiate		Evaluate		Integrate		Generate	
MOC	DOC	*N*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
FB	M	28	9.39	7.73	21.39	12.71	4.86	3.09	8.29	5.97	1.82	2.13
	S	29	8.72	5.71	18.66	11.82	5.21	4.51	8.24	6.08	1.79	3.44
	Total	57	9.05	6.73	20	12.23	5.04	3.85	8.26	5.98	1.81	2.84
HH	M	29	10.97	7.50	30.93	13.32	5.52	2.80	10.24	5.66	1.83	1.67
	S	33	7.36	5.46	26.55	13.62	6.55	5.48	11.42	7.84	2.21	2.80
	Total	62	9.05	6.69	28.6	13.55	6.06	4.43	10.87	6.88	2.03	2.33
Total	M	57	10.19	7.59	26.25	13.77	5.19	2.94	9.28	5.85	1.82	1.89
	S	62	8	5.57	22.85	13.31	5.92	5.06	9.94	7.20	2.02	3.10
	Total	119	9.05	6.68	24.48	13.58	5.57	4.18	9.62	6.57	1.92	2.58

Taken together, these findings suggest that although no significant differences existed in conversational time among the four conditions, the means of control—handheld tablet controller or full-body interaction—affected visitors' conversations in some way. The next section examines how those differences manifested in terms of the learning goals of the exhibit.

4.7.2 Impacts of Means and Distribution of Control on the Quality of Learning Talk

Each code was classified as described above and in section 4.4 according to its relevance to the exhibit's learning goals. Here the relative frequencies of high, middle, and low-relevance codes are presented for each condition. Table 3 provides the mean number of codes in each category per condition.

TABLE 3. FREQUENCY OF THE CODES OF EACH LEVEL OF RELEVANCE APPLIED WITHIN EACH EXPERIMENTAL CONDITION

			Low-relevance		Mid-relevance		High-relevance	
MOC	DOC	*N*	Mean	*SD*	Mean	*SD*	Mean	*SD*
FB	M	28	34.71	21.21	7.46	4.59	3.57	3.83
	S	29	29.72	17.13	7.9	7.41	5	4.49
	Total	57	32.18	19.23	7.68	6.14	4.3	4.20
HH	M	29	45.76	20.24	9.24	5.13	4.48	3.53
	S	33	38.12	19.92	9.48	7.63	6.48	6.41
	Total	62	41.69	20.27	9.37	6.53	5.55	5.32
Total	M	57	40.33	21.28	8.37	4.91	4.04	3.68
	S	62	34.19	18.99	8.74	7.51	5.79	5.60
	Total	119	37.13	20.27	8.56	6.37	4.95	4.84

The average number of low-relevance, mid-relevance, and high-relevance codes by MoC and DoC were compared. A two-way ANOVA examining the number of low-relevance codes yielded a main effect for the Means of Control, $F(1, 115) = 7.23$, $p = .008$, with a significantly higher number of lower relevance codes for handheld sessions ($M = 41.69$, $SD = 20.27$) than full-body ($M = 32.18$, $SD = 19.23$). No other significant main or interaction effects were found for MoC or DoC in any of the other relevance levels. The high standard deviations present in these tallies and in the overall number of codes applied as reported in Section 4.7.1, along with the lack of statistically measurable differences by these metrics suggest the need for a more powerful method of analysis to tease apart effects of these design variations. This need is addressed by the weighted content scores, discussed next.

4.7.3 Comparison among Conditions on Learning Talk: Content Scores

As discussed in Section 4.5, this section of the analysis assigns weights to each code according to its relevance classification, with high relevance codes weighted as 3, mid-relevance

2, and low-relevance 1. Then weights for all codes applied to a session were summed to produce a content score for each session providing an overall picture of the quality of that session's learning talk. Overall content scores ranged from 10 to 242, with a mean of 69.42 and a standard deviation of 42.60. Table 4 presents descriptive statistics by condition.

TABLE 4. WEIGHTED CONTENT SCORES, BY EXPERIMENTAL CONDITION

			Content Score	
MOC	DOC	*N*	Mean	*SD*
FB	M	28	60.61	37.91
	S	29	61.03	42.07
	Total	57	60.82	39.72
HH	M	29	77.93	36.32
	S	33	76.79	50.23
	Total	62	77.32	43.93
Total	M	57	69.42	37.80
	S	62	69.42	46.89
	Total	119	69.42	42.60

A two-way ANOVA on the scores yielded a main effect for the means of control, $F(1, 115) = 4.52$, $p = .036$, with a significantly higher weighted content score for handheld sessions ($M = 77.32$, $SD = 43.93$) than full-body ($M = 60.82$, $SD = 39.72$). The main effect for distribution of control was non-significant, $F(1, 115) = .002$, $p > .05$, as was the interaction effect, $F(1,115) = .010$, $p > .05$. These findings reveal that, contrary to expectations, the handheld condition better supported learning talk among visitors.

As identified in section 4.4.1, a potential confound of weighted scoring is the numerical relationship between ordinal values, in this case low, middle, and high relevance to learning goals. To assess whether and how the 1-2-3 weighting system impacted the outcomes, two alternatives were tested. Alternative 1 gave low-mid-high codes numerical values of 1-3-5, and Alternative 2 used 1-5-10. Table 5 shows the means and standard deviations of these scores in comparison with the original 1-2-3 weighting, and Figure 15 presents profile plots for easier comparison.

TABLE 5. ALTERNATIVE WEIGHTED CONTENT SCORES, BY EXPERIMENTAL CONDITION.

			Content Score		Alternative 1		Alternative 2	
MOC	DOC	*N*	Mean	*SD*	Mean	*SD*	Mean	*SD*
FB	M	28	60.61	37.91	74.96	47.94	107.75	72.52
	S	29	61.03	42.07	78.41	56.28	119.21	90.79
	Total	57	60.82	39.72	76.72	51.91	113.58	81.80
HH	M	29	77.93	36.32	95.79	46.27	136.62	69.55
	S	33	76.79	50.23	99	69.01	150.39	113.82
	Total	62	77.32	43.93	97.5	59.03	143.95	95.21
Total	M	57	69.42	37.80	85.56	47.84	122.44	71.88
	S	62	69.42	46.89	89.37	63.72	135.81	104.05
	Total	119	69.42	42.60	87.55	56.47	129.4	89.96

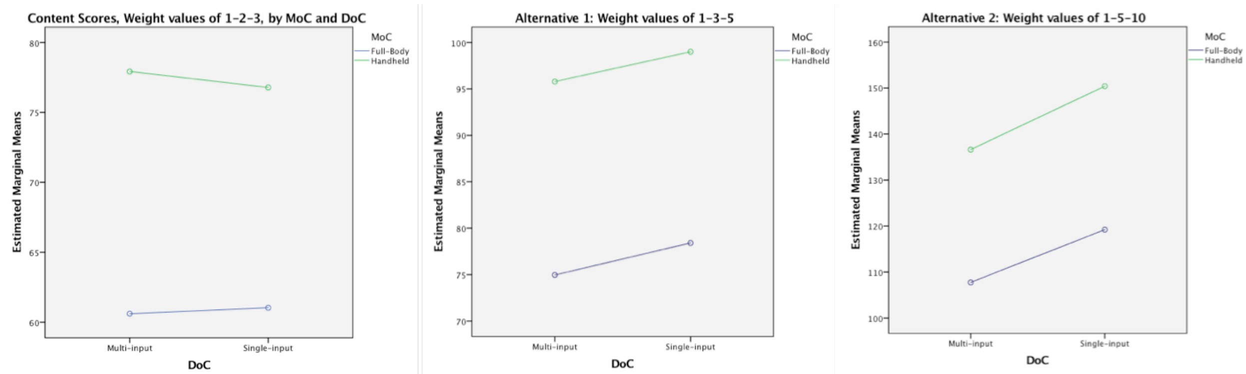


Figure 15. In all three weighting systems, the Handheld groups outperformed Full-Body groups.

These alternative scoring methods demonstrate consistently higher scores for groups using the handheld controller. As with the 1-2-3 weighted content scores, a two-way ANOVA for Alternative 1 (1-3-5 weighting) revealed a main effect for the means of control, $F(1, 115) = 4.02$, $p = .047$, with a significantly higher number for handheld sessions ($M = 97.50$, $SD = 59.03$) than full-body ($M = 76.72$, $SD = 51.91$). The main effect for distribution of control was non-significant, $F(1, 115) = .104$, $p > .05$, as was the interaction effect, $F(1,115) < .000$, $p > .05$. The second alternative weighting (1-5-10), showed the same relationship between MoC conditions, but these differences were not statistically significant, $F(1,115) = 3.33$, $p = .071$. The DoC and interaction effects for Alternative 2 were also non-significant.

A judgment is being made whenever a researcher (or teacher) makes the decision to assign value to an educational outcome, e.g., whether one test question should be worth the same amount of “points” as another. Because the three weightings presented here yielded consistent results with respect to the trends in performance in the various conditions, it can be concluded that any of the weighting systems would be valid for comparing conditions. This dissertation asserts that the 1-2-3 weighting is most appropriate because it gives proper credit to the low-level foundational talk needed to establish a shared understanding of the data, and offers more possible ways for companions and observers to make meaning from the statements, a critical value in a social learning setting.

4.7.4 Comparison on Learning Talk Across Conditions: Efficiency

In a free choice learning context with no time limit on the interaction, it may not matter how long it takes a visitor group to get to its content score, but assuming that for a variety of reasons hold times at any given display will vary, it may be beneficial from an exhibit design standpoint to have an efficient interaction in order to maximize the value of visitors' time spent at the display. Therefore, contents scores (using the 1-2-3 weighting) were also normalized by active session time to investigate differences in learners' score per second among the conditions. Normalized content scores were subjected to a two-way analysis of variance for the two means of control (full-body and handheld) and two distributions of control (single-input and multi-input). All effects were statistically significant at the .05 significance level.

The main effect of means of control yielded an *F* ratio of $F(1,115) = 32.47$, $p < .001$, indicating that the mean session score was significantly greater for handheld conditions (M = .464, SD = .160) than for full-body conditions (M = .311, SD = .133). The main effect of distribution of control yielded an *F* ratio of $F(1,115) = 4.82$, $p = .030$, indicating that the mean session score was significantly greater for single-input conditions (M = .421, SD = .187) than for multi-input conditions (M = .357, SD = .132). The interaction effect was significant, $F(1,115) = 4.40$, $p = .038$, indicating that the MoC effect was greater in the single-input condition than the multi-input condition.

The combination of the higher content scores for handheld conditions overall and the slightly shorter interaction times for the HHS condition led to the handheld single-input conditions producing the highest amount of data talk in the shortest amount of time.

4.7.5 Summary

Expectations based on theories of embodied cognition, our understandings of the challenges of sharing technologies in museums, and pilot work for this project were that a full-body interaction design and individualized control through multi-input controllers would best support visitor engagement with data and productive learning dialogue. However, this was not the case. Quite to the contrary, visitors using the handheld tablet controller overall engaged in conversations that were more strongly aligned with the exhibit learning goals over a shorter period of time, particularly when using a single-input controller. These differences cannot be accounted for based solely on differences in how much talking visitors did, as conditions were found to be comparable (see Section 4.6.2). Section 4.8 seeks to understand these surprising results by investigating visitors' interactions with the exhibit that may have affected the learning talk.

4.8 Research Question 1B: Exploring whether Differences in Interactions arising from Variations in Means and Distribution of Control Affect Learning Talk

The measures of learning dialogue reported in section 4.7 indicate surprising findings that, contrary to expectations, the handheld controller better supported visitor learning talk, with the Handheld Single-Input (HHS) condition yielding significantly higher scores than the other conditions. Prior work tells us that differences in interactions with an exhibit can greatly impact the experience and learning during such an interaction, and that different designs can produce different interactions (see Section 2.2). This section seeks to account for the unanticipated learning talk results by exploring main and interaction effects of the means of control and distribution of control on interactions; namely the amount of data rendered in each interaction (section 4.8.1) and the amount and frequency of control actions (section 4.8.2). To better understand the results in the preceding sections, I also undertook an examination of whether the quantity of operational talk (a proxy for how much visitors struggled with understanding how to operate the exhibit) was related to the quality of content talk (section 4.8.3), and if the distribution of productive talk over the sessions might indicate a longer learning curve with the exhibits in certain conditions (section 4.8.4).

4.8.1 Investigating the Relationship between Data Rendered and Quality of Content Talk

In this interactive exhibit, visitors were able to control not only what data were presented (via their selections at the kiosk prior to the interaction), but also the amount of data rendered during their interactions due to the control actions they took. Upon stepping into the interaction area, two datasets were initially rendered: each visitor saw his selected dataset for the first category (typically heritage) in one decade (typically 2010). Any further rendering of data was only accomplished through control actions to switch the category and decades. Because the design of the controller through which visitors interacted was manipulated in the experimental design of this study, differences were examined among conditions for how many datasets were *rendered* during each session. A dataset was considered to be rendered when it shown on the display for 2 seconds or more, giving the visitors the opportunity to engage with it. Both interacting visitors in a session had the potential to view their own self-selected data in four categories (heritage, household size, house type, and industry) over three different decades (1990, 2000, and 2010), but not all visitors explored all options. A binary representation (See

Appendix C: Data Rendering Coding Sheet) indicates by shading which combinations (category + year) were made present on the display by each user during the session. This metric shows which of the twenty four potential data sets (4 categories x 3 decades x 2 visitors) were rendered—shown on the display—during the session. For multi-input (M) conditions, any number between 2 and 24 datasets could be rendered; for single-input conditions only even numbers of datasets (2, 4, … 22, 24) could be rendered because one user's set could not be altered along a dimension without the other similarly being altered.

Of the 119 sessions analyzed in this study, 44 (37%) rendered all available data (24 datasets). The minimum number of rendered datasets was 8 (5 groups). The mean was 20, with a standard deviation of 4.7. A moderate but significant correlation was found between the number of datasets rendered and the content score of a session, $r(119) = .346$, $p < .001$.

To relate this finding to the controller designs, a two-way ANOVA on the number of rendered datasets yielded a main effect for the distribution of control, $F(1, 115) = 4.14$, $p = .044$, with a significantly higher number of rendered datasets for single-input sessions ($M = 20.84$, $SD = 4.77$) than multi-input ($M = 19.0$, $SD = 4.76$), which could possibly be explained by the "two-for-one" rendering attained by using the single-input controls. The main effect for MOC was non-significant, $F(1, 115) = .043$, $p > .05$, as was the interaction effect, $F(1,115) = 2.52$, $p > .05$.

The more data made available during a session meant that visitors had more opportunities to engage with data. Although differences in distribution of control were statistically significant, on average single-input groups saw fewer than two more datasets than their multi-input counterparts. This difference alone is unlikely to have greatly impacted the overall learning talk of the session, but it may have been partially responsible for differences among conditions. The next section looks more closely at engagement by examining differences in numbers of control actions.

4.8.2 Investigating the Relationship between Numbers of Control Actions and Quality of Content Talk

The two means of control tested here were both "embodied" in that they required physical movements on the part of the operator in order to complete. The full-body condition utilized bigger motions, requiring full-arm gestures and whole-body movements around the floor, while the handheld condition used the same shape of gesture but on a smaller scale, interacting within a 9-inch tablet screen (see section 3.1.1). Three types of control actions were available to the users: changing the category of data, changing the decade of data, and controlling the aggregation level of the data. In both MoC conditions, a change to the category affected all datasets regardless of whether it was in the multi- or single-input condition; at no time in any condition could one map show two different categories of data (e.g. household size and industry) at the same time. The aggregation and decade changes, however, were individualized in the multi-input conditions, allowing two users to view their own data in the decade and aggregation level they chose, regardless of the state of their companion's data (see Figure 14). In single-input conditions, decade and aggregation level for the two visitors were linked. Because of this difference in number of potential changes, DoC groups are analyzed here separately.

The first comparison is in the overall number of control actions taken by each group. Independent samples t-tests were conducted to compare number of control moves for the full-body and handheld conditions. In the multi-input conditions, participants in the handheld conditions made a higher number of control actions (M = 55.59, SD = 27.69) than did those in the full-body conditions (*M* = 39.00, *SD* = 14.38), $t(42.40) = -2.85$, $p = .007$. A t-test on the single input conditions indicated that no significant difference existed between handheld (*M* = 32.24, *SD* = 20.88) and full-body (*M* = 33.59, *SD* = 19.73) conditions, $t(60) = .259$, $p > .05$.

Session scores and total number of control moves were moderately positively correlated in the multi-input conditions, $r(57) = .51$, $p < .01$ and were strongly correlated in single-input conditions, $r(62) = .62$, $p < .001$. In particular, a strong correlation existed between the number of decade moves and session score in the single-input condition, $r(62) = .61$, $p <.001$. Decade moves were positively but weakly correlated with session scores in the multi-input condition, $r(57) = .27$, $p = .042$ (see Table 6).

