Designing for the Cooperative Use of Multi-user, Multi-device Museum Exhibits

by

Leilah Blakeney Lyons

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Computer Science and Engineering)
in The University of Michigan
2008

Doctoral Committee:

Professor Elliot Soloway, Chair
Professor John E. Laird
Associate Professor Mark S. Ackerman
Assistant Professor Christopher L. Quintana

DEDICATION

I would like dedicate this work to my family, whose support and encouragement helped me keep my stamina throughout this very long process. First and foremost, I must acknowledge my husband, Brett Lyons, a true partner who has stood by me throughout, even when it was at some cost to himself and his ambitions. My mother, Martha Wright Toth, who not only inspired me to look to science education as an area wherein I could make a difference, but who to this day is willing to wield her editorial pen over my writing, regardless of the lateness of the hour (or the "loomingness" of the deadline). My sister, Valerie Wright Toth, with whom I have shared so many interests and activities. My newfound family through Brett, especially Alona and Sharon Moon, who not only raised the man I'm lucky to share my life with, but whose unfailing support and love have made me feel blessed to have two mothers-in-law.

Finally, but not least, I would like to dedicate this work in the memory of two irreplaceable men: my father, Gerard Stephen Toth, and my father-in-law, John Ian Lyons. I know that they would have been proud.

ACKNOWLEDGEMENTS

This work would not have been possible without the inspiration, advice, and direct assistance of many people. Foremost, I would like to thank my adviser, Elliot Soloway, for his innovative suggestions and his willingness to allow me to explore even out-of-the-ordinary ideas. I owe a tremendous debt to Christopher Quintana, who over the years provided reams of practical advice and material support, and was always willing to make time for discussions.

This study would not have been possible without Sherry Hsi, who took me under her wing and devoted more energy than I could ever have hoped for to expertly mentor me while at the Exploratorium. Likewise, this entire project would not have been possible without Joseph Chigwan Lee, my former partner-in-crime, without whom the initial MUSHI framework would never have seen the light of day.

I am incredibly grateful to those at the Exploratorium who put aside their own work to assist me, especially Cassie Byrd and Melody Yang, who were always ready to help lug equipment, operate video cameras, and provide their pleasant company. The staff of the Exploratorium's Department of Visitor Research & Evaluation graciously allowed me to make use of their top-notch research lab, and I would particularly like to thank the Director Sue Allen for agreeing to the imposition, and Adam Klinger, for taking time out of his busy schedule to bring me up to speed.

I must thank Zbigniew Pasek, who over the years offered advice more akin to a mentor than an employer, and introduced me to the Museum Studies Program. I am grateful to Ray Silverman and Brad Taylor of the Museum Studies Program, who made room for an out-of-the-ordinary engineering student.

Although my other two committee members were not involved in the design of this research, they still graciously agreed to come on board for the final stages. I would like to thank them for that, and specifically thank John Laird for helping a young undergraduate realize that games and research *can* mix, and Mark Ackerman for providing professional mentoring above-and-beyond the call of duty.

Julie Weber was always available to provide last-minute critiques of my writing and made sure I got out for regular exercise; her friendship made these last few years much more enjoyable.

I would be remiss if I did not acknowledge the others who were there in the early days to help with the initial MUSHI implementation, Richard Vath and Makiko Kawamura, and Linda Kendall and John Merlin Williams, who helped "seed" MUSHI through the GROCS program.

The members of the UM3D lab (Prof. Beier, Eric, Lars, Steffen, Scott, & Shawn) helped in hundreds of small ways over the years, and didn't mind that I made the lab my second home.