TABLE 6. CORRELATIONS OF VISITOR MOVEMENTS AND TALK CONTENT SCORES

		Content Score	Total Moves	Aggregation Moves	Decade Moves	Category Moves
Content Score	Pearson Correlation	1	.511**	.531**	.270*	0.116
	Sig.(2-tailed)		0.000	0.000	0.042	0.39
	N	57	57	57	57	57
Total Moves	Pearson Correlation	.511**	1	.588**	.734**	.449**
	Sig.(2-tailed)	0.000		0.000	0.000	0.000
	N	57	57	57	57	57
Aggregation Moves	Pearson Correlation	.531**	.588**	1	0.083	0.025
	Sig.(2-tailed)	0.000	0.000		0.541	0.855
	N	57	57	57	57	57
Decade Moves	Pearson Correlation	.270*	.734**	0.083	1	0.015
	Sig.(2-tailed)	0.042	0.000	0.541		0.914
	N	57	57	57	57	57
Category Moves	Pearson Correlation	0.116	.449**	0.025	0.015	1
	Sig.(2-tailed)	0.39	0.000	0.855	0.914	
	N	57	57	57	57	57

These correlations suggest that physical interactions with an exhibit may provide an advantage for supporting learning talk. This finding would be in alignment with both theories of the affordances of interactive data visualizations (see Sections 1.1.1 and 2.3) and of embodied cognition (Section 1.3.3). However, this dataset does not support an analysis to tease apart which of these theories is more at play in this scenario; further work in this area is warranted.

If learning talk is correlated with control actions, the ease of completing those control actions is an important consideration for exhibit designers. While whole body interaction may have inherent affordances for cognition, those affordances may be negated by difficulties in completing the actions. Visitors were not directly surveyed about perceived difficulty of controlling the exhibit, so impacts of physical barriers cannot be measured using this dataset. The difficulty visitors had in *understanding* the control actions can be examined, however. A proxy for measuring this concern with the available data is presented next.

4.8.3 Investigating the Relationship between Operational-Researcher Intervention Time and Quality of Content Talk

It is to be expected that any novel interactive system will require some amount of time for users to learn and become comfortable with the controls and basic operation of the interactive features. Dialogue in this study dedicated to the operation of the system was marked *operational*; these operational statements focused on how to work any of the functions of the exhibit and did not fall into any of the management categories determined to be productive learning talk (see section 4.3.1). Though some operational statements included and were coded for instantiations, no other codes were applied to operational statements; if they served some other function, they were not considered operational. Additionally, researchers overseeing the data collection stayed in the exhibit room during sessions but only interacted with the participants during their interactions if the participants requested help or clarification. Instances of dialogue by researchers were coded *researcher intervention*. By summing the amount of time in seconds during each session devoted to operational or researcher intervention talk, I was able to create a new metric for comparing sessions, Operational and Researcher Intervention (OPRI) time.

OPRI time was interaction time that visitors were unable to engage in data talk, as their attention was elsewhere, so differences among conditions in how much time was spent on these control issues could affect the amount of productive data talk visitors could do and therefore impact visitors' session scores. In particular, it was hypothesized that visitors in the full-body conditions likely had a more difficult time learning the mechanics of the novel interaction than those in the handheld condition, and that multi-input users may have had to devote more talk to understand the control actions than single-input users, where one visitor in the interaction could "drive" the session. OPRI (operational-researcher intervention) time was calculated to account for this difference and provides a proxy for assessing how difficult it was for visitors to understand how the exhibit worked.

As reported in section 4.6.1, the HHS conditions had a lower mean interaction time of 145.86s, compared to the overall average active session time for all 119 sessions of 182.85 seconds, although a posthoc Tukey test didn't reveal any pairwise significant differences with any of the other 3 conditions. Because of the variation in session time, it is valuable to compute OPRI time as a proportion of total session time. Visitors using the handheld devices spent only about 8% of their active interaction time on operational talk, compared to nearly 20% in the Full-Body conditions. When examining the percentage of session time devoted to OPRI dialogue, a two-way analysis of variance yields a main effect for MoC, $F(1,115) = 52.55$, $p < .001$, indicating that users in the handheld condition indeed devoted a smaller percentage of their interaction to figuring out the mechanics of the system (*M* = 7.94, *SD* = 7.66) than those in the full-body condition (*M* = 19.47, *SD* = 9.50). The main effect for DoC was non-significant $F(1,115) = .515$, $p > .05$, as was the interaction effect, $F(1,115) = .002$, $p > .05$. This finding may be promising for embodied cognition theories: if participants in the whole-body interactive conditions spent on average 20% of their interaction time dealing with the mechanics of the system, it could follow that their content scores would be affected. Indeed, a significant but weak negative correlation was found between content scores and percentage of OPRI time across all sessions, $r(119) = -.248$, $p = .007$.

Subtracting the OPRI time from the active session times can therefore give a revised metric for viewing session length. This *non-OPRI* time is the amount of session time visitors were not actively distracted by the mechanics of the interactive system, and a two-way analysis of variance comparing MoC and DoC on this metric found the main effect of MoC was non-significant, $F(1,115) = .053$, $p > .05$, the main effect of DoC was non-significant, $F(1,115) = 2.22$, $p > .05$, and the interaction effect was non-significant, $F(1,115) = 3.48$, $p > .05$, indicating that once the time spent talking about the mechanics of the system was subtracted from the amount of time spent interacting, all conditions spent a comparable amount of time during which they could have been engaging in productive learning talk. Rather than differences in active non-OPRI time accounting for variations in session scores, this finding makes these content score differences—particularly between full-body and handheld conditions—even more striking.

In fact, a two-way analysis of variance examining session scores normalized by the active, non-OPRI time (in seconds) yielded a main effect for the means of control, $F(1,115) = 14.97$, $p < .001$, indicating that even when subtracting out the time spent learning the control of the system, visitors in the handheld sessions had a higher score-per-second (*M* = .50, *SD* = .17) than those in the full-body conditions (*M* = .39, *SD* = .16). The main effect for the distribution of control was non-significant, $F(1,115) = 3.09$, $p > .05$. The interaction effect was significant, $F(1,115) = 4.20$, $p = .043$, indicating that the MoC effect was greater in the single-input condition than in the multi-input condition.

These results suggest that it was not only the time spent learning the control of the system that accounts for variations in learning talk among conditions. The subsequent interaction analysis looks at the distribution of learning talk over the course of the sessions to determine whether the learning curve required to master the full-body interaction systems may have significantly affected the quality of talk early in the session but, once visitors caught on to the control actions, learning talk increased to levels comparable to the more easily operated handheld condition.

4.8.4 Exploring the Distribution of Learning Talk throughout Sessions

Previous sections reported on the differences among conditions on learning talk (as evidenced by content scores, section 4.7.2) and amount of time learning to operate the interactive system (section 4.8.3), finding that overall, the full-body conditions had lower scores and higher percentages of time spent on operational dialogue and researcher intervention. This final section examines the distribution of scores over the course of a session to identify whether time spent familiarizing themselves with the exhibit may have slowed the dialogue for visitors in the full-body conditions initially, thus pulling down their overall session scores.

This analysis was conducted first by dividing the total session time into thirds. For analyses conducted thus far using durations, the *active session time* duration was used, identified by a code marking the amount of time visitors were in the space while data were being displayed.

For most sessions, the active session time was uninterrupted, resulting in a single duration and start time for the active session. However, some sessions contained multiple active codes, for example if a session began and then was interrupted by technical difficulties or the instruction slides being re-triggered. In these cases, the *active session time* is the sum of the durations of the two (or more) *active* codes. For this final analysis, duration was recalculated to include the breaks between *active* codes by subtracting the initial start time from the final end time. For example, session FBM-S10 had two separate *active* codes. The first began at 34.7s through 1:04.5. At that time the participants requested the instructions be played again because they had missed them. The researcher restarted the instruction slides, which played until 1:38.1, and then the pair remained active until 4:25.6. For calculations reported previously in this study, this session was marked as having 197.2 seconds of active time (29.7s from the first segment and 167.5s from the second). For purposes of analyzing the distribution of comments, this session is counted as lasting 230.9 seconds, from the start time of 34.7 to the final end time at 4:25.6 (or 265.6 seconds). Fourteen of the 119 sessions in this analysis were adjusted in this way. To avoid confusion, this adjusted metric for session duration will be referred to as *real time duration*.

Real time durations were calculated for each session and divided by three, to grossly define the beginning, middle, and end segments of each session. The first third of a session was from the begin time of the initial active code; the middle third began at begin+(real time duration/3); and the third began at begin+((real time duration/3)*2). Any comments made during the instruction slides (before the official begin time of the session) were included in the first third's tally.

Because no significant main effect was found on DoC in the percentage of time users spent learning the system, this analysis focused on the MoC. T-tests were conducted to investigate differences between full-body and handheld conditions on the total number of codes applied in each third of the session, and the number of codes at each relevancy level according to learning goals in each third. Handheld conditions had significantly more codes applied in the middle of their sessions, $t(117) = -2.678$, $p = .008$ than full-body conditions, primarily because more low-relevance codes were applied in the middle portions of the handheld sessions than in the full-body sessions, $t(117) = -2.940$, $p = .004$. No other significant differences were found.

Session scores were also examined at the beginning, middle, and end thirds of visitors' sessions. An independent-samples t-test yielded no significant differences for MoC in the beginning or end thirds of visitors' sessions. In the middle third segment of sessions, a t-test yielded a significant difference between MoC conditions, $t(117) = -2.35$, $p = .021$, indicating higher scores for handheld conditions ($M = 28.18$, $SD = 17.28$) than full-body sessions ($M = 21.14$, $SD = 15.25$).

If barriers due to difficulties learning the interactive system had been responsible for significant differences between MOC conditions, we would have expected to see lower scores in the beginning of full-body control sessions and then scores catching up to their handheld counterparts in the middle and final segments of the interaction. However, these patterns did not occur. Handheld and full-body sessions were comparable in the number of codes applied and the score value of those codes in the initial and final segments of the session. Handheld users demonstrated more productive talk in the middle of their sessions, which is not directly attributable to the learning curve of the interactive system. This analysis can therefore not conclusively attribute differences between full-body and handheld sessions to initial difficulties in learning the control actions.

4.9 Summary

This chapter reported results from a quasi-experimental study investigating main and interaction effects of the *means of control* (MoC) and *distribution of control* (DoC) on visitor dialogue around an interactive data map exhibit. Participants in this study interacted in pairs with one of four versions of the *CoCensus* exhibit: full-body, multiple input (FBM); full-body, single input (FBS); handheld, multi-input (HHM), or handheld, single-input (HHS). Visitor dialogue was coded according to categories of talk known to be productive for learning both with data visualizations and in informal learning environments. These codes were then classified according to their relevance to the learning goals of the exhibit. Section 4.7 examined differences among conditions in the type of codes applied, the frequency of codes applied in each session in each relevance classification, and the learning talk content scores of each session as calculated by codes weighted according to their relevance to the exhibit's learning goals. Prior work and theories of embodied cognition suggested that full-body interaction and distributed control would productively mediate dialogue. However, the analyses presented here demonstrate that the handheld sessions outperformed the full-body sessions on most metrics, in particular on content scores.

To account for these surprising results, variations in other aspects of the interaction were inspected in Section 4.8 to determine whether the amount of data rendered during a session, the number of control actions engaged in, the amount of dialogue dedicated to learning the interactive system, or the learning curve for understanding the interactivity could have impacted the visitors' learning talk.

While main and interaction effects were demonstrated in these interaction measures, none fully explain the results. Single-input users saw more data rendered during their interactions, but only an average of two additional datasets out of 24 available, which is unlikely to be enough to account for differences. Moreover, no differences were found in MoC on data rendering. Some correlations between specific control actions and content scores were observed, but the data here do not support thorough analysis of the underlying causes. Future work in this area would be of great interest. Full-body system users were found to have spent more time talking about the operational aspects of the exhibit, but removing time devoted to operational talk and interactions with researchers only exaggerated the differences in content scores between MoC conditions. Finally, the distribution of dialogue over the course of the sessions was examined to understand whether full-body participants had a slower start in discussing data due to the difficulties learning the system, thus dragging down their overall score. If this learning curve issue had been a major factor, we would have expected a difference in the beginning part of sessions, but this did not occur. Additional hypotheses for the success of the handheld—and specifically the handheld single-input—design that are not explorable with the current dataset are discussed in Chapter 6. These hypotheses will suggest directions for future work.

5 MEDIATIONAL EFFECTS OF A PSYCHOLOGICAL TOOL — ANALYSIS AND FINDINGS

The previous chapter reported impacts of a *physical tool* mediating visitors' interactions with a data exhibit. This chapter looks at another kind of cultural tool, a *psychological* tool called perspective taking (see section 2.4) to understand how it might be mediating visitors' experiences with the exhibit. Specifically, I am examining a form of perspective-taking, *actor perspective taking* (APT), that is present in this setting.

5.1 Research Question 2A: Relating Actor Perspective Taking to Learning Talk

Perspective taking has been studied in a variety of contexts and generally refers to how learners are positioned with respect to the content or learning environment (see section 2.4 for a deeper review). Some studies directly manipulate the perspective learners take by assigning them roles (Filipi & Wales, 2004) or designing a digital environment to induce a particular perspective (Lindgren, 2009). This study follows Ochs, Gonzales, and Jacoby (1996) in analyzing the spontaneous use of *actor perspective taking* (APT) in meaning-making conversations. This spontaneous adoption of the actor perspective—that is, when visitors speak of the data in the first person, making statements like "that's me" or "I'm all on the north side"—was observed in visitor dialogue in pilot work for this project. Analysis of these pilot sessions showed a correlation between usage of the actor perspective and data talk (Roberts et al., 2014). This study seeks to further that analysis to deepen our understanding of what perspective taking can be good for in this context.

Idea units were coded for APT when the visitor's choice of words indicated she was viewing herself as part of the dataset. For example, when describing a dispersed dataset a visitor could say, "They are all over the place," or "It's all over the place," or "The blue is all over the place," or "Germans are all over the place." Each of these statements positions the data as something external to the visitor. By contrast, sometimes visitors phrased such a statement differently, as in, "I'm all over the place." This small linguistic difference doesn't alter the meaning or purpose of the statement, but it may indicate a difference in how the visitor is thinking about the data and how connected she feels to it. When a visitor verbally assumes an actor perspective, I propose that he is using that perspective either to concretize his own understanding or to help his companion understand or see the data in a particular way, in short, making it a tool to aid in the group's meaning making.

Importantly, this analysis does not attempt to identify individual cognitive use of APT for understanding the exhibit content. As with the analysis presented in the previous chapter, APT coding only considered the content visitors put into the shared social space through dialogue. While some visitors may have been mentally adopting a first-person perspective more often than instantiated, just as they were likely doing more reasoning about the data silently than they said aloud, this analysis is concerned with the use of APT as a tool for mediating the conversation, not for mediating individual cognitive tasks. For that reason, I also coded statements where the speaker placed his or her companion within a certain perspective (for example, "That's you over there") as APT, because it is positioning the speaker's partner as an actor "within" the data set. As with the previous chapter's analysis, that contribution to the group dialogue is the focus of this socioculturally-framed study.

Of the 119 user sessions included in this study, 54 included instances of visitors spontaneously adopting the Actor perspective at least once, for a total of 122 idea units coded as APT. Of those 122, three idea units were found to be in need of further segmentation because, as will be discussed in section 5.3, the function of APT in those statements changed even though the broader purpose of the idea unit did not, resulting in a total of 125 coded APT statements. In half of the APT sessions (*N* = 27), APT was utilized only once. The maximum number of APT statements coded in a session was 9 (1 session). Of the 54 sessions using APT, the average number of uses was 2.3, with a standard deviation of 1.8. So while APT is not a dominant method of expression used by visitors, it does spontaneously occur in a significant minority of sessions.

This study hypothesized that there may be a relationship between APT and learning talk. Session codings and content scores for groups who used APT and those who didn't were compared using independent samples t-tests, and it was found that sessions in which APT was used at least once had a significantly higher total number of learning talk codes applied (M = 58.22, SD = 28.78) than sessions where no APT was used (M = 44.35, SD = 27.46), $t(117) = -2.684$, $p = .008$, and APT sessions had a higher average content score (M = 78.46, SD = 43.76) than sessions where no APT was used (M = 61.34, SD = 39.65), $t(117) = -2.238$, $p = .027$. These findings confirm the positive relationship between APT and learning talk seen in pilot data; the following sections inquire more closely into the mechanics of that relationship.

5.2 Actor Perspective Taking and Group Composition

It is known that group composition affects dialogue in museum learning (Ash, 2003; Falk, 2006; Atkins et al., 2009; Allen, 2002; Schauble et al., 2002; see section 2.1.1). Therefore it was of interest whether mixed-age (i.e. family groups, *N* = 57) were using APT differently than were homogenous adult-adult or child-child (peer) groups (*N* = 62). The mixed age groups comprised 34 adult-child pairs, 11 adult-teen pairs, one teen-child pair, and 12 family groups (3 or more visitors comprising at least one adult and one child). The peer groups comprised 41 adult pairs, two groups of 3 or more adults, 9 child pairs, and 9 teen pairs. A t-test determined no significant differences existed in overall content scores between peer groups (M = 68.00, SD = 46.54) and mixed age groups (M = 70.28, SD = 37.62), $t(117) = 117$, $p > .05$.

Content scores were then subjected to a two-way analysis of variance for two group composition types (peer and mixed) and whether or not APT was used in the session. Consistent with results reported above, the main effect for APT use was significant, $F(1,115) = 4.907$, $p = .029$. No significant main effect was found for composition class (peer or mixed), $F(1,115) = .165$, $p > .05$. However, an interaction effect was found, $F(1,115) = 4.910$, $p = .029$, indicating that the APT effect was greater in the peer groups than the mixed age groups. In fact, the mixed groups had virtually identical content score means regardless of whether they used APT or not: M = 70.28 for mixed groups with no APT (N = 36, SD = 39.13) compared to a mean of 70.27 for those who did use APT (N = 22, SD = 35.90). The peer groups, however, showed a drastic difference. Peer groups who did not use APT (N = 29) had a mean score of 50.24 (SD = 38.08) while peer groups who used APT (N=32) had a mean score of 84.09 (SD = 48.17).