And of course, I must thank the support staff at the University of Michigan, without whom none of this would be possible: Dawn Freysinger, Becky Turanski, and Karen Liska.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

APE	Active Prolonged Engagement
AR	Augmented Reality
CAI	Computer-Aided Instruction
CAS	Complex Adaptive Systems
CBI	Computer-Based Instruction
CSCL	Computer-Supported Collaborative Learning
CSCW	Computer-Supported Collaborative Work
DBR	Design-Based Research
GOMS	Goals-Operators-Methods-Selection Rules
HCI	Human-Computer Interaction
IrDA	Infrared Data Association
LCD	Learner-Centered Design
MMUI	Multi-Machine User Interface
MUSHI	Multi-User Simulation with Handheld Integration
NFC	Near Field Communication
O-UI	Opportunistic User Interface
PDA	Personal Digital Assistant
SDG	Single-Display Groupware
STEM	Science, Technology, Engineering, and Math
TCP	Transmission Control Protocol
UCD	User-Centered Design
UDP	User Datagram Protocol
VR	Virtual Reality

CHAPTER 1

Introduction

Museum visitors, and visitors to science-oriented museums in particular, seldom attend alone – they almost always attend in social groups, and also tend to prefer to engage in shared learning experiences. More and more museum visitors are attending museums with mobile computing devices, like cellular phones, in their pockets. How might museums design computer-based activities for *groups* of users to engage with an exhibit by using their own personal devices? That is the design space this research explores.

1.1 Summary

The proposal here is to allow visitors to employ their own personal computational devices (e.g., cell phones, Personal Data Assistants, etc.) as Opportunistic User Interfaces (O-UIs) to an exhibit hosted on a museum floor. This work aims to begin exploring the design space for O-UIs employed to help groups of museum visitors jointly interact with a shared simulation of a scientific phenomenon (see Figure 1). The simulation used in this work was that of cancer growth in human tissue. After joining, visitors are given the task of working together to treat the cancer by interactively applying analogues of real-world cancer treatments.

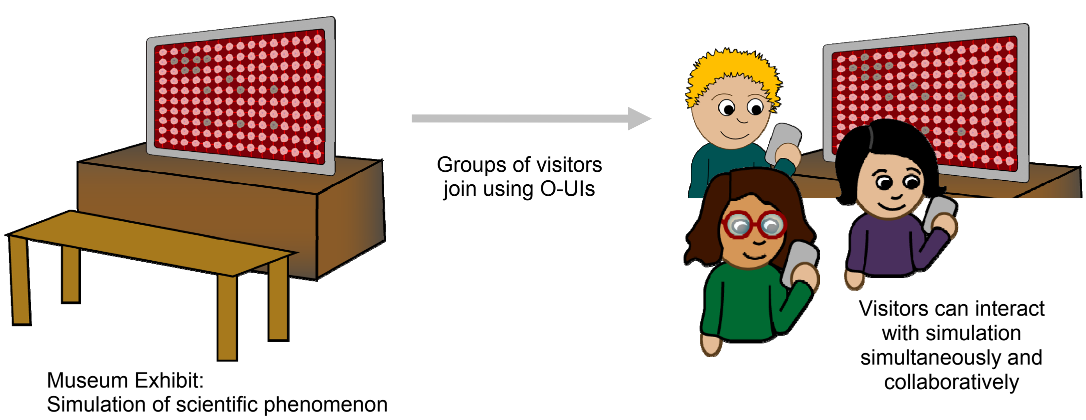


Figure 1. Depiction of proposed multi-user computer-based exhibit paradigm. Visitors use their own personal devices as Opportunistic User Interfaces (O-UIs) to join and collaboratively interact with a simulation of a scientific phenomenon.

Personal computational devices grow increasingly sophisticated in computational, graphical, and communication capabilities each year, which allow for a wide range of interaction possibilities. As always, though, just because one *can* do something with technology, doesn't mean that one *should*, especially if one's primary aim is to educate. There is evidence from museum practice and educational psychology that if designers take full advantage of O-UI capabilities, especially their dynamic graphical capabilities, the resulting experience might actually impair learning. Essentially, by asking visitors to divide their attention between too many stimuli (the exhibit, the visitor's personal O-UI to that exhibit, and the visitor's companions), designers might unintentionally make it harder for visitors to have the sort of collaborative learning experience museums would like to promote. For that reason, the main experiment presented in this work seeks to bracket the graphical design space for O-UIs, contrasting a condition where the O-UIs use no graphics at all (the "Simple" condition) against a condition where the O-UIs employ detailed, dynamic graphics (the "Complex" condition). Figure 2 depicts this design space. The purpose of the experiment was to discover two things: (1), to confirm if more graphically "complex" O-UIs do in fact promote poor visual attention management (known as the "heads-down phenomenon"), and (2) to determine to what extent, if any, the use of more "complex" O-UIs impact the exhibit's ability to support collaborative learning.

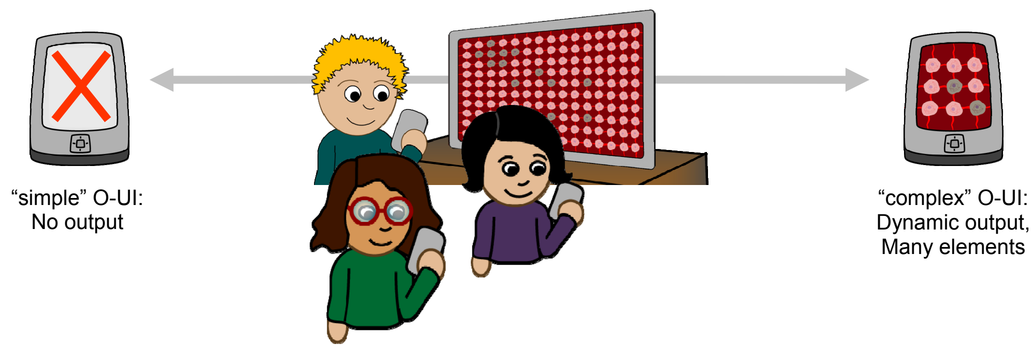


Figure 2. Design space explored by the controlled experiment in this research. One on end, the "Simple" condition provides no output. On the other end, the "Complex" O-UI provides dynamic output with many discrete graphical elements.

Groups of visitors to the Exploratorium, a hands-on science museum in San Francisco, were recruited from the floor of the museum to take part in the experiment, which was conducted in a controlled lab off the main floor of the museum. A repeated-measures-with-rotation experimental design was used, so that the behavior of the visitors in one condition could be compared against their behavior in the other condition. Around thirty-five participants were recruited altogether, with an average group size of about three. To collect the data needed to address the research questions, all user actions within the software were logged, questionnaires were administered to capture self-reported perceptions, and the experimental sessions were videotaped (from the audio of which a dialogue transcription was made).

The primary evidence for poor visual attention management comes from determining what users look at when, so the moment-to-moment target of each participant's gaze was coded. From this evidence, it is very clear that displaying "complex" output on O-UIs promotes poor visual attention management – the "Complex" participants gazed at their O-UIs proportionally more often, and for longer unbroken durations, than "Simple" participants. The "heads-down" visual attention behaviors led "Complex" participants to miss out on engaging with other elements of the shared context as well, like the shared display and their companions, as compared to the "Simple" condition. This is exactly what would be expected from the "heads-down phenomenon," first observed in the users of audio-visual guides in museums[1].

The evidence for whether or not collaborative learning was supported falls into three categories: the extent to which visitors displayed an *Awareness of Goals* within the joint experience, the quality of *Interaction* between visitors, and the *Equity* of visitor participation and performance. The first two categories (goal awareness and interaction) are commonly considered to be prerequisites to collaborative learning, and the third (equity) is of special interest for museums, which are tasked with producing exhibits that can help a broad range of visitors learn. Because the participants in the "Simple" condition have more of a What-You-See-Is-What-I-See (WYSIWIS) experience with the exhibit, one might assume that they would be more likely (and able) to collaborate. The evidence shows quite the opposite, however: the "Complex" condition better supported collaborative learning in each of the three categories. The logs of user actions show that "Complex" participants have a better *Awareness* of the activity *Goals*. Analysis of visitor conversations shows that while "Simple" participants talk more, "Complex" participants have much higher-quality *Interactions*, as measured by the proportion of conversation devoted to the learning activity at hand. Finally, groups in the "Complex" condition show more participation and performance *Equity*, meaning that no group members were being left out during the activity. This is especially true when gender is taken into account: male participants in the "Simple" condition were more likely to participate than female participants.