These results suggest that, in alignment with prior work, a relationship exists between visitors' use of the *Actor* perspective in dialogue and their ability to productively talk about data in this environment, but these data show that the affordances for this effect only apply to peer groups, not adult-child pairs or family groups. In light of these findings, the following sections look at the ways visitors used APT to mediate their conversations, with special attention paid to the differences in how mixed and peer groups utilized these affordances.

5.3 Research Question 2B: Characterizing the Applications of Actor Perspective Taking in Visitor Talk

The previous section established a relationship between visitors' spontaneous use of the actor perspective and higher average content scores, though this relationship only manifested in peer groups. Further exploration into potential causality would be of great interest (see section 6.2), but the data collected for this study do not support investigation into a causal relationship. However, developing an understanding of how APT was used by groups can provide insight into its potential mechanism as a tool for reasoning about data in this context, and to how APT usage here relates to work investigating perspective taking as studied in other disciplines. This section identifies the different ways APT was applied by different learners.

5.3.1 Defining Different Applications of APT in Learning Talk

If the actor perspective is a tool, it is one that can be used differently by different people throughout their interactions. Each APT statement in this dataset was examined to determine how it was being applied, using an interpretive coding approach (Miles & Huberman, 1994) intended to group similar uses together while maintaining nuance. Twelve unique applications of APT use were identified (see Table 7). Though some of the twelve categories themselves align with codes applied to learning talk as described in the previous chapter (for example, the *comparison* APT code corresponds to the INTEGRATE: compare learning talk code, and the *characterize* APT code matches the EVALUATE: characterize learning talk code), the low overall usage of APT doesn't allow us to responsibly use statistics to determine if the different APT applications correlate with particular codes of learning talk or if higher content scores can be attributable to any specific APT application.

One other difference between APT codes and learning talk codes is that the categories presented here are mutually exclusive: an APT statement is categorized here according to the one descriptor best illustrating how it is being used. For example, a statement such as, "I'm overpowering you," is a comparison between "I" and "you," but because it has a competitive component it is categorized as *competitive* rather than *direct compare.* I argue that when one takes a competitive stance towards another person, this is a fundamentally different application of perspective taking than when one is taking a merely comparative stance – it has strong implications both for how the listener will receive the statement, and for how the speaker will behave after taking that perspective. These twelve codes were applied by a second coder to an 11% sample of the APT statements, or 14 idea units, with a 92.8% rate of agreement. Cohen's kappa was run to determine if there was agreement between both coders on application of the 12 APT codes, and there was strong agreement, $\kappa = .915$, $p < .001$.

TABLE 7 APPLICATIONS OF ACTOR PERSPECTIVE TAKING OBSERVED IN SPONTANEOUS VISITOR TALK

Interpretive Code	Description	Examples
Avatar	APT statements describing how and where the visitor as a member of this dataset would live.	"You live downtown, I would live in Queens." "Ah, that's where I'd work."
Competitive	Make a comparison, but with a competitive tone.	"You dominate." "I win."
Direct Compare	Compare spatial or quantitative aspects of two datasets.	"There's more of me than you."
Direct Connection	Identifying oneself in the dataset.	"There's one English person, that is you."
Extrapolate	Draw a conclusion about the population based on the represented data.	"We have more opportunities there." "We work harder."
Generalize	Make a general descriptive statement about the dataset.	"We don't move." "Our people don't disappoint in health and social services."
Notice Absence	Identifying a lack of one's data visible in the representation.	"I'm not anywhere." "Where's our presence in the Bronx?"
Orienting	Using APT to interpret the legend and establish the connection between themselves and the representation.	"I'm purple" "That's me."
Quantitative Characterization	Use APT to describe the quantity of a dataset.	"There's a lot of you." "We're depleting."
Quantitative Inferences	Draw a quantitative conclusion.	"I'm so common." "I'm the majority."
Representation	Use APT to directly connect oneself to the representation (the bubbles or the color).	"You mixed us together." "Gather yourself together."
Spatial Characterization	Use APT to describe the spatial distribution of a dataset.	"You're really spread out." "Wow, I'm like everywhere."

5.3.2 Exploring how Peer and Mixed Groups Applied Actor Perspective Taking

To attempt to explain the differences in peer and mixed groups in content scores related to APT use, the frequency of usage of each APT application was analyzed by group type. Figure 16 shows the average per-session frequency of the different APT applications made by peer versus mixed groups. The most common APT application for both group types was *orienting*, which is not surprising: these statements represented a very basic connection to the data, often literally saying "That's me." Such simple connections were useful in establishing the overarching design of the exhibit, i.e. that the visitors' responses to the survey they just took at the kiosks were now being displayed on the map. This grounding set the stage for data exploration.

Other applications, in particular *avatar*, in which a visitor pretends to be an individual in the dataset (e.g. "I would live in Queens."), were drastically higher for peer groups, approaching the frequency of *orienting*. By contrast, mixed groups used APT more often for *directly comparing* two datasets (e.g. "There's more of me than you.") than did peer groups. Some of the other APT categories were more common in peer groups (*competitive comparison*, *direct connection*, *extrapolation, notice absence, spatial characterization*), but the frequencies were fairly low, preventing strong claims. Nonetheless, it seems that several of the categories favored by peer groups (*avatar*, *direct connection*, *extrapolation*) may be interpreted as requiring the participant to more thoroughly inhabit the kind of first-person actor perspective most akin to the way the actor perspective has been studied in prior work, in which the visitor is connecting to the data on an individual level. Other applications of APT found here indicate a different connection seemingly unique to this context in which the visitor is projecting himself to a larger group, an idea I will return to in greater depth in Section in 5.4.1. The small sample size of this dataset does not permit reliable statistical analyses of correlations between these APT applications and content scores, but future work in this vein would be of great interest in attempting to understand differences between peer and mixed-age groups.

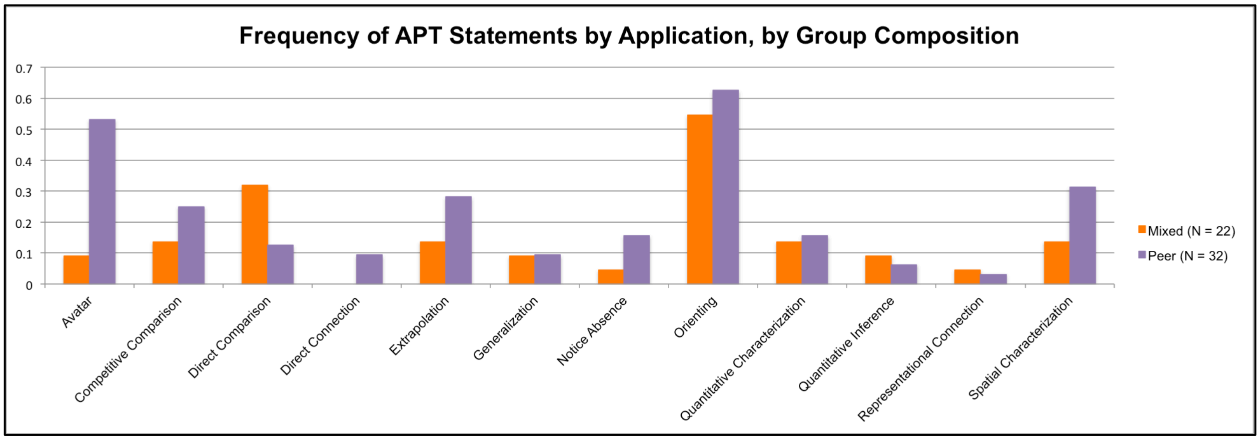


Figure 16. Average use of the different APT applications in peer versus mixed groups.

5.3.3 Exploring Temporal Applications of APT by Peer and Mixed Groups

The analysis of the intersection of interaction and learning talk in Section 4.8.4 broke down sessions into thirds (beginning, middle, end) to see if participants changed their engagement with the exhibit over time. Perspective-taking may also be a practice which changes over time, and when in a session visitors employ APT may shed important light on how it is being used. The preceding section, 5.3.2, showed that both peer and mixed groups engaged in a large amount of *orienting* APT, which we might expect to happen early on in a session. But peers and mixed groups differed in the application of other APT categories. Could the distribution of APT throughout the session provide insight into the different perspective-taking applications of peers and mixed groups? If, for example, peer groups tended to adopt certain applications of APT earlier in their session, it might set the stage for further APT applications, or frame the remaining data interpretation in a more meaningful way.

To further break down APT application, the next step of analysis returned to the distribution of APT throughout the session using the beginning, middle, and end segments introduced in Section 4.8.4, looking at mixed and peer groups separately. The overall distribution of APT statements over the session (beginning, middle, and end) was comparable between peer and mixed groups. Peer groups had a total of 86 APT codes applied. Of these, 34 (or 40%) were in the beginning segment of the sessions, 21 (24%) were in the middle, and 31 (36%) occurred at the end of the session. In mixed groups the 39 APT statements were distributed as 14 (36%) in the beginning, 10 (26%) in the middle, and 15 (38%) at the end. The lack of difference between mixed and peer groups in when APT (in general) was used indicates its position in the session does not contribute to the differences between content scores of these groups. Looking at the different specific ways APT was applied, though, begins to reveal a potential explanation.

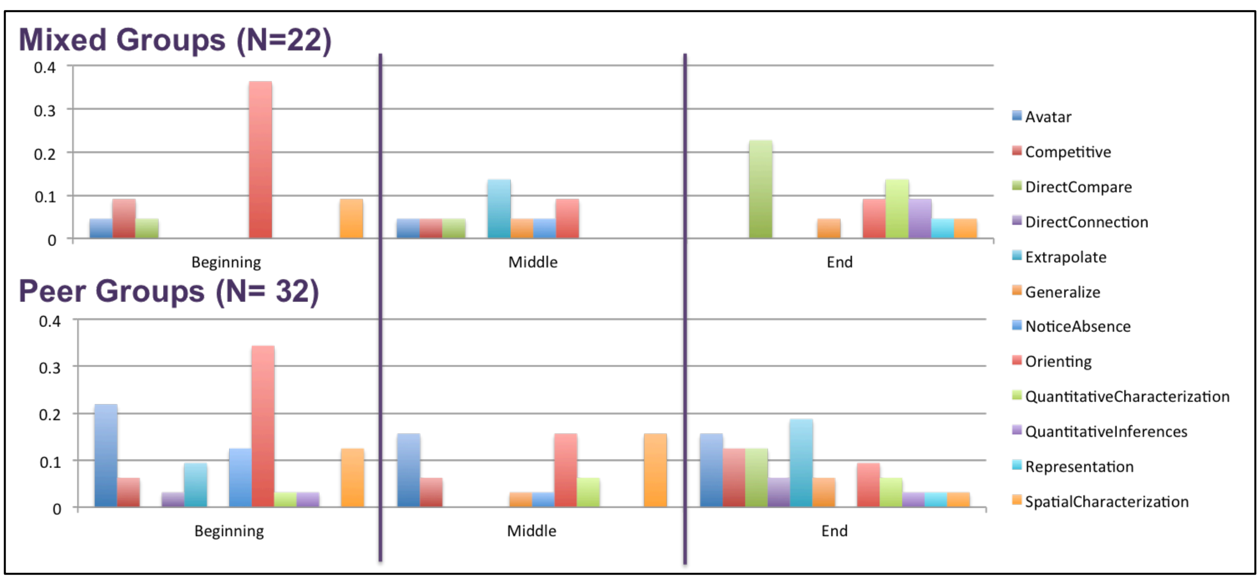


Figure 17. The frequency of APT application by each third (beginning, middle, and end) of a session, for mixed groups versus peer groups.

Figure 17 illustrates which specific APT applications were used in the beginning, middle, and end thirds of the sessions for mixed versus peer groups. Unsurprisingly, both groups see the majority of *orienting* APT in the first third of the session, but apart from that, the group types exhibit very different trends. This decomposition reveals an interesting and potentially fruitful difference between peer and mixed groups overall. Particularly in the beginning and end of sessions, peer groups seemed to be using APT in a wider variety of ways than mixed groups. While peer groups in the beginning of their sessions overall used APT for 9 of the twelve identified applications, mixed groups only used 5 of them, and most APT statements were categorized as *orienting*. A similar pattern exists in the end of sessions, where peer groups used 11 of the 12 APT applications and mixed groups only 7. The flexibility in APT application by peer groups may indicate that they are approaching their perspective-taking in a more "thorough" way, i.e., that for these visitors, there is a stronger perceived connection between the visitor and the data. Section 5.1 established the positive correlation between APT and content scores. Might a stronger visitor-data connection, as indicated by a richer application of APT, be the reason behind the higher content scores of the peer groups?

At the individual session level, most peer and mixed groups applied APT between one and four different unique ways, with peer groups averaging 1.81 (SD = 1.06) and mixed groups averaging 1.55 (SD = 0.86). A single peer group used APT five unique ways. Two outliers in the peer groups were removed because their content scores fell more than 2.5 standard deviations above the mean, and the number of APT functions used by each group was plotted against that session's content scores (Figure 18). Over all remaining APT sessions, a correlation was found between the number of unique APT applications and the session's content scores $r(52) = .486$, $p < .001$. Broken down by group composition, the mixed groups showed a stronger correlation, *r*(22) = .651, *p* = .001 than peer groups, but scores and APT functions were still correlated in peer groups, *r*(30) = .386, *p* = .035.

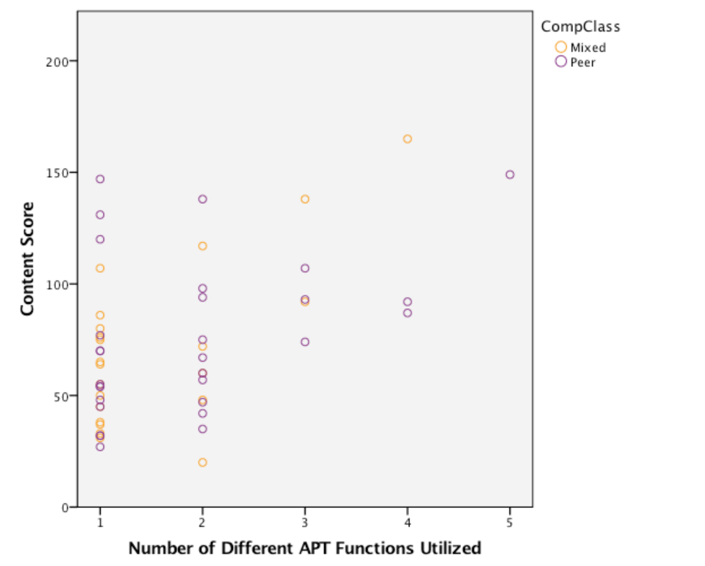


Figure 18. Correlations between variety of applications of APT and content scores

From this dataset we cannot make firm claims about a causal relationship, but the correlation suggests potentially interesting areas for future study. The next section identifies patterns in these functions to demonstrate how they supported different kinds of relationships between the visitor and the data.

5.4 Two Ways of Blending Identities

In close analysis of the APT statements, it became clear that the different applications of APT identified in Section 5.3 likely afford different kinds of relationships between the visitors and the data. The temporal breakdown of APT in Section 5.3.3 implied that many visitors used APT at a very basic level to orient themselves to the data, typically at the beginning of their sessions. *Orientation* statements, such as “That’s me,” or “I’m purple,” help visitors recognize the connection between themselves and the data, setting the stage for more productive talk. Of the 54 sessions in which APT occurred, 26 used it for orientation. In 17 sessions, orientation was the only way APT was used. This APT use may be helpful in laying the foundation for future talk, but it doesn’t seem to deeply represent a true blending of self-and-data as described by Ochs et al. (1996), and is therefore not the focus of the following analysis.

Ochs and colleagues posit that the *indeterminate constructions* used by physicists form an “extreme form of subjectivity in which the distinction between the scientist as subject and the physical world as object is blurred.” In making sense of a complex physical phenomenon, they claim, the physicists in a particular lab referred to themselves as the particle being studied in order to understand its trajectory through various states as represented by a graphic display. This linguistic oddity was most frequently initiated by the most senior scientist, the head of the lab, though his students picked it up in group discussions. Furthermore, it was found to be tied to the use of gesture and graphic representations; Ochs et al. conclude that “graphic representations can referentially constitute scientists and physical entities as simultaneous, co-existing participants in events.”

The following sections will demonstrate that in this informal context of data interpretation and reasoning, non-experts (museum visitors with no assumed training in demographic data or geographic information systems) construct similar blended identities as they interpret a graphic representation (the interactive census data map), and furthermore they situate themselves as "co-existing participants" with the data in two unique and productive ways: Some visitors used APT to link the visitor to an individual in the map in a type of *role play*, in which visitors sought to answer, broadly, *if I were a person in this map, where would I be?* This manifests in statements like, "You live downtown, I would live in Queens," and "So where do you work?" This type of *role play* dialogue using APT is discussed in section 5.4.1 below. Other APT statements found in this dataset used personal pronouns—including singular pronouns—to project their identity to a group of people. For example, several groups made statements like, "There are more of me than you" or "I'm congregating in Queens." A single individual can't congregate, and one person is equal to, not more than, one person, so we can infer that the singular pronoun was referring to the whole category of data, indicating the visitors' *projection* of himself as part of the dataset as a whole. This APT purpose of *projecting* is discussed in section 5.4.2.