The data suggests that those interested in designing O-UIs for collaborative, software-based museum exhibits should probably use "Complex" O-UIs, especially if one wishes to promote *Equity*. The data shows, though, that when "Complex" participants engaged in more "heads-up" behaviors (gazing more at the shared display, at more frequent intervals) they were able to obtain better outcomes within the activity. So the first design recommendation to emerge from this work is:

> **DR1:** *If designing O-UIs for museum exhibits, one can use "Complex" output, but should incorporate mechanisms to remind or encourage users to direct attention to shared display periodically.*

The means by which a designer should go about directing O-UI users' attention is still very much an open question. Standard prompting "stick" approaches (e.g., pop-up prompts, passive indicators) have not met with much success in other multi-device applications. It may be that an incentivized "carrot"

[1] Because the shared display and a visitor's companions are more likely to actively attract a visitor's attention than, say, a stone bust of a Roman general, there was some hope that "Complex" O-UI users would not suffer from the heads-down phenomenon observed in AV guide users.

approach would be more successful, like information fission, when critical information is distributed across devices.

A deeper analysis of the data shows that the WYSIWIS nature of the “Simple” condition, rather than encouraging collaboration, was actually encouraging emergent competition between group members. In Computer-Supported Collaborative Work (CSCW) literature, it has been thought that collaboration can be encouraged (or even enforced) by tightening the input and output “coupling” of the system. A WYSIWIS system has the tightest possible output coupling – the users are sharing the exact same output. By way of contrast, the “Complex” condition in this experiment offered loosely-coupled output: although the users shared the exhibit’s large display, they each had their own individual displays on their O-UIs. One other research group found that tight output coupling encouraged competition, not collaboration, and recommended that designers choose between two imperfect alternatives: either accept that competition will emerge, or restrict the input to the system so that independent input is impossible so as to enforce collaboration. (This latter approach is an example of tight input coupling). This work suggests a third, possibly more palatable approach, which is encoded into a design recommendation:

DR2: *When designing collaborative activities that use a shared output device, one can reduce competition by providing private, loosely-coupled outputs to each participant in addition to the shared output.*

It bears mentioning, when considering coupling and its impact on collaboration, that “input coupling” is usually thought of only in an action-level, tactical sense. For example, a collaborative drawing program with tight input coupling might require that, in order to draw a rectangle, one user must control the placement of the top-left coordinate, while the other must drag the bottom-right coordinate into place. An alternative approach is to consider input coupling on a strategic level. Educational research on collaboration in classroom settings is rife with the use of interdependent roles to encourage (or even enforce) collaboration. For example, in a “jigsaw” activity, each learner is assigned a unique role that is necessary to the completion of the collaborative task. Although never labeled as such, jigsawing is an example of tight strategic input coupling. In the experiments in this work, the input coupling was extremely loose from both a tactical and strategic perspective, and was not manipulated. Oftentimes users find tight tactical input coupling to be inconvenient and an encumbrance, even though it does act to enforce collaboration. What might happen if, rather than manipulating the tactical input coupling to encourage collaboration, the strategic input coupling was tightened instead? An exploratory study was conducted on the floor of the Exploratorium, wherein visitors could assume different roles while engaging with the exhibit’s simulation (each role reflected a real-world cancer treatment option). Analysis is ongoing, but should shed light on the potential utility of strategic input coupling to encourage collaboration in multi-user, multi-device museum exhibits.

1.2 Designing for the Cooperative Use of Multi-user, Multi-device Museum Exhibits

1.2.1 Motivation

Science museums provide an excellent platform for performing Science, Technology, Engineering, and Mathematics (STEM) educational outreach – a single exhibit can literally reach tens of thousands of visitors a year. In a study performed by the author's own work, a single exhibit in a small science museum was used by over 16,000 visitors in a single year, roughly 14% of the city's population (Lyons & Pasek, 2006). Another reason to look to museums as locations for STEM education is that they may be well positioned to influence the course of children's lives. In a large-scale study that tracked thousands of students from early school years through their undergraduate years of education, the single largest predictor of whether or not they would pursue a STEM career was *not* aptitude, socioeconomic class, or grades: it was an early expressed interest in science, technology, engineering, and math topics (Tai, Liu, Maltese, & Fan, 2006). Museums were cited as one of the premier places where students had experiences that piqued their interest in science, technology, engineering, and math topics.