These two types of relationships between self-and-data emerged concurrently with the APT applications identified in the previous section as it became clear that some APT statements referred to groups (*projection*) and some to individuals (*role play*). The projection and role play codes were applied to individual APT statements, but it was found that these categories generally aligned with the APT applications defined above. For example, comparisons and characterizations aligned with group *projections* and avatar and direct connections fell under individual *role play*. Two applications —extrapolate and notice absence—were used at both the individual and group level. Figure 19 illustrates the classifications of each function code. The same 11% sample of APT statements coded by a second researcher for the APT applications as reported in Section 5.3.1 were coded for self-to-data relationships, and again a 92.8% agreement was reached (13 of 14 idea units coded identically), with a Cohen's kappa indicating there was strong agreement, $\kappa = .882$, $p < .001$.

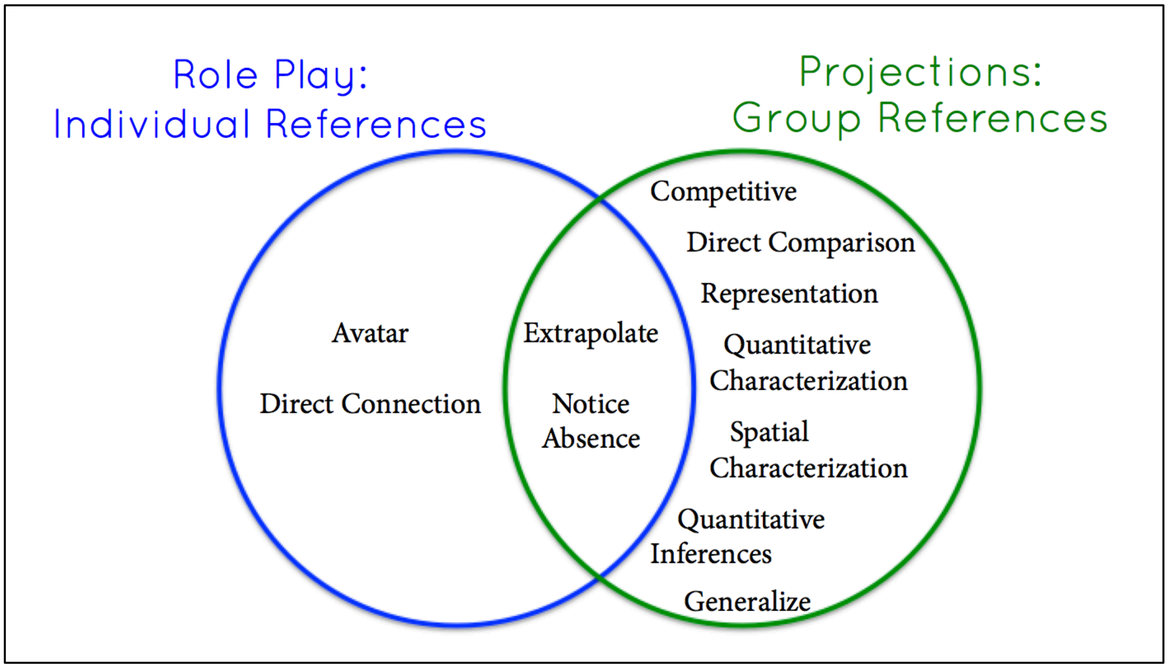


Figure 19. Venn diagram showing relationship between APT applications and self-to-data relationships

5.4.1 Role Play: Qualitative humanizing of census data

One of the overarching learning goals of the *CoCensus* exhibit is to get people to playfully engage with and productively talk about census data. The challenge of this task lies in the somewhat abstract and opaque nature of this immense dataset. Though census data is intended to directly reflect people, people can have a hard time finding themselves in the numbers. *CoCensus* addresses this problem through several design decisions to personalize and humanize the dataset, including narrowing the corpus of data to variables expected to be especially relevant to people's lifestyles, connecting the data to a familiar geographic area, and making the display interactive to let people explore what they found interesting rather than imposing a predefined narrative.

APT in general can be considered a successful manifestation of this goal: people adopting the actor perspective are clearly connecting to the data in some way, literally saying, “That’s me!” However, role play as identified here is a special function in which visitors are more fully adopting the subjective perspective of a character in the map. We saw this phenomenon first in an earlier version of the CoCensus project, with two visitors referred to as Belle and Peg (Roberts et al., 2013; Roberts, Lyons, and Radinsky, 2013; Lyons & Roberts, 2014, see section 2.4.2). This pair of adult colleagues playfully adopted their personas as “upscale” British people living on the lakefront and partying in boats and Polish “farming folk … working in the fields” based on the distribution of the two populations in relation to Lake Michigan in Chicago. The pair used APT to connect to the data, identify patterns, and pose inferences and ask questions about those patterns.

Visitors using the iteration of the exhibit studied here made similar statements. Some were neutral statements like, “This is where I would work,” and “I’d live downtown,” but many also included evaluative and extrapolation statements like, “Oh, so you’re saying I’d live in like the cool area,” and “I was always important.” The latter kind of statement is particularly valuable for the exhibit goal of humanizing census data: visitors employing this kind of APT use are not just identifying quantitative patterns but are thinking more deeply about what it means to be a person with certain demographic characteristics in a certain geographical place. To demonstrate this mode of APT use in action, in particular how—as noted by Ochs and colleagues (1996)—participants unproblematically transition from participant-centered, data-centered, and blended perspectives frequently within and across turns, Table 8 below adopts the representation used by

Ochs et al. to illustrate one excerpt of a role playing group. This pair of adults had spent over 90 seconds in the exhibit space without really understanding the purpose of the exhibit. Their dialogue to this point was almost entirely instantiations, occasionally discussing what to do next with each other or the researcher. The segment below—read left to right in the table—shows the moment they begin to understand the exhibit: in line [34] the male uses APT to put himself in the data, saying "So that's where we'd live." This triggers the female participant's understanding ("Oooh"), which she follows with an immediate adoption of similar APT role play ("and that's where I'd work.").

TABLE 8. TRANSCRIPT EXCERPT OF SESSION HHS-S32, DEMONSTRATING *ROLE PLAY* APT USE

Line	Participant-Centered	Data-Centered	Blended Role Play (Self-to-Individual)
[34]	M: What do you choose on your one?		
[35]		F: What, house type?	
[36]		M: Yeah.	
[37]		F: Single detached house.	
[38]	M: Yeah, so did I.		
[39]	F: Yeah.		
[40]	M: So we both took the same one.		
[41]			M: So that's where we'd live.
[42]			F: Ooohhh, and that's where I'd work.
[43]			M: And that's where I work, construction, and you'd work in educational services.
[44]			F: So I'm like by the Hudson River.
[45]			M: Yeah.

It was only after this initial APT-mediated role play that the two began productively engaging with the display and discussing the content in depth. Their final content score for the session was 107, almost a full standard deviation above the mean (for all sessions). Of that score, 28 points came before APT was used and 79 came after. It was only once they began seeing themselves in the data that they were able to engage in productive learning talk (see Figure 20).

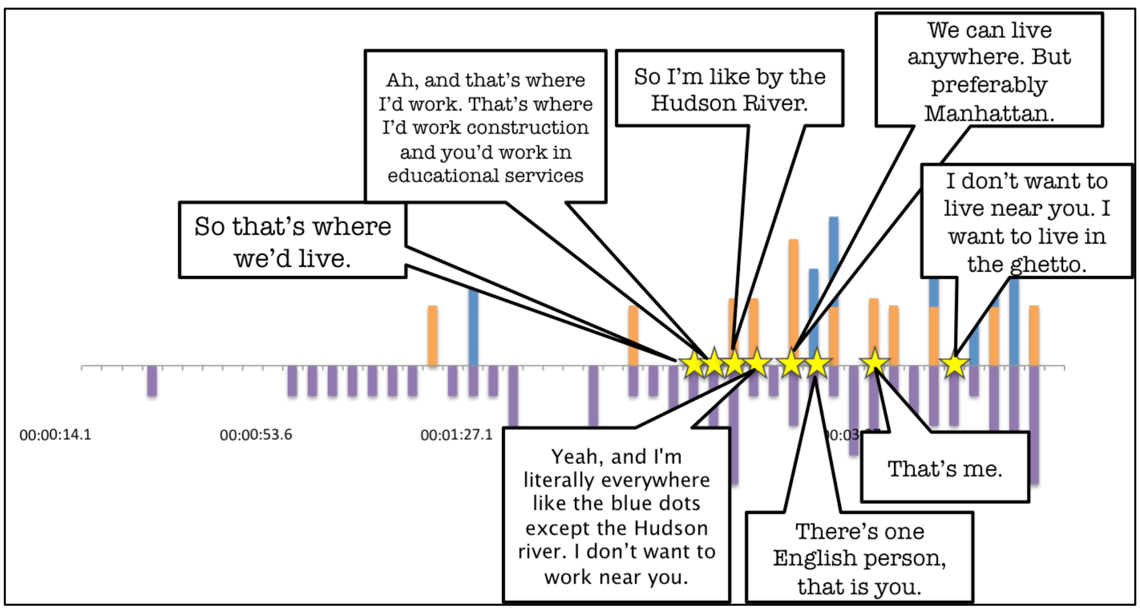


Figure 20. Timeline of HHS-S32 session, with APT statements annotated

This kind of blend—the visitor substituting himself for a person represented by the data—is most akin to the indeterminate constructions identified by Ochs and colleagues. It is not, however, the only kind of blended identity occurring in this dataset. Many of the APT statements connected a visitor not to an individual but to a group. This self-to-group projection is discussed next.

5.4.2 Projection: Blending "I" and "They"

Another important goal of the *CoCensus* exhibit, beyond humanizing census data as described in the previous section, is to help visitors discuss quantitative and spatial characteristics of the data. The *role play* blended identity discussed above doesn't necessarily support this goal. Some visitors, therefore, used APT instead to speak for a group of people, projecting themselves onto the whole dataset. *Projection* APT statements were comments such as, "You guys dominate," "There's a lot of me," "We make a pattern," and "You're really spread out." In each case, "I" and "you" become referents not to any one person but to the whole dataset, allowing the visitor to make observations and inferences about the whole population. The below excerpt is from an adult group of three women in the handheld multi-input condition (HHM-S68, Table 9). In this excerpt the two participants start using APT to make a quantitative statement about the data representation, specifically that visitor B's circle representing her heritage ("Other Hispanic") is "dominating" the map in 2010: she had just aggregated the data to the city level, so the single pink dot representing her data was covering most of the map. A few seconds later (line [8]), they had gone back on the timeline to 1990 and had both disaggregated to the census tract level, but user A still uses the "dominate" characterization to compare the quantities of Other Hispanic and Indian (user A's selected heritage). A few seconds after that, the two switched to a *role play* perspective when they couldn't find user B's dataset for household size, further indicating the flexibility with which they employed APT in their dialogue.

TABLE 9. TRANSCRIPT EXCERPT FROM HHM-S68, ILLUSTRATING *PROJECTION* APT USE

Line	Participant-Centered	Data-Centered	Blended Role Play (Individual)	Blended Projection (Group)
[1]	A: I live around that area, though.			
[2]		B: Where? Where all the purples are?		
[3]		A: Yeah.		
[4]	B: Why is mine so big?			
[5]				A: 'Cause you dominate.
[6]	B: Oh now we could look at... let's look at 1990.			
[7]		B: Wow, still...		
[8]				A: There's so little. You still dominate.
[9]		B: Household size.		
[10]		A: Five.		
[11]			A: Where are you?	
[12]	B: I don't know.			
[13]			B: Oh there I am.	
[14]	A: Oh.			

The content score distribution before and after APT usage in this group was more balanced than in the previous example, with 61 points of the content score occurring before the visitors started using APT around 2 and a half minutes into their session and 86 points after (in the remaining 3 minutes), but most of the the pre-APT statements were low-relevance codings.

The mid-relevance and high-relevance codes were much more frequent after the group used APT, see Figure 21.

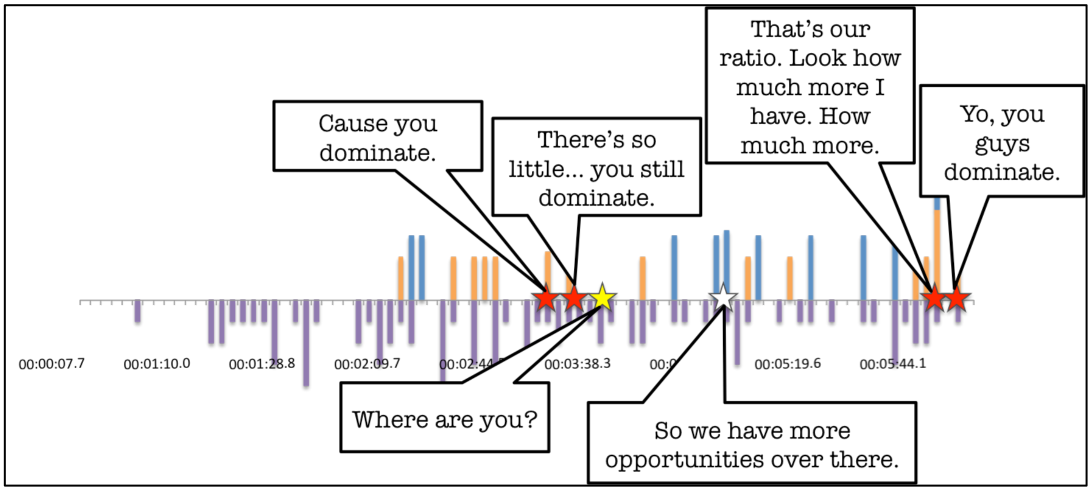


Figure 21. Timeline of HHM-S68 session, with APT statements annotated

Of particular interest in this excerpt is the choice of *competitive* language to characterize the size/quantity of user B's data with the word "dominate." This competitive language was seen in conjunction with APT in pilot work for this project, with visitors often talking of "winning" or even asserting, "I slaughter you!" to compare one visitor's data to another (Roberts et al., 2013). Even though this type of competition is associated with self-to-group *projection* APT usage (rather than the more personalized self-to-individual *role play*), it seems to be indicative of a personal connection that is productive for visitor's reasoning, particularly about quantitative relationships. Exploration of the utility of competitive talk in mediating dialogue in this context would be an interesting subject for future work.

5.4.3 Blending the Blends: Mixing uses of APT

The previous sections showed three distinct uses of APT: as a tool for *orienting* visitors to the representation, as a tool for connecting visitors' selves to the data from the perspective of an individual living in the dataset that they could *role play*, and as a tool for *projecting* a visitor's identity to the larger group of data they were embodying. Of the 54 groups using APT, 36 (67%) used APT as only one kind of tool: 17 for *orienting* only, 13 for *projection* only, and 6 for *role play* only. Two APT statements were coded as "ambiguous" because the visitors' use of the first-person plural pronoun "we" made it unclear whether they were referring to an individual or group ("We're in different areas," in FBM-S63 and "There we are!" in FBM-S81); these statements both occurred in sessions using only other *projection* APT. Of the 16 remaining groups, 14 used APT two ways, and two used it all three ways. Figure 22 graphs the usage of APT by each group in relation to the content scores of that group's interaction.

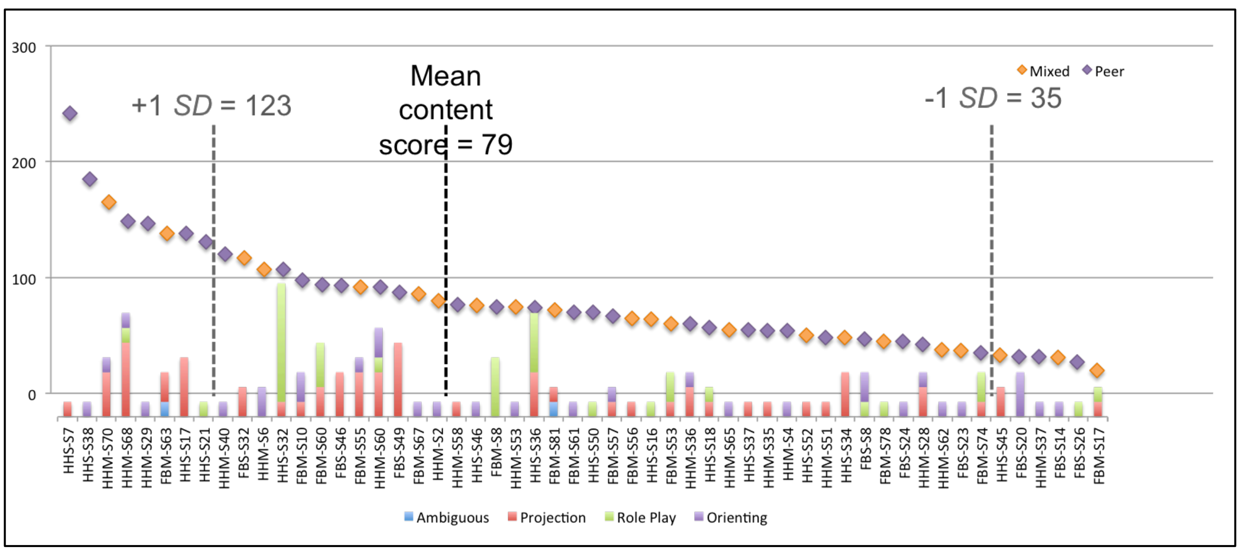


Figure 22. Graph depicting content scores and APT usage by tool type. Bars are color-coded to indicate the number of statements of each visitor-to-data relationship.