Many science museums already use physically interactive exhibits to present STEM phenomena, and by employing interactive computing technology as a basis for exhibits, even wider ranges of phenomena can be addressed. Because well above 90% of all science museum visitors attend in groups, often with the express purpose of sharing a learning experience, it is important to try to design some of those computer-based exhibits to support simultaneous use by groups of visitors.

1.2.2 Proposal: Opportunistic User Interfaces

This work explores the design of software-based museum exhibits where groups of visitors can employ their own personal mobile devices as impromptu user interfaces to the exhibit. Personal devices commandeered into service in this fashion will be dubbed Opportunistic User Interfaces (O-UIs) in the remainder of this work. Because museum visitors, especially visitors to science museums, usually prefer to engage in shared learning experiences, emphasis is placed on how to design software interfaces to support collaborative learning. To study the issue, a Design-Based Research (DBR) approach was taken. DBR emphasizes theoretically-grounded design, and alternates exploratory *in situ* formative testing phases (which produce artifacts that actually function in their intended use contexts) with retrospective experimental phases (which are used to verify, amend, or challenge the theories underlying the artifact's design). The purpose behind using a DBR methodology was to construct an exemplar of this type of multi-device exhibit that would have credible external validity, while also conducting more traditional experiments to explore the O-UI design.

There is no tailor-made body theory, as yet, to guide the design of software-based museum exhibits, so three analyses, of (1) museums as a context, (2) existing computer-based exhibits found in museums, and (3) computer support of collaborative processes in both work and classroom contexts, were performed to generate theoretically-grounded guidelines for the design of software-based museum exhibits.

The guidelines were then used to inform the design of the software-based exhibit that was created as a testbed for this research. During the formative phase of this work, the exhibit was refined via extensive testing on the floor of the Exploratorium, a hands-on science museum in San Francisco, CA. The testbed exhibit presents an open-ended dynamic simulation of a complex system: specifically, a simulation of cancer growth in human tissue. Museum visitors use O-UIs to log into the exhibit and try their hands at applying analogues of real-world cancer treatment options within the simulation, observing the emergent results of their efforts, and working together to eliminate the cancerous cells from the simulated tissue.

1.2.3 The Experimental Research Questions

The experimental phase of this work examined the impact of O-UI design on (1) the visual attention management and (2) collaborative learning behaviors of visitors. The concern was that by utilizing the full graphical display capabilities of mobile devices, designers might inadvertently make it harder for visitors to engage in the sort of collaborative learning that museums would like to promote. (Handheld audio-visual guides have been observed to monopolize visitor attention in museums, and educational psychology research warns against asking learners to divide their attention between too many visual stimuli). For the experiment, an O-UI design that did not display any graphical output (the “Simple” condition) was contrasted against an O-UI design that displayed multi-element, dynamically animated graphics (the “Complex” condition). These two conditions were chosen to effectively bracket the design space for O-UI output “complexity,” so that least complex display (i.e., no display at all) is contrasted against a worst-case scenario.