We can see from this graph that of the groups whose content scores fell above the mean, most used APT more than once, and of those, most used it a variety of ways. Inspection of the dialogue revealed that these different usages were not necessarily in different sections of the dialogue, which might have indicated visitors were in a mental mode, so to speak, of relating to the whole dataset (*projection*) or relating to an individual experience (*role play*). In fact, the statements identified in section 5.1 as needing to be split after their initial idea unit coding for the APT analysis were all due to mixed role play and projection usages. For example, a participant in the adult peer group FBM-S60 stated:

"Alright, so I'm in Manhattan. So there's a lot of me right there."

The first sentence in this statement is an individual *avatar* usage of APT: I—as this person represented by the data—am located in Manhattan. Immediately in the next sentence, the visitor switches to a group characterization by saying "there's a lot of me," meaning "there are a lot of people like me." This rapid flip from self-to-individual connection to self-to-group connection seemed to come naturally to the visitor producing it and was unquestioned by his companion, who had seemingly no problems understanding. Another group, HHS-S36, showed similar flexibly in using *role play* and *projection* APT. See Table 10.

TABLE 10. TRANSCRIPT OF SESSION HHS-S36, DEMONSTRATING BOTH *ROLE PLAY* AND *PROJECTION* BLENDS

Line	Participants	Data	Blended Role Play (Individual)	Blended Projection (Group)
[1]			M: So where do you work?	
[2]		F: Yeah, teach!		
[3]				M: I'm congregated in Queens.
[4]				F: Yeah, I'm all over the place.
[5]	F: Can we move down?			
[6]			F: I wasn't so popular.	
[7]		M: In the nineties.... spread out.		
[8]			F: I was always important.	
[9]			M: You're, I was going to say, you're always in the city.	
[10]		M: Right?		
[11]		F: Yeah.		
[12]				F: Mostly in the city, but I'm all over the place.
[13]		M: Alright, household type.		
[14]		M: Single detached house. That probably didn't change very much, right?		
[15]		F: There would probably have been more. I assume, yeah.		
[16]	M: We're looking at the same thing. Because we had the			

Line	Participants	Data	Blended Role Play (Individual)	Blended Projection (Group)
	same thing.			
[17]	F: Well, yeah.			
[18]		M: Staten Island has become more... populated.		
[19]	M: You do it.			
[20]		F: I'm... the next are pretty much all the same.		
[21]	M: {We can} look at it.			
[22]		F: More populated.		
[23]		M: It'd be interesting to look, like, household sizes of like 6 increased, compared to household sizes of like two.		
[24]	F: Because I mean we're all similar, so...			
[25]		F: Look at that.		
[26]		F: It's going to decrease.		
[27]		F: No! I thought it would decrease.		
[28]		F: That it?		
[29]		M: They might have to update their data in Brooklyn.		
[30]		M: {Okay} look, Williamsburg has like... see?		
[31]		F: Yeah. Hm.		

Though their session content score was only 74, this adult pair had one of the shortest interaction times (105 seconds) and one of the highest percentages of mid-relevance and high-relevance codings. They quickly jumped in with their flexible APT usage, which may have helped set the stage for the productive talk throughout their brief session. After they engaged in flexible APT usage through line [13], they began making predictions (lines [14-15], [26]), characterizations ([18], [20], [22]), and inferences about other data and geographies ([23],[29]). Figure 23 illustrates their interaction.

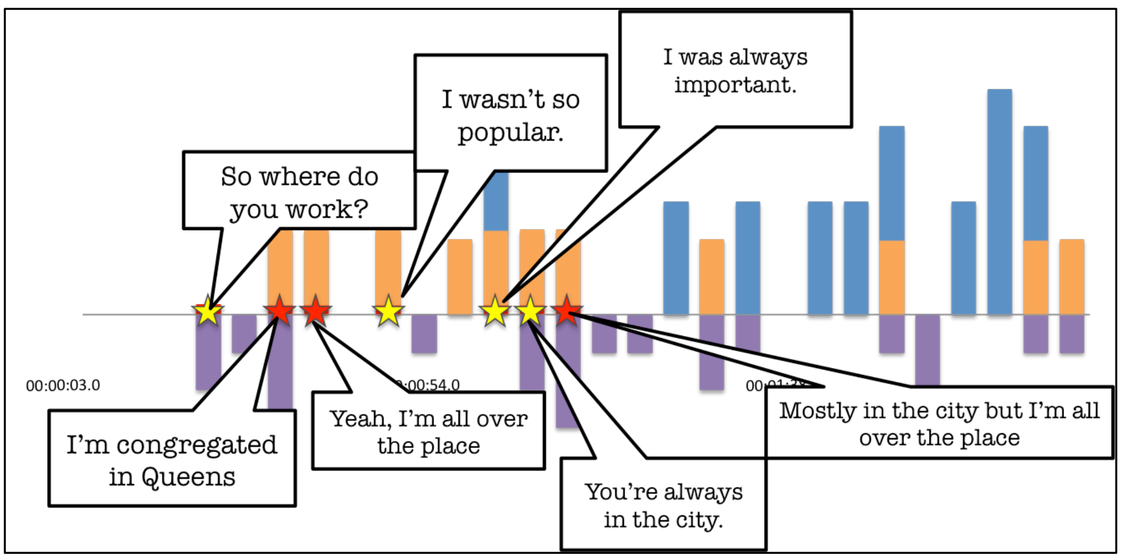


Figure 23. Session timeline of HHS-S36 with APT annotated. Yellow stars indicate *role play* and red indicate *projection* usage.

In summary, the twelve APT applications outlined in section 5.3 filled three broad productive purposes: orienting visitors to the representation, helping them make statements about whole populations by projecting their identity to the whole dataset, and role playing individual stories. These three types of self-data relationships provide unique perspectives to the visitor for making sense of different aspects of the presented data. The flexibility with which peer groups and high performing mixed groups used APT may be an indication of its role as a cognitive tool for thinking about data. While the sample sizes present in this dataset for each function are insufficient to draw statistics-based claims about causal relationships, interesting trends in the numbers of APT functions used by each group and the differences in groups in usage of *role play* and *projection* self-data relationships may warrant attention in future work.

5.5 Relationship between Controller Design and Perspective Taking

While the analyses presented here examined controller design and perspective taking separately, it was expected based on prior work that there might be a relationship between the two. Specifically, the full-body *means of control* and multi-input *distribution of control* were both anticipated to help foster APT through the greater physical and individualized connections to the datasets (Roberts et al., 2013; Roberts et al., 2014). Preliminary analysis of APT usage showed promise for the impact of distribution of control, with a higher percentage of multi-input users engaging in APT at least once in both the full-body (15/28, 39.5%) and handheld conditions (17/29, 58.6%) compared to single-input full-body (9/29, 31.0%) and handheld (14/33, 42.4%).

Statistical analyses were run to determine whether MoC and DoC were predictors of whether APT was used and whether they impacted the total number of APT statements made. To determine whether controller design impacted whether visitors used APT at all, a logistic regression analysis was conducted to predict any usage of APT during the session with MoC and DoC as predictors. A test of the full model over a constant-only model was not significant, indicating that the MoC and DoC predictors as a set did *not* reliably predict APT usage, (X^2= 4.851, p = .088 with df = 2). Naglekerke's R^2 of .053 indicated a weak relationship between prediction and grouping. Prediction success overall was 58.8% (81.5% for no APT and 31.5% for APT). The Wald criterion demonstrated that neither MoC (p = .265) nor DoC (p = .055) made a significant contribution to prediction.

A two-way ANOVA was run to determine whether the *means of control* (MoC) or *distribution of control* (DoC) affected the number of APT instances occurring in a session. No main effect was found for MoC, $F(1,115)$ = .385, p = .536, or DoC, $F(1,115)$ = 1.542, p = .217, and no interaction effect was found between the two, $F(1,115)$ = .417, p = .520. These results indicate that neither the means nor distribution of control were more likely to produce a greater number of APT statements in a session.

In summary, although higher proportions of sessions in the multi-input conditions spontaneously used APT at least once during their sessions, these differences were not significant at the .05 level. No differences were observed between full-body and handheld conditions in APT usage, though as will be discussed in Section 6.1 below, elements of the design of the full-body control version of this exhibit may have limited the affordances of embodiment that may have been instrumental in supporting spontaneous APT usage (i.e. strong physical connection). Further work investigating whether highly embodied controller designs can impact spontaneous use of the first-person actor perspective would be warranted.

5.6 Conclusion

This chapter examined visitors' spontaneous adoption of the Actor perspective as they engaged with each other and the data representation during their sessions. It was found that overall, content scores were higher for groups using *actor* perspective taking (APT) at least once during their session. Deeper examination revealed that the differences between APT and non-APT sessions only manifested in peer groups; mixed age groups had identical mean content scores regardless of the presence of APT. The analysis then focused on functions of APT statements: what were visitors trying to accomplish using APT? Interpretive coding of the 125 APT statements identified twelve unique functions of APT, which then fell broadly into three patterns of self-to-data relationships. Some statements served as *orientation* tools. These statements did not "blend" identities but rather helped visitors translate and connect the representation to themselves. Other APT functions served the purpose of connecting a visitor to an individual living in the dataset in a *role play* relationship. These functions allowed the visitor to imagine "if I were this person with these demographic characteristics living in this map, where would I live?" This type of usage facilitated qualitative humanizing of census data, which can be productive for relating otherwise abstract data to prior knowledge. Other functions of APT connected the visitor to the whole dataset, *projecting* his identity to the whole group. These constructions allowed the visitor to use singular first- and second-person pronouns to refer to an entire group of people, affording quantitative and spatial comparisons and characterizations, sometimes with a particularly engaged competitive slant.

Ochs, Gonzales, and Jacoby (1996) posited the theory that use of blended identities, as marked by "indeterminate constructions" of grammatical talk—where the physicists being studied similarly used personal pronouns to refer to external objects—was not only common, but was most associated with expert thinking (i.e. most frequently used by the head of the lab). It was found to be tightly linked to interpretation of a data representation and was used fluidly and flexibly, with scientists moving in and out of different constructions unproblematically. The data presented here demonstrate that even non-experts not only are able to blend their own identities with abstract data as represented by a graphical display, but that they do so in multiple ways. This study builds on the findings of Ochs and colleagues in the following ways: (1) APT is not limited to experts, but may be an indication of data comprehension or literacy. In particular peer groups who did not use APT had session content scores significantly below the mean. This may be a sign that APT is a natural way of talking about data with which one feels comfortable. For example, the adult pair in HHS-S32 (section 5.4.1) interacted with the exhibit for almost 2 minutes before they really understood what to do with it. Eventually one participant said, "So that's where we'd live" — using *avatar* APT—to which his companion responded "Oooohhhh," finally understanding the exhibit. Prior to that point their talk comprised almost entirely low-relevance comments. Once they understood the exhibit they started using APT and engaging in productive learning talk. (2) This analysis identified three broad functional uses of APT. Two of these uses are subdivisions of "blended identities" as described by Ochs and colleagues in which the speaker blends subject and object to make sense of a phenomenon. Both of the identified blends—*role play* and *projection*—connect a visitor to the dataset in a way that is productive to the kinds of learning talk the exhibit aims to promote, and the flexible usage of these blends, particularly by high-performing peer groups, suggests APT is not a single-purpose cognitive tool but rather is a multi-tool capable of mediating dialogue and meaning making in multiple ways. Implications of these findings for future work are discussed in the final chapter.

6 DISCUSSION AND IMPLICATIONS FOR FUTURE WORK

This dissertation presented two analyses of visitors' dialogue at an interactive data museum exhibit. Chapter 4 evaluated effects of the *means of control* and *distribution of control*—and the differences in exhibit interactions caused by these variations—on visitors' productive learning talk during their interaction sessions. Chapter 5 examined visitors' use of *actor perspective taking* (APT) to understand the ways it mediated their dialogue. This final chapter discusses implications for these findings and for the methods used to discover them on future work.

6.1 The Curious Case of Handheld Single Input

Theories of embodied learning suggested that higher levels of physical involvement would productively mediate interactions. In the study presented here, this was not the case. Visitors in the full-body controller conditions produced significantly less learning talk as measured by content scores than those in the handheld conditions, whether in the single-input or multi-input variations. Section 4.8 attempted to account for these differences by analyzing differences in how people interacted with the exhibit in the conditions, including the amount of data rendered, the number of control actions engaged in, and their difficulties in learning the operation of the system (as measured by the amount of dialogue dedicated to learning the interactive system and evidenced within the learning curve for understanding the interactivity). None of these factors were fully able to explain the results. This section looks at additional factors that may have been at play in this scenario but that are not testable in the current dataset, and discusses implications for future work.

6.1.1 Intuitiveness of Control Gestures

The CoCensus gesture suites for the full-body and handheld conditions were designed to be parallel to each other in an attempt to isolate the size of the gesture—the amount of the body involved in the action—as the independent variable. However, this may have led to an unintended confound where full-body controls were less intuitive than their handheld counterparts. It is well documented that off-the-desktop interactions are hardly "natural" (Norman, 2010; Cafaro, 2015). In the current era of technology immersion even very young children are comfortable pinching and swiping touch screens and tablets, but though game systems like the Wii and Xbox are popularizing full-body control games, they are still novel. Kaptelinin (2013) reviews the role of modern technologies and cites Ihde (1990), who describes "hermeneutic" relationships with technology. In these relationships we become aware of the technology because we have the need to interpret both it and the world we are acting upon through it. This could be how the full-body conditions affected people: requiring them to interpret the controls as they attempt to interpret the data. While not a traditional learning curve, in the sense of visitors being *unable* to operate the controls for the exhibit, the extra work required for this joint interpretation could well have been enough to deflect attention from data interpretation tasks to interaction interpretation, resulting in less learning dialogue. Indeed, the analysis presented here examined visitors' learning curve over the course of their session by looking at dialogue content scores in the beginning, middle, and end of the interaction (section 4.8.4) and found a difference between full-body and handheld users during the middle portion of the sessions. The handheld users generated more low-relevance codes during the middle section of use, perhaps indicating a greater fluency with the controls. The relatively short durations of these sessions may not have been enough for visitors in the full-body condition to become comfortable enough to reach an "embodiment" relationship with the technology, in which they act through the technology without being aware of it (Kaptelinin, 2013, citing Idhe, 1990). Extended interactions may have led to smaller differences between conditions.

Cafaro (2015) studied the role of *embodied schemata* in shaping how visitors can intuit an understanding of a suite of control gestures. Through "framed guessability" studies with participants using *CoCensus* in a lab setting and *in situ* at the museum, he found that complementary suites of gestures and body movements can be more easily discovered by participants when a certain *frame* is activated for the visitors, such as that of working out with a weighted ball or standing in front of a funhouse mirror. Participants in the handheld conditions of the study presented here may have—by virtue of being handed a familiar tablet device—inadvertently been given a "frame" for how to interact with the system. They were able to employ a suite of familiar gestures commonly used with touchscreen devices: swiping, pinching, and button pressing. Perhaps more importantly, these gestures mapped clearly to familiar outcomes: swiping changed the "picture" of the map to another category just as swiping photos in a camera's gallery application changes to a new picture, pinching the tablet controller was somewhat akin to a zoom that would be expected in a photo or map application (and in fact many visitors referred to aggregation changes as "zooming"), and tapping a decade button selected that decade as one would expect in any menu. No comparable frame exists for the full-body equivalents of these gestures. While none of the full-body control actions are inherently counter-intuitive or idiosyncratic, and they were all gestures suggested by visitors in early guessability studies (Cafaro, 2015), they do not complement each other or feel like part of a unified suite of actions that are a part of normal life like those on a touchscreen.

Because of this inherent framing, the handheld-controlled systems could be viewed as a form of higher-order or embedded mediation (Kaptelinin, 2013). Manipulating the data represented on the display requires manipulating the controller. This dissertation took the perspective that the controller, whether it was a tablet external to the user or the user's body itself, was a mediational means in the interaction, a kind of *physical* (Wertsch, 1998) or *technical* (Vygotsky, 1978) tool. From the perspective of embodied cognition, removing the physical controller and placing the locus of control in the learner's body fosters a stronger connection between user and system. However, the research may have overlooked that the physical tool of the handheld tablet controller brought with it an embedded *psychological* tool, the frame for understanding and relating the control actions to each other. Because the full-body condition had no such frame uniting the gestures, removal of the physical device effectively may have removed an additional mediational means, resulting in a disconnect between the control actions and the exhibit's response.

Furthermore, the control actions in this exhibit activated discrete, not continuous, responses from the system. Completing any of the three actions triggered a change, but until the event was triggered the user received no feedback from small movements, potentially decreasing their ability to develop fluency of control (Snibbe & Raffle, 2009; Norman, 2010). Early iterations of *CoCensus* used distance from the display to control the transparency of an individual's data bubbles (Roberts et al., 2012; Roberts et al., 2013), with a movement front-to-back in the interaction area triggering a transparency change matched to the extent of the ground covered in the movement (fully opaque at the front of the room and nearly fully transparent at the back). The initial idea behind this component of the interactivity was that altering the transparency would allow visitors to choose whether they wanted to focus on the data itself or the underlying geography, by making one or the other more visible, so they could reason about both quantity and spatiality. We now suspect that this design decision had the side effect of fostering stronger physical and conceptual links between the person and the data. This connection was supplemented with a small jiggle to the data bubbles when a user moved to provide additional visual feedback, with faster movements triggering larger jiggles. These interaction components were abandoned following findings that the front-to-back timeline configuration was more intuitive and enjoyable for users (Roberts et al., 2014), so the vertical axis (perpendicular to the display) became the timeline controller. I suspect, though, that reincorporating an interaction design utilizing continuous feedback has high potential to boost the affordances of full-body interaction and would be of great interest in future work.

Finally, it must be noted that two of the three control actions in the full-body conditions were touchless. Swipes were performed in the air, and while some visitors finished the aggregation gesture of drawing their arms together in an exaggerated pinch with a clap, this clap was not a necessary part of the control action. The tendency of visitors to do this clap may speak to the desire for some sort of haptic feedback for these control actions. The field of tangible, embedded, and embodied interactions (TEI) is based on the idea of building systems that allow humans and computers to interact with fewer intermediaries, using “increased physical engagement and direct interaction”(Hornecker, 2011). While the full-body design studied here did have fewer intermediaries (i.e. the removal of the external controller), it could be argued that the increased size of the control actions did not actually lead to “increased physical engagement.” After all, what is more physical than touch? Chattopadhyay & Bolchini (2013) examined touchless interfaces in wall-sized displays and emphasized the importance of visual feedback in absence of haptic feedback. The observed differences between full-body and handheld conditions in learning talk suggest that the full-body version of *CoCensus* as tested here may not have adequately accommodated for this lack of touch.

6.1.2 Challenging the notion of "heads down" technologies

As technology has been increasingly incorporated into museum exhibitions, designers have worried about the impacts of use of digital portable screens on the social nature of a museum exhibit. Hsi (2003) reported such "heads down" activity with visitors using the Electronic Guidebook system, in which visitors not only were observed dedicating more attention to the screen than to their companions and the exhibits in the museum, but they also self-reported feelings of isolation. Similar findings were reported by Bellotti, Berta, Gloria, & Margarone (2002) using a mobile application in an aquarium. Heath and vom Lehn (2008) reviewed screen-based interactives at multiple institutions, reporting that visitors tended to passively "wait their turn" to use a screen-based exhibit, leading to isolation of visitors and decreased social interactions. Lyons (2009) confirmed that complex handheld interfaces led to more frequent and longer gazes at the device rather than at the museum exhibit it was mediating. However, she noted that visitors' conversations and collaborative activities were not adversely affected (Lyons, 2009), concluding that "the heads-down phenomenon is not as problematic for multi-user activities as it has been found to be for single-user activities in museums." The device tested by Lyons was highly interactive and explicitly designed for social use, much like the tablet devices tested here. The tablets in this dissertation study solely served the purpose of controlling the large, shared display and did not convey any content themselves, making them less likely to draw visitors' attention from the shared experience.

All these earlier studies were conducted prior to the wide release of smartphones and touchscreen technologies. According to a report by the Pew Research Center (2015), while the majority of U.S. adults already owned cell phones and desktop or laptop computers in the 2000s, the iPhone wasn't introduced until 2007, and 2010 was the year the iPad tablet computer was introduced to the market. In 2010, an estimated 35% of U.S. adults owned a smartphone; that number has now risen to 68%, and tablet ownership since 2010 has climbed to 45% of U.S. adults (Pew Research Center, 2015). This ubiquity of tablets and smart phones presents the possibility that touchscreen devices are not only *not* novel, they are increasingly used as mediators of conversation: picture a group of friends around a table, talking and passing phones around to illustrate points, share videos and photos, or look up just-in-time information relevant to the conversation. In designing this interactive exhibit, it was thought that a full-body interaction design would help keep people's "heads up" and their attention on the shared social space. In light of the unexpected success of the handheld conditions, it may be time to re-evaluate the concerns designers and practitioners have for heads-down visitors to take into account the fluidity and flexibility with which people mediate their social engagements through technological devices.

6.1.3 Moving Forward: Implications for future work

This study has shown that while full-body interaction may have its benefits, it is susceptible to problems, in particular that the connection between the user and the output is not necessarily easily made. Based on the results here, I propose that because of our current cultural uses of technologies (based in touch and haptic feedback), full-body interactions don't have the inherent frame for helping people use them without an extra level of interpretation work. These frames for assisting people in interpreting the technology can be supplied to learners (Cafaro, 2015), but they won't yet happen naturally until touchless technology interactions are fluid and ubiquitous. Until then, designers must take care to attend to the psychological tools (e.g. frames) embedded within physical tools. That is not to say work in the meantime cannot address questions of the role embodied cognition can play in facilitating learning in this kind of environment. For example, a repetition of this study using large physical controllers to complete the control actions (e.g. a large wheel or slider or flip panel to change the categories) would better isolate the independent variable of gesture size while removing the touch/touchless confound that seems to have emerged in this study.

In addition, multiple measures that may have indicated benefits of the full-body controllers were not taken here, such as visitors' enjoyment of the system. This metric could be quantified through a survey or studied qualitatively through exclamations and other affective responses (Allen, 2002), and may show a preference of museum visitors for more physical interactions. One mother in the HHS condition, for example, said sarcastically as her son was handed the tablet controller, "Wow, I'm so glad we came to the museum today to get you away from your iPad," and some participants in the full-body conditions—particularly younger users—clearly enjoyed jumping around the timeline. Measures of individual cognitive gain could also demonstrate affordances of full-body interaction on cognition that may not have manifested in dialogue. Though not aligned with the sociocultural framework adopted in this study, a cognitive assessment could provide more insight to how movement may have impacted thinking. Similarly, this analysis only gave cursory attention to the balance of dialogue in sessions but otherwise did not look at individual interactions to measure "participation equity" (Kapur & Kinzer, 2007) or "task division" (Lyons, 2009). An analysis narrowing the focus to individuals' contributions within sessions, particularly in relation to their movements and gestures, could deepen our knowledge of how movement mediated individuals' experiences.

6.2 Actor Perspective Taking

Initial findings that APT users had higher average content scores were incredibly promising, but identifying mechanisms for that relationship was difficult and left many questions unanswered. In particular, the stark contrasts between peer groups who used APT (with a mean content score of 83) and those who didn't (M = 50) appeared to be fruitful at first glance, but the data and analysis presented here were not able to develop a theory as to the underlying causes.

Much work studying the conversational strategies used by museum visitors—particularly family groups—has established that members of the group assume different identities and roles (Falk, Moussouri, & Coulson, 1998; Crowley & Jacobs, 2002; Allen & Gutwill, 2009; Atkins et al., 2009; Ash, 2003; Borun, Chambers & Cleghorn, 1996; Falk, 2006; Schauble et al., 2002). This dissertation laid a framework for understanding a role not previously explored in museum literature, actor perspective taking, by characterizing how APT is used in spontaneous visitor conversations and identifying twelve distinct applications of APT in this context. Analysis of these applications, and the broader self-to-data relationships of role playing and projection that they revealed, could be a fruitful direction for future work in understanding visitors' roles. The framework presented here for both the APT application codes and the two human-data relationships they help foster—self-to-individual *role play* and self-to-group *projection*—can be useful for exploring APT use not just in an expanded study in the current context—from which statistical causal claims could potentially be made—but in other data interpretation contexts as well.

This work built heavily on the findings of Ochs, Gonzales, and Jacoby (1996) in their examination of physicists' use of "indeterminate constructions," in which they noted that "the construction of this indeterminate referential identity plays an important role in scientists' efforts to achieve mutual understanding and arrive at a working consensus." The data presented here indicate that this linguistic construction may be beneficial to non-experts in collaboratively interpreting data, too. An investigation into the ways these applications are utilized in other settings by experts and non-experts could serve to inform the design of future learning environments.

Ochs et al. note the importance of studying these indeterminate constructions as "part of embodied interpretive practices," saying "it is therefore crucial to pursue an analysis that integrates the language, gestural practices, and visual arrays which comprise physicists' interpretive activities." The analysis presented here did not investigate gestural interactions among participants and how they positioned themselves physically in relation to each other and the map display. Future work attending to these interactions would be of interest in further investigating parallels and differences between experts and non-experts. Moreover, one of the theorized benefits of full-body controls in shared spaces is that users are better able to witness the expressivity of their companion's motions – it may be the case that, if a more continuous control design is used (per section 6.1.1), the controls could serve the dual purpose of interacting with the system and communicating nuance to companions.

6.3 Methodological Implications

Understanding visitors' dialogue as they interact naturalistically with museum exhibits is of great interest to museum researchers but has traditionally been difficult to capture. Much prior work on museum learning relied on overheard talk from live observations or transcripts of audio recordings taken from microphones planted in the space or worn by participants. This dissertation employed a methodological approach that maintained the context of the dialogue by coding all talk directly from the video recordings of the interaction. This tactic not only illuminated the referents in visitors' conversations but also allowed segmentation of dialogue into a meaningful unit for the spontaneous, flowing discourse occurring in museums: the "idea unit." Because idea units as defined here, modified from the broader definition used by Jacobs et al. (1997), were not restricted to continuous talk—they were often interrupted by other participants or the original speaker's own separate thought—and because they could flow across multiple users as visitors co-constructed the talk, they most accurately reflected the nature of the dialogue occurring. Segmenting dialogue this way addresses a problem addressed by Allen (2002) in dealing with visitor discourse that tends to be "fragmented, ambiguous, or lacking clear referents" and that frequently involves repetition of words and phrases as members of a group echo each other. Allen dealt with this issue by coding only for the presence or absence of a type of talk during the entire interaction. While this strategy provides some information about the learning talk, it doesn't allow for the more nuanced understanding of the depth of conversation. Breaking the discourse into idea units that are then coded individually for presence of absence of a type of talk allows a clearer picture of the content of the dialogue to emerge. A paired analysis (using idea units and more traditional segmentation) on the same set of data could illuminate the power of idea unit coding for understanding learning talk at exhibits, and would likely be of interest to the visitor study community.

Deeper analysis of these idea units throughout the visitors' interactions could examine more closely their trajectory over the course of the session to understand how the conversations build and what ideas are returned to throughout the course of a visit. For this analysis, whose aim was to assess differences among conditions with respect to learning talk, a single quantitative measure was all that was needed. This was accomplished through the combination of two methods. By using *simultaneous coding* (Saldaña, 2009) in which a single idea unit could be coded with any number of codes, this analysis was able to capture not only the content of each statement but also the depth and complexity of the talk. Because each code applied was, though productive, not equally relevant to the learning goals of the exhibit, this analysis also employed *magnitude coding* (Saldaña, 2009; Miles & Huberman, 1994; Tashakkori & Teddie, 2010) in order to differentiate codes according to their relevance to the exhibit's learning goals, much the way a teacher uses a rubric to quantify a student's piece of creative writing. By assigning weighted values to the codes applied, differences in the dialogue generated in each session were quantifiable and available for statistical comparisons. This method cannot detect all differences between groups. In particular, groups with less overall talk but all or most of it at in the "highly relevant" category may have similar or lower content scores than other groups who talked more but mostly made mid and low-relevance statements and rarely got to the "highly relevant" talk. In a formal classroom setting, this distinction might be crucial for assessing learning. However, in the museum setting where the overarching learning goal is to help visitors engage with and talk about content in a way that is meaningful and relevant to them and their companions (Falk & Dierking, 2000; Ash, 2003; Packer & Ballantyne, 2005; Allen, 2002; Atkins et al., 2009), this method most directly measures that goal. As museums embrace the dialogic model of education, they must concurrently embrace research methods suited to that model. The methodology presented here takes steps toward that goal.

6.4 Conclusions

This dissertation addressed multiple facets of an emerging problem space for informal designers and researchers. The increased social access and physical engagement of "off the desktop" interactive technologies make full-body, multi-input systems appealing to museum exhibit designers and educators hoping to actively engage their visitors in interpretation of content. Theories of embodied cognition suggest that increased physical engagement may productively mediate visitors' learning. However, this dissertation has called that assumption into question. When conducting design-based research, outcomes that do not match theoretical predictions provide opportunities for reflection, inviting a re-examination of both the theory and the implementation. In this case I have identified several aspects of the design which may have fallen short of allowing this exhibit to fully support embodied cognition as theory would require. These new hypotheses add nuance to the discussion concerning the design of effective full-body educational experiences: (1) It may be the case that continuous interaction controls are critical to realizing the benefits of embodiment, and that without such fluid controls visitors never fully embrace using their bodies as tools to control the experience. (2) It may be the case that a lack of an interpretive frame can stand in the way of users' comfort with fluently using haptic-free gesture-only controls. Because people are used to acting on or acting with objects, touchless controls fundamentally lack the framing users need to seamlessly adopt their use. These are valuable human-computer interaction issues to raise, but the implication for educators is that in the absence of embodied controls that learners can fully adopt and inhabit, learners are not able to reinforce or even offload part of their cognition onto their physical selves, one of the advantages for learning that embodied cognition theory offers to designers.

While this research generated more questions than it did answers about the use of embodied controls in educational settings, it did provide an opportunity to generate a framework for documenting learners' interactions with data outside of formal classroom settings. As interest grows for museums to be able to present data as the product of science to their visitors, there is a need to understand how the body of work examining learning through graphic displays—overwhelmingly situated in lab or classroom settings—can be adapted to informal free-choice learning environments such as museums. This dissertation presented a framework for analyzing dialogue that accommodates for the free-choice, fluid nature of talk in social museum spaces while attending to the aspects of graph interpretation known to be productive.

Finally this dissertation investigated the psychological tool of actor perspective taking (APT). Though this phenomenon has been investigated across multiple contexts and disciplines, its application in spontaneous dialogue has been under-studied. This dissertation has shown that APT is used unproblematically and flexibly by learners interpreting data and may be a marker for understanding. Further efforts to capitalize on the affordances of APT and the potential for interactive technologies to impact visitors' adoption of an actor perspective has great potential to improve exhibit design going forward.

The dialogue among visitors, the perspectives they employ, the exhibit content, and the design of the both physical space and an exhibit's interactive features are all *mediational means* (Wertsch, 1998) contributing to the exhibit interaction experience, each affecting the learning taking place. The work presented here took initial steps to understand how these mediational means work together—or in conflict—to affect learning outcomes. More of this work is needed to better understand the rich, multi-layered learning experiences provided in technology-mediated informal settings.

7 APPENDIX A: VISITOR TRACKING SHEET

Handheld Single Input MASTER COPY Visitor Tracking Sheet: Jessica's Study (rev 5/15/15)

Consent?	Date & Time	Users	Sex	Age	Shirt color	Relationship	Tech notes	Initials
Y N		Kiosk A	M F	C T A				
If yes, Session ID:		Kiosk B	M F	C T A				

Consent?	Date & Time	Users	Sex	Age	Shirt color	Relationship	Tech notes	Initials
Y N		Kiosk A	M F	C T A				
If yes, Session ID:		Kiosk B	M F	C T A				

Consent?	Date & Time	Users	Sex	Age	Shirt color	Relationship	Tech notes	Initials
Y N		Kiosk A	M F	C T A				
If yes, Session ID:		Kiosk B	M F	C T A				

Consent?	Date & Time	Users	Sex	Age	Shirt color	Relationship	Tech notes	Initials
Y N		Kiosk A	M F	C T A				
If yes, Session ID:		Kiosk B	M F	C T A				

Consent?	Date & Time	Users	Sex	Age	Shirt color	Relationship	Tech notes	Initials
Y N		Kiosk A	M F	C T A				
If yes, Session ID:		Kiosk B	M F	C T A				

Record all users;
Assign consecutive session ID for all consenting.