1.2.4 Background: The Problem with “Complex” Handheld Device Displays

1.2.4.1 Handhelds in Museums

Museums were relatively early adopters of mobile technology, using early Personal Data Assistants as audio/visual (A/V) replacements for the traditional audio-only guides that museum visitors were already well familiar with (e.g., Acoustiguide, n.d.). The research on Human-Computer Interaction (HCI) issues for handheld devices in museums, which began in earnest in the late 1990s, thus tended to focus on the use of handhelds as museum guides. One of the earliest observations was that museum visitors, while using these new A/V guides, would get so involved in interacting with the devices that they would fail to attend to their surroundings – an effect dubbed the “heads-down phenomenon” (Walter, 1996). Visitors reported feeling isolated from the museum experience and from their companions (Bellotti, Berta, Gloria, & Margarone, 2002; Exploratorium, 2005; Fleck et al., 2002; Hsi, 2002, 2003, 2004; Hsi, Semper, Brunette, Rea, & Borriello, 2004; Wessel & Mayr, 2007). Several researchers attempted to ameliorate the problem by “simplifying” the interfaces, by removing the audio portion of the experience (Hsi, 2002, 2003; Wessel & Mayr, 2007) or by recommending the use of “simple” graphics in the user interface (Bellotti et al., 2002; Fleck et al., 2002; Yatani, Sugimoto, & Kusunoki, 2004). Although these guides were designed as single-user experiences, it suggests that the proposed use of O-UIs to support multi-user museum exhibits may result in poor attention management behaviors in visitors if the interfaces are too "complex." This concern provided the motivation for the experimental study conducted in this research.

1.2.4.2 Defining O-UI "Complexity"

The prior research on handhelds in museums does not provide clear guidance on just what constitutes a "simple" handheld user interface. For the purpose of this research, there is a need for a definition that addresses the demands on visual attention that a user interface will impose – the heads-down effect's primary symptom is a handheld device's monopoly on visual attention. Some HCI researchers have attempted to define UI complexity by looking at factors that increase visual search times, like element size (smaller objects take longer to register visually), local density of elements (it takes longer to process dense arrangements), alignment (aligned elements are easier to scan), and grouping (clustering elements into functional groups reduces eye-travel time) (Miyoshi & Murata, 2001; Parush, Nadir, & Shtub, 1998). Of course, humans are also more or less "hard-wired" to respond to certain types of stimuli regardless of what a person consciously sets out to attend to, like larger, closer, and moving stimuli (Knudsen, 2007). Because O-UIs are already physically closer to the user than any other stimuli, they are likely to attract visual attention on that basis alone, but by placing graphics with many moving elements on the O-UI display, visitor attention is even more likely to be monopolized.

1.2.4.3 A Case for Not Using O-UI Displays

Educational psychologists have found that when a learner must divide his or her visual attention between stimuli located in different spatial locations, learning is impeded, a phenomenon known as the "Split Attention Effect" (Sweller, Chandler, Tierney, & Cooper, 1990). The underlying principle is that dividing one's visual attention between different spatial locations puts a high "cognitive load" on a one's working memory, leaving less capacity for managing other aspects of the learning task. For this reason, it may be better to not use the output displays of O-UIs whatsoever, and allow visitors to provide input via their device's hardware buttons (as if it was a remote control), so they may focus their full attention on the main exhibit (which, for the purposes of this research, is assumed to be a single large shared display).

Computer-Supported Collaborative Learning (CSCL) and Computer-Supported Collaborative Work (CSCW) researchers have also experimented with different combinations of devices and displays to support synchronous, co-located collaborative activities (i.e., collaborative activities that take place at the same time in the same location). A concept arising form this work is that of "coupling," used to describe how "tightly" the inputs and outputs allowed to the users of a collaborative system are tied together (Dewan & Choudhard, 1991). A system with "tight" coupling requires a high degree of simultaneous focus and action, and a system with "loose" coupling allows users to have more distinct foci and for them to take individual actions. For this reason, a system's degree of coupling is thought to also determine the degree of collaboration present during the joint activity, with loose coupling merely enabling collaboration, and tight coupling encouraging or even enforcing collaboration (Benford et al., 2000). Returning to the O-UI design issue at hand, then, by utilizing the displays on O-UIs, visitors are being presented with loosely-coupled output, since the device displays present each user with an independent view. If the O-UI displays are not utilized, the visitors are being presented with tightly-coupled output, since they are only able to view the single, shared display of the main exhibit. According to the notion of coupling, then, by not using the O-UI displays, visitors will be more strongly encouraged to work together.