Tech notes: x in box if yes, write full description
on back of sheet (reference Session ID#)

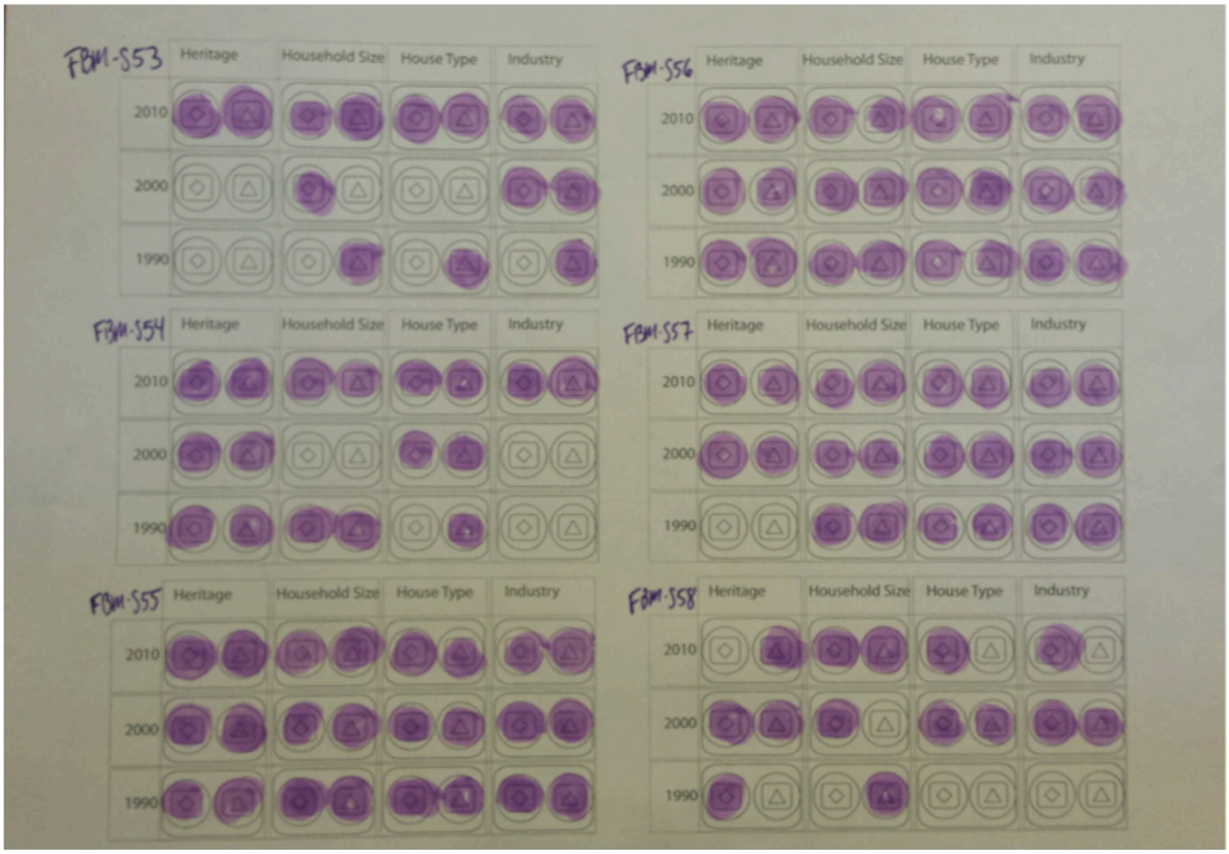
Age: C = child; T = teen; A = adult

8 APPENDIX B: LEARNING TALK CODING DICTIONARY

Code	Description
INSTANTIATE category	Heritage, Household Size, House Type, or Industry
INSTANTIATE dataset	The selected variable, e.g. Irish, White, English, Four (person household), construction, apartment
INSTANTIATE decade	1990, 2000, 2010
INSTANTIATE geography	e.g. Manhattan, New York, Prospect Park
INSTANTIATE outside knowledge	A fact about the geography or dataset that is said aloud for use in meaning making
INSTANTIATE representation	Bubbles, blue, red, circles
INSTANTIATE self	I, me, you
MANAGE ask interpretive question	Ask what the exhibit is or does or what the data are showing
MANAGE purpose of exhibit	Tell companion what the exhibit is or does or what the data are showing
MANAGE clarify	Give more information about a dataset, the census, or the representation to clarify it to companion
MANAGE direct co-visitor's movements	Tell a companion to do a particular control action, e.g. "Step forward," or "Go to 1990."
MANAGE suggest action	Suggest a control action, e.g. "Should we go to the next one?" Distinct from direct movements in that it involves the companion in the decision, instead of giving a command.
MANAGE narrate intentionality	Say aloud what one is going to do, "I'm going to look at 1990." Similar to suggest action but isn't asking for input from companion
MANAGE negotiation of control	Coordinate movements or exhibit control with companions. Similar to suggest action but involving negotiation when visitors have different movements in mind and have to agree.
MANAGE direct co-visitor's attention	Point out specific item of interest to companion. Must give direct referent, e.g. "Look at the big cluster in Queens." Ambiguous statements, e.g. "Look at that!" are not coded unless "that" is specified.
EVALUATE characterize	Qualify the data in some way, using words like "so many" or "a lot" or "spread out"

Code	Description
EVALUATE question census categories	Point out that the interpretation of the data is dependent on the definition of the category by the census, e.g. "Well it depends on what they mean by White."
EVALUATE win	Qualify a comparison or characterization by adding a competetive aspect, e.g. "I'm winning," "You dominate," "I'm overpowering you."
INTEGRATE connect simple	Make a connection across two *types* of data, e.g. "This is construction in Manhattan."
INTEGRATE connect multiple	Make a connection across more than two *types* of data, e.g. "This is construction in Manhattan in 1990."
INTEGRATE compare	Note the similarity or dissimilarity between two of the same type of data, i.e. between two datasets, across decades, across geographies. Can be explicit, e.g. "There are more Mexicans than Puerto Ricans." or implicit, "There's still more."
INTEGRATE challenge interpretation	Respond to a companion's statement by giving a conflicting interpretation. E.g. A: It increased. B: No it didn't, it decreased.
GENERATE negotiate meaning	Similar to challenge interpretation, but support argument by incorporating outside knowledge
GENERATE contextualize	Use outside knowledge to make sense of data, e.g. "Do you remember when we were in Manhattan and we saw all those tall buildings that we said were banks?"
GENERATE confirm	Indicate that data matches outside knowledge/expectations
GENERATE notice surprising pattern	Indicate that data doesn't match outside knowledge/expectations
GENERATE identify knowledge gap	Indicate that outside knowledge would be necessary to make a claim or contextualize the data but that the visitors don't have that knowledge
GENERATE make prediction	State what patterns or differences are expected. Can be testable in the exhibit (e.g. "I bet if we go back to 1990 it will be less." or not testable, "In the next census that's going to go down."
GENERATE pose inference	Make a claim about the underlying causes for data patterns visible in the map.

9 APPENDIX C: DATA RENDERING CODING SHEET



CITED LITERATURE

Abrahamson, D., & Lindgren, R. (2014). Embodiment and embodied design. In R. K. Sawyer (Ed.), The Cambridge handbook of the learning sciences (2nd ed.) (pp. 358-376). Cambridge, UK: Cambridge University Press.

Allen, S. (2002). Looking for learning in visitor talk: A methodological exploration. Learning conversations in museums, 259-303.

Allen, S., & Gutwill, J. (2004). Designing with multiple interactives: Five common pitfalls. Curator: The Museum Journal, 47(2), 199-212.

Allen, S., & Gutwill, J. P. (2009). Creating a program to deepen family inquiry at interactive science exhibits. Curator, 52(3), 289–306.

Atkins, L. J., Velez, L., Goudy, D., & Dunbar, K. N. (2009). The unintended effects of interactive objects and labels in the science museum. Science Education, 93(1), 161-184.

Ash, D. (2003). Dialogic inquiry in life science conversations of family groups in a museum. Journal of Research in Science Teaching, 40(2), 138-162.

Bao, P., & Gergle, D. (2009, April). What's this you say?: the use of local references on distant displays. In Proceedings of the 27th international conference on Human factors in computing systems (pp. 1029-1032). ACM.

Beheshti, E., Van Devender, A., & Horn, M. (2012, November). Touch, click, navigate: comparing tabletop and desktop interaction for map navigation tasks. In Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces (pp. 205-214). ACM.

Bell, B., Bareiss, R., & Beckwith, R. (1993). Sickle Cell Counselor: A Prototype Goal-Based Scenario for Instruction in a Museum Environment. Journal of the Learning Sciences, 3(4), 347-386.

Bellotti, F., Berta, R., Gloria, A., & Margarone, M. (2002). User testing a hypermedia tour guide. Pervasive Computing, 1(2), 33-41.

Bernard, H. R. (2006). Research methods in anthropology: Qualitative and quantitative approaches. Walnut Creek, CA:AltaMira Press

Block, F., Horn, M. S., Phillips, B. C., Diamond, J., Evans, E., & Shen, C. (2012). The DeepTree Exhibit: Visualizing the tree of life to facilitate informal learning. IEEE Transactions on Visualization and Computer Graphics, 18(12), 2789-2798.

Borun, M., Chambers, M., & Cleghorn, A. (1996). Families are learning in science museums. Curator, 39(2), 262-270.

Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives pronouns modulate perspective taking during narrative comprehension. Psychological Science, 20(1), 27-32.

Cafaro, F. (2015). Using Framed Guessability to Design Gesture Suites for Embodied Interaction. Retrieved from indigo.uic.edu

Cafaro, F., Panella, A., Lyons, L., Roberts, J., & Radinsky, J. (2013) I See You There! Developing Identity-Preserving Embodied Interaction for Museum Exhibits. Accepted for publication in the Proceedings of the 2013 ACM int'l conference on Human Factors in Computing Systems. ACM, New York, NY, USA.

Card S, Mackinlay JD, Shneiderman B (eds) (1999) Readings in information visualization: Using vision to think. Morgan Kauffman Publishers, Inc., San Francisco, CA

Chattopadhyay, D., & Bolchini, D. (2013). Laid-back, touchless collaboration around wall-size displays: visual feedback and affordances. CHI'13 POWERWALL workshop.

Chi, M. T. H. (1997). Quantifying Qualitative Analyses of Verbal Data: A Practical Guide. Journal of the Learning Sciences, 6(3), 271–315.

Corbin, J. M. & Strauss, A. L. (2008). Basics of qualitative research : Techniques and procedures for developing grounded theory. Los Angeles, Calif.: Sage Publications.

Correia, N., Mota, T., Nóbrega, R., Silva, L., & Almeida, A. (2010). A multi-touch tabletop for robust multimedia interaction in museums. In A. Krüger, J. Schöning, D. Wigdor, & M. Haller (Eds.), ACM International Conference on Interactive Tabletops and Surfaces (Vol. 151, pp. 117-120). ACM. doi:10.1145/1936652.1936674.

Crowley, K., & Jacobs, M. (2002). Building islands of expertise in everyday family activity. In G. Leinhardt, K. Crowley & K. Knutson (Eds.), Learning conversations in museums. Mahwah, NJ: Lawrence Erlbaum.

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. Journal for Research in Mathematics Education, 18, 382–393.

Davis, P., Horn, M. S., Schrementi, L., Block, F., Phillips, B., Evans, E. M., ... & Shen, C. (2013). Going deep: Supporting collaborative exploration of evolution in natural history museums. In Proceedings of 10th International Conference on Computer Supported Collaborative Learning.

Devine, S., and Bernstein, S. (2015) Responsive in Planning and Practice: Comment Kiosks at the Brooklyn Museum. Exhibitionist, Spring 2015, 42-45. Retrieved from http://name-aam.org/uploads/downloadables/EXH.spr_15/11%20Devine_Bernstein%20EXH%20Spring%202015-11.pdf

DiPaola, S., & Akai, C. (2006, July). Designing an adaptive multimedia interactive to support shared learning experiences. In ACM SIGGRAPH 2006 Educators program (p. 14). ACM.

Diamond, J., Luke, J. J., & Uttal, D. H. (2009). Practical Evaluation Guide: Tools for Museums & Other Informal Educational Settings (2 ed.). Lanham, MD: AltaMira Press.

Dourish, P. (2001). Where the action is: the foundations of embodied interaction. MIT press.

Eberbach, C., & Crowley, K. (2005). From living to virtual: Learning from museum objects. Curator: The museum journal, 48(3), 317-338.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. Journal of the Learning Sciences, 8(3-4), 391-450.

Enyedy, N., Danish, J. A., DeLiema, D. (2013). Constructing and Deconstructing Materially-Anchored Conceptual Blends in an Augmented Reality Collaborative Learning Environment (full paper). In S. Puntambekar, N. Rummel, M. Kapur M. Nathan (Eds.), Proceedings of the 10th International Conference on Computer Supported Collaborative Learning. Madison, WI.

Falcão, T. P., & Price, S. (2009). What have you done! The role of 'interference' in tangible environments for supporting collaborative learning. In Proceedings of the Conference on Computer Supported Collaborative Learning. CSCL'09. ISLS, 324-334.

Falk, J. H. (2006). An identity-centered approach to understanding museum learning. Curator: The museum journal, 49(2), 151-166.

Falk, J. H., & Dierking, L. D. (2000). Learning from museums: Visitor experiences and the making of meaning. Altamira Press.

Falk, J. H., Moussouri, T., & Coulson, D. (1998). The effect of visitors 'agendas on museum learning. Curator: The Museum Journal, 41(2), 107-120.

Few, S. (2009). Now you see it: simple visualization techniques for quantitative analysis. Analytics Press.

Filipi, A., & Wales, R. (2004). Perspective-taking and perspective-shifting as socially situated and collaborative actions. Journal of pragmatics, 36(10), 1851-1884.

Fischer, M. H., Dewulf, N., & Hill, R. L. (2005). Designing bar graphs: Orientation matters. Applied Cognitive Psychology, 19(7), 953-962.

Forlines, C., & Shen, C. (2005, October). DTLens: multi-user tabletop spatial data exploration. In Proceedings of the 18th annual ACM symposium on User interface software and technology (pp. 119-122). ACM.

Forlines, C., Esenther, A., Shen, C., Wigdor, D., & Ryall, K. (2006, October). Multi-user, multi-display interaction with a single-user, single-display geospatial application. In Proceedings of the 19th annual ACM symposium on User interface software and technology (pp. 273-276). ACM.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications.Journal for Research in mathematics Education, 32(2), 124-158.

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. Studies in science education, 47(2), 183-210.

Gutwill, J. P. (2003). Gaining visitor consent for research II: Improving the posted‐sign method. Curator: The Museum Journal, 46(2), 228-235.

Gutwill, J. P., & Allen, S. (2012). Deepening students' scientific inquiry skills during a science museum field trip. Journal of the Learning Sciences, 21(1), 130-181.

Heath, C., & vom Lehn, D. (2008). Configuring 'Interactivity' Enhancing Engagement in Science Centres and Museums. Social Studies of Science,38(1), 63-91.

Horn, M., Atrash Leong, Z., Block, F., Diamond, J., Evans, E. M., Phillips, B., & Shen, C. (2012, May). Of BATs and APEs: an interactive tabletop game for natural history museums. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2059-2068). ACM.

Hornecker, E. (2008, October). “I don’t understand it either, but it is cool”-visitor interactions with a multi-touch table in a museum. In Horizontal Interactive Human Computer Systems, 2008. TABLETOP 2008. 3rd IEEE International Workshop on (pp. 113-120). IEEE.

Hornecker, E. (2010, January). Interactions around a contextually embedded system. In Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction (pp. 169-176). ACM.

Hornecker, E. (2011). The role of physicality in tangible and embodied interactions. interactions, 18(2), 19-23.

Hsi, S. (2003). A study of user experiences mediated by nomadic web content in a museum. Journal of Computer Assisted Learning, 19(3), 308-319.

Ihde, D. (1990). Technology and the lifeworld: From garden to earth (No. 560). Indiana University Press.

Jacobs, J. K., Yoshida, M., Stigler, J. W., & Fernandez, C. (1997). Japanese and American teachers' evaluations of mathematics lessons: A new technique for exploring beliefs. The Journal of Mathematical Behavior, 16(1), 7-24

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. Journal of Educational Psychology, 106(1), 86.

Kaptelinin, V. (2013). The mediational perspective on digital technology: Understanding the interplay between technology, mind and action. The SAGE handbook of digital technology research, 203-213.

Kapur, M., & Kinzer, C. K. (2007). Examining the effect of problem type in a synchronous computer-supported collaborative learning (CSCL) environment. Educational Technology Research and Development, 55(5), 439–459. http://doi.org/10.1007/s11423-007-9045-6

Kisiel, J., Rowe, S., Vartabedian, M. A., & Kopczak, C. (2012). Evidence for family engagement in scientific reasoning at interactive animal exhibits.Science Education, 96(6), 1047-1070.

Krygier, J., & Wood, D. (2011). Making maps: A visual guide to map design for GIS (2nd ed. ed.). New York: Guilford Press.

Leinhardt, G., & Crowley, K. (1998). Conversational elaboration as a process and an outcome of museum learning. Museum Learning Collaborative Technical Report (MLC-01). Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh.

Leinhardt, G., Crowley, K., & Knutson, K. (Eds.). (2003). Learning conversations in museums. Taylor & Francis.

Libarkin, J. (2014, February). Communicating Science: Best Practices for Visualizations. Lecture conducted from the University of Illinois-Chicago, Chicago, IL.

Libarkin, J., & Brick, C. (2002). Research methodologies in science education: Visualization and the geosciences. Journal of Geoscience Education, 50(4), 449-455.

Lindgren, R. (2009). Perspective-Based Learning in Virtual Environments. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.

Lindgren, R. (2012). Generating a learning stance through perspective-taking in a virtual environment. Computers in Human Behavior, 28(4), 1130-1139.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment six precepts for research on embodied learning and mixed reality. Educational Researcher, 42(8), 445-452.

Lindgren, R., & Moshell, J. M. (2011, June). Supporting children's learning with body-based metaphors in a mixed reality environment. In Proceedings of the 10th International Conference on Interaction Design and Children (pp. 177-180). ACM.

Lloyd, W. J. (2001). Integrating GIS into the undergraduate learning environment. Journal of Geography, VOL 100(Part 5), 158-159-163.

Lozano, S. C, Hard, B. M., & Tversky, B. (2006). Perspective taking promotes action understanding and learning. Journal of Experimental Psychology: Human Perception and Performance, 32(6), 1405-1421.

Lyons, L. (2009, June). Designing opportunistic user interfaces to support a collaborative museum exhibit. In Proceedings of the 9th international conference on Computer supported collaborative learning-Volume 1 (pp. 375-384). International Society of the Learning Sciences.

Lyons, L., and Roberts, J. (2014, June). Scaffolding the Collaborative Interpretation of Data in a Museum Exhibit. In Leilah Lyons (chair) Synergistic Scaffolding of Technologically-enhanced STEM Learning in Informal Institutions. Symposium conducted at the 11th International Conference of the Learning Sciences. Boulder, CO.

Lyons, L., Tissenbaum, M., Berland, M., Eydt, R., Wielgus, L., & Mechtley, A. (2015, June). Designing visible engineering: supporting tinkering performances in museums. In Proceedings of the 14th International Conference on Interaction Design and Children (pp. 49-58). ACM.

MacEachren, A. M. (2005). Moving geovisualization toward support for group work. Exploring geovisualization, 461.

MacEachren, A. M., & Brewer, I. (2004). Developing a conceptual framework for visually-enabled geocollaboration. International Journal of Geographical Information Science, 18(1), 1-34.

Malinverni, L., & Burguès, N. P. (2015, June). The medium matters: the impact of full-body interaction on the socio-affective aspects of collaboration. In Proceedings of the 14th International Conference on Interaction Design and Children (pp. 89-98). ACM.

Marsh, M., Golledge, R., & Battersby, S. E. (2007). Geospatial concept understanding and recognition in G6-college students: A preliminary argument for minimal GIS. Annals of the Association of American Geographers, 97(4), 696-712.

Marshall, P., Morris, R., Rogers, Y., Kreitmayer, S., & Davies, M. (2011, May). Rethinking'multi-user': an in-the-wild study of how groups approach a walk-up-and-use tabletop interface. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 3033-3042). ACM.

McCabe, C. A. (2009). Effects of data complexity and map abstraction on the perception of patterns in infectious disease animations. Unpublished MS, The Pennsylvania State University.

McManus, P. (1994). Families in museums. In Miks L. & Zavala A. (Eds.), Towards the museum of the future (pp. 81–118). London: Routledge.

Melendez, J. Roberts, J. Radinsky, J. (2010, November). GIS census data browsers: Tools for constructing social and historical narratives. In A. Beveridge (Chair), New Directions in Historical GIS(1): Substantive Findings and Web-Based Resources. Paper session conducted at the meeting of the Social Science History Association, Chicago, IL.

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook. Sage.

Monmonier, M. (1991). How to lie with maps. Chicago, IL: The University of Chicago Press.

Murray, T. S., Kirsch, I. S., & Jenkins, L. B. (1998). Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey. US Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328. Web site: http://nces. ed. gov..

National Research Council. (2009). Learning Science in Informal Environments: People, Places, and Pursuits. The National Academies Press, Washington, DC.

Norman, D. A. (2010). Natural user interfaces are not natural. Interactions,17(3), 6-10.

Ochs, E., Gonzales, P., & Jacoby, S. (1996). " When I come down I'm in the domain state": grammar and graphic representation in the interpretive activity of physicists. Studies in Interactional Sociolinguistics, 13, 328-369.

Packer, J., & Ballantyne, R. (2005). Solitary vs. shared: Exploring the social dimension of museum learning. Curator: The Museum Journal, 48(2), 177-192.

Paris, S. G. 1997. Situated motivation and informal learning. Journal of Museum Education 22 (2): 22-27

Pew Research Center (2015). Technology Device Ownership: 2015. Smartphones, Tablets Grew in Recent Years; Other Devices Declined or Stayed Flat. Retrieved from http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/pi_2015-10-29_device-ownership_0-01/

Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educationally effective animations and simulations. Journal of Computing in Higher Education, 21(1), 31-61.

Pousman, Z., Stasko, J. T., & Mateas, M. (2007). Casual information visualization: Depictions of data in everyday life. Visualization and Computer Graphics, IEEE Transactions on, 13(6), 1145-1152.

Povis, K. T., & Crowley, K. (2015). Family Learning in Object-Based Museums: The Role of Joint Attention. Visitor Studies, 18(2), 168-182.

Price, S., Sakr, M., & Jewitt, C. (2015). Exploring whole-body interaction and design for museums. Interacting with Computers, iwv032.

Radinsky, J., Melendez, J., Roberts, J. (2012, April) Do the Data Strike Back? Students' Presentations of Historical Narratives about Latino Communities using GIS. In Josh Radinsky (Chair), Tools for Constructing Historical Narratives: Teaching African American and Latino Histories With GIS Census Maps. Symposium conducted at the meeting of the American Educational Research Association, Vancouver, B.C., Canada.

Richards, T. (2012). Using kinesthetic activities to teach Ptolemaic and Copernican tetrograde motion. Science & Education, 21(6), 899-910.

Rittschof, K., & Kulhavy, R. (1998). Learning and remembering from thematic maps of familiar regions Springer Boston.

Roberts, L. C. (1997). From Knowledge to Narrative: Educators and the changing museum. Washington: Smithsonian Institution Press.

Roberts, J., Lyons, L., Radinsky, J. (2013) Become one with the data: Technological support of shared exploration of data in informal settings. In Anne Knowles (Chair), From Visualizing to Understanding Historical Change: Using GIS Tools on the Web, in Class, and in Museums. Paper session conducted at the meeting of the Social Science History Association.

Roberts, J., Radinsky, J., Lyons, L., Cafaro, F. (2012, April) Co-Census: Designing an Interactive Museum Space to Prompt Negotiated Narratives of Ethnicity, Community, and Identity. In Josh Radinsky (Chair), Tools for Constructing Historical Narratives: Teaching African American and Latino Histories With GIS Census Maps. Symposium conducted at the meeting of the American Educational Research Association, Vancouver, B.C., Canada.

Roberts, J., Cafaro, C., Kang, R., Vogt, K., Lyons, L., & Radinsky, J. (2013). That's Me and That's You: Museum visitors' perspective-taking around an embodied interaction data map display. In Proceedings of the 10th International Conference on Computer-Supported Collaborative Learning. Madison, WI, USA. 343-344.

Roberts, J., Lyons, L., Cafaro, F., Eydt, R. (2014). Interpreting Data from Within: Supporting Human-Data Interaction in Museum Exhibits Through Personalization. In Proceedings of the 13th International conference on Interaction Design & Children. Aarhus, Denmark.

Rounds, J. (2006). Doing identity work in museums. Curator: The Museum Journal, 49(2), 133-150.

Saldaña, J. (2009). The coding manual for qualitative researchers. Sage

Schauble, L., Gleason, M., Lehrer, R., Bartlett, K., Petrosino, A., Allen, A., ... & Street, J. (2002). Supporting science learning in museums. Learning conversations in museums, 425-452.

Schiano, J. D., and Tversky, B. (1992). Structure and strategy in encoding simplified graphs. Mem. Cogn. 20: 12–20.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. Educational Psychology Review, 14(1), 47-69.

Snibbe, S. S., & Raffle, H. S. (2009, April). Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. InProceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1447-1456). ACM.

Stahl, G. (2000). A Model of Collaborative Knowledge-Building. In B. Fishman & S. O'Connor-Divelbiss (Eds.), Fourth International Conference of the Learning Sciences (pp. 70-77). Mahwah, NJ: Erlbaum.

Tashakkori, A., & Teddlie, C. (2010). Sage handbook of mixed methods in social & behavioral research (2nd ed.). Los Angeles: SAGE Publications.

Uttal, D.H., (2000), Seeing the big picture: Map use and the development of spatial cognition. Developmental Science, V. 3, p. 247-286.

van Dam, L. L. (1992). A picture is worth 1,000 numbers. Technology Review (00401692), 95(4), 34.

Vom Lehn, D., Heath, C., & Hindmarsh, J. (2001). Exhibiting interaction: Conduct and collaboration in museums and galleries. Symbolic interaction,24(2), 189-216.

Vygotsky, L. S. (1978). Mind and society: The development of higher mental processes.Cambridge, MA: Harvard University Press.

Wells, G. (2000). Dialogic inquiry in education. Vygotskian perspectives on literacy research, 51-85.

Wertsch, J. V. (1998). Mind as action. Oxford University Press, USA.

Williams, A., Kabisch, E., & Dourish, P. (2005). From interaction to participation: Configuring space through embodied interaction. Proceedings of the Ubicomp 2005, LNCS 3660 (pp. 287-304).

Yatani, K., Sugimoto, M., & Kusunoki, F. (2004). Musex: A System for Supporting Children's Collaborative Learning in a Museum with PDAs. In Proceedings of the 2004 IEEE International Workshop on Wireless and Mobile Technologies in Education, Taoyuan, Taiwan, (pp. 109-113). IEEE Computer Society.

Yiannoutsou, N., Papadimitriou, I., Komis, V., & Avouris, N. (2009). "Playing with" museum exhibits: designing educational games mediated by mobile technology. In Proceedings of the 8th International Conference on Interaction Design and Children, Como, Italy, (pp. 230-233). ACM

Zimmerman, H. T., Reeve, S., & Bell, P. (2008). Distributed expertise in a science center: Social and intellectual role-taking by families. *Journal of Museum Education*, *33*(2), 143-152.

Jessica Ann Roberts

www.jessicaannroberts.com

Education

University of Illinois at Chicago, PhD in Learning Sciences 2016

Specializations: Data Visualization and Geographic Information Systems, Informal Learning
Geospatial Analysis and Visualization Certificate (College of Urban Planning)
Dissertation: Connecting Visitors to Data: Exploring Tools for Mediating Learning Talk at an Interactive Museum Exhibit

National-Louis University, Alternative Certification Program for Elementary Education 2004

Self-contained K-9 Certification

Northwestern University, Bachelor of Science in Interdepartmental Studies in Communication 2003

Specialization: Theatre lighting design

Skills Summary

Research	Seven years of experience in study design, data collection and management, IRB documentation, participant recruitment, video and audio transcription, design research, mixed methods, open coding, and grounded theory
Design	Design and development of graphical user interface, map display, and interaction for CoCensus museum exhibit; over 30 theatrical productions as lighting director, lighting designer, assistant lighting designer, or sound technician
Software & Coding	Proficient in Adobe Illustrator CS5, Adobe InDesign CS5, ESRI ArcGIS 10, NVivo, Inqscribe, Wordpress, MaxQDA12; experience with Processing 2.0, HTML5, CSS, PHP, MySQL, SPSS, Adobe Photoshop CS5, SketchUp, D3, R, ATLAS.ti, Qualtrics
Teaching	Classroom experience in K-8 elementary, after-school adult math class (English and Spanish), and Kaplan test prep LSAT, GMAT (math only), and ACT; led workplace training to prepare attorneys and paralegals for participation in domestic violence pro bono clinic and to introduce software to fellow graduate students, including NVivo, RefWorks, and Scrivener; developed and conducted training for museum explainers on research methods in informal environments

Research Experience

Designing Digital Rails to Foster Scientific Curiosity around Museum Collections 2015 - present

Postdoctoral Researcher, Northwestern University & Field Museum of Natural History

- Develop innovative data collection methods for conducting *in situ* studies evaluating visitors' use of digital touchscreens installed in the Cyrus Tang Hall of China
- Liaise with museum staff to ensure the needs of all stakeholders are being met
- Manage and analyze collected data and prepare and present regular updates to team members
- Supervise first-year PhD student in project activities
- Collaborate with team members in design-based research activities

CoCensus: Collaborative Exploration of U.S. Census Data 2011 - 2015

Research Assistant, Learning Sciences Research Institute, University of Illinois at Chicago

- Worked with an interdisciplinary team of researchers to iteratively design and implement an interactive census data map exhibit at two museums
- Managed data collection and analysis of over 400 participants in multiple ongoing *in situ* studies
- Trained museum Explainers on facilitation strategies
- Led project branding efforts including logo development, direction and editing of promotional video, and website management
- Disseminated findings at international academic and professional conferences

Creating and Disseminating Tools to Teach with Demographic Data Maps and Materials 2009 - 2012

Research Assistant, Learning Sciences Research Institute, University of Illinois at Chicago

- Assisted in data collection and analysis for students and teachers' use of GIS data maps in classroom activities
- Recruited UIC faculty for design teams and collaborated with team members to create lesson plans
- Disseminated findings at international academic and professional conferences

Selected Publications

Roberts, J., & Lyons, L. (submitted for review). Bigger May Not Be Better: Impacts of Embodied Interaction and Distributed Control on Visitors' Learning Talk at an Interactive Data Museum Exhibit. Under review.

Roberts, J., Hall, A., & Goldman, S. (in press) Learning Sciences. In K. Peppler (Ed.), *SAGE Encyclopedia of Out of School Learning*.

Roberts, J., Banerjee, A., Matcuk, M., McGee, S., Horn, M. (2016). Uniting Big and Little Data to Understand Visitor Behavior. Poster presented at the annual meeting of the Visitor Studies Association, July 2016, Boston, MA.

Roberts, J., Lyons, L., Cafaro, F., & Eydt, R. (2015). Harnessing motion-sensing technologies to engage visitors with digital data. In Proceedings of the 19th annual Museums and the Web conference. Available online: http://mw2015.museumsandtheweb.com/proposal/harnessing-motion-sensing-technologies-to-engage-visitors-with-digital-data/

Roberts, J., Lyons, L., Cafaro, F., & Eydt, R. (2014, June). Interpreting data from within: Supporting human-data interaction in museum exhibits through perspective taking. In Proceedings of the 2014 Conference on Interaction Design and Children (pp. 7-16). ACM.

Lyons, L., and **Roberts, J.** (2014). Scaffolding the collaborative interpretation of data in a museum exhibit. In Leilah Lyons (chair), Synergistic Scaffolding of Technologically-enhanced STEM Learning in Informal Institutions. Symposium conducted at the International Conference of the Learning Sciences.

Roberts, J., Lyons, L., Radinsky, J. (2013) Become one with the data: Technological support of shared exploration of data in informal settings. In Anne Knowles (Chair), From Visualizing to Understanding Historical Change: Using GIS Tools on the Web, in Class, and in Museums. Paper session conducted at the meeting of the Social Science History Association.

Roberts, J., Cafaro, C., Kang, R., Vogt, K., Lyons, L., & Radinsky, J. (2013). That's me and that's you: Museum visitors' perspective-taking around an embodied interaction data map display. In Proceedings of the 10th International Conference on Computer-Supported Collaborative Learning (pp. 343-344).

Cafaro, F., Panella, A., Lyons, L., **Roberts, J.**, & Radinsky, J. (2013, April). I see you there!: Developing identity-preserving embodied interaction for museum exhibits. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1911-1920). ACM.

Roberts, J., Lyons, L., Radinsky, J., Cafaro, F. (2012). Connecting Visitors to Exhibits through Design: Exploring United States census data with CoCensus. In Proceedings of the 2012 International Conference of the Learning Sciences.

Roberts, J., Radinsky, J., Lyons, L., Cafaro, F. (2012, April) Co-Census: Designing an interactive museum space to prompt negotiated narratives of ethnicity, community, and identity. In Josh Radinsky (Chair), Tools for Constructing Historical Narratives: Teaching African American and Latino Histories With GIS Census Maps. Symposium conducted at the meeting of the American Educational Research Association.

Lyons, L., Becker, D., & **Roberts, J. A.** (2010). Analyzing the affordances of mobile technologies for informal science learning. Museums & Social Issues, 5(1), 87-102.

Melendez, J., **Roberts, J.**, Radinsky, J. (2010, November). GIS census data browsers: Tools for constructing social and historical narratives. In A. Beveridge (Chair), New Directions in Historical GIS(1): Substantive Findings and Web-Based Resources. Paper session conducted at the meeting of the Social Science History Association.

Doctoral Consortia & Invited Events

Cyberlab IGNITE Research Challenge Invited Participant (2015). Hatfield Marine Science Center, Oregon State University. Newport, Oregon.

Museums and the Web (MW 2015). Designing Exhibit Interactivity to Support Group Exploration of Digital Data. Graduate Student Colloquium. Chicago, Illinois.

Tangible, Embedded, & Embodied Interactions (TEI 2015). Exploring Effects of Full-body Control in Perspective-based Learning in an Interactive Museum Data Display. Doctoral Consortium. Palo Alto, California.

Computer-Supported Collaborative Learning (CSCL 2013). Designing an Interactive Exhibit for Exploring Complex Data in Informal Learning Environments. Doctoral Consortium. Madison, Wisconsin.

Service

Reviewer: Tangible, Embedded, and Embodied Interaction (TEI); SIGCHI Conference on Human Factors in Computing Systems (CHI)

Instuctional Workshops: Introduction to RefWorks (Learning Sciences Student Association - LSSA), NVivo (LSSA), Scrivener Demonstration (Learning Sciences Research Institute (LSRI) Brown Bag), Informal Research Training Course for Practitioners (New York Hall of Science Explainers)

Technology Demonstrations: Introduction to Museum Research (LSRI visiting high school students), CoCensus demo (NSF Assistant Director visit; UIC Board of Visitors)

Committee: Museums and the Web Local Committee (2015)

Community Involvement

Director of Student Relations and Outreach, Do the Write Thing Challenge Chicago, 2009-present

Secretary, Gardens of Ravenswood Condominium Association, 2010-2013

Young Alumni President, Northwestern University Alumni Club of Chicago, 2007-2009

Additional Experience

Editorial Assistant, Journal of the Learning Sciences — 2012 - 2014

- Managed over 150 manuscripts annually from initial submission to final decision
- Corresponded with authors and editorial team about journal business
- Organized quarterly and annual reports and meetings for editorial team

Litigation Paralegal, Latham & Watkins, LLP — 2006 - 2009

- Organized, indexed, and maintained files of case-related correspondence and court filings
- Prepared documents for depositions and court filings
- Managed logistics of pro bono domestic violence clinic, including developing training presentations and resource materials for attorneys and paralegals, compiling reports for mayor's office and press release, and supervising paralegals and project assistants on site at domestic violence courthouse

Elementary Teacher, Chicago Public Schools — 2003 - 2006

- Constructed curriculum for 4th, 7th, and 8th grade self-contained classes using primary and secondary sources
- Adapted lesson plans to reach varying levels in split 7th and 8th grade classroom
- Maintained highest standardized test scores, highest attendance rates, highest uniform policy compliance, and lowest suspension rates in upper grade classrooms