1.2.5 Experimental Study Design

The prior section presented several arguments for why designers should pause before blindly taking advantage of the sophisticated graphical display capabilities of mobile devices when designing for O-UIs. The controlled experiment conducted during this research was designed to examine whether or not the poor visual attention management behaviors museum researchers warned of would come to pass when O-UIs displayed “complex” graphics (after all, the exhibit’s shared display also displays dynamic graphical elements, which may better serve to attract the visual attention of visitors than, say, an unmoving stone sculpture). It was also designed to determine if, on the one hand, the educational psychologists’ concerns about learners splitting their visual attention, and on the other, the HCI researcher’s notions of coupling encouraging collaboration, would predispose the graphics-free (or “simple”) version of the O-UI to better supporting collaborative learning. Before delving into the structure of the experiment, however, the software that was used as a testbed for the experiment must first be described.

1.2.5.1 The Design of the Multi-User, Multi-Device Exhibit Testbed

Hands-on science museums (unlike art or history or even natural history museums) do not usually present objects (like paintings, or artifacts, or fossils) to visitors. Rather, they are most often in the business of presenting *phenomena* to visitors, by constructing hands-on exhibits to allow visitors to directly interact with phenomena like gravity, electricity, or human visual perception. Computers can be very valuable in that they can present phenomena that otherwise could not be contained by a physical exhibit, whether by reason of scale (too small or too large, as with atomic interactions versus cosmic orbits), time (too fast or too slow, as with avalanches versus glaciation), or hazard (as with explosions). Via simulations, which present analogues of real-world phenomena, computer-based exhibits can give visitors the same degree of hands-on interactions and experimentation provided by more traditional physical exhibits. For the testbed exhibit, then, a simple phenomenon needed to be selected for presentation; one that was uncomplicated enough to be easily accessible to visitors, but rich enough to provide ample opportunity for experimentation and manipulation for multiple simultaneous users.

The desired richness could be provided by a simulation based on a cellular automaton model, which is a simulation comprised of many small simulated entities, or automata, that each obey a small rule set, but by virtue of interacting with one another can nonetheless generate near-endless varieties of emergent phenomena. Starting from the simplest possible cellular automaton simulation, the *Game of Life* (Gardner, 1970), a simulation was constructed to emulate cancer growth in human tissue. In this simulation, there are three types of automata are cancer cells, healthy cells, and blood vessel segments. Each automaton maintains information about its current state in the form of variables. For example, a healthy cell will maintain a current "health" variable, a variable that tracks cell age, and a "cumulative radiation exposure" variable. Each automaton shares a rule base with other automata of its type (i.e., all healthy cells obey the same set of rules, all blood vessel segments obey their own shared set of rules, etc.). There is no "controlling hand" to the simulation – the next state of the simulation is an outgrowth of each automaton performing its own state update. The rule sets were designed so that the simulation exhibits several emergent phenomena that are hallmarks of cancer in real life, like tumor growth and its associated angiogenesis, metastasis, and radiation-induced secondary cancers.

Visitors can interact with the simulation by administering analogues of real-world cancer treatments, like surgery, radiation, and proton beam radiation, via their O-UIs, which are wirelessly connected to the simulation. The impact of the different treatments on the automata variables was tuned to so as to emulate how the real-world treatments impact cancer cells and normal human tissue. Groups of visitors need to work together to operate on the patient, as the simulation is tuned so that no one user can eliminate the cancer cells before they take over 50% of the simulated tissue (at which point the simulated patient "dies," and the simulation restarts with new, randomly-generated seed parameters).

1.2.5.2 Experimental Conditions

The experimental need to control variables meant that, for the duration of the experiment, visitors were restricted to administering the same treatment: surgery. Two O-UIs were developed: one graphics-free and one that employed dynamic graphics (see Figure 3). Extensive formative testing was conducted on the floor of the museum to ensure that the two interfaces impacted the simulation in the same way, and that museum visitors would find them usable. The "Simple" O-UI was initially designed to show only a blank, black screen after the visitors joined the simulation, but formative testing showed that visitors thought the device was turning itself off, so a simple static image (brief instructions on how to use the hardware buttons) was substituted instead. The "Complex" users were also exposed to a very similar static instruction screen, which gets replaced by the main graphical interface after joining. The "Complex" O-UI was designed to display a magnified region of the cancer simulation. All of the cells on the O-UI screen are updated in real-time to match the current simulation status, and grow and change color dynamically as the automata's parameters are updated.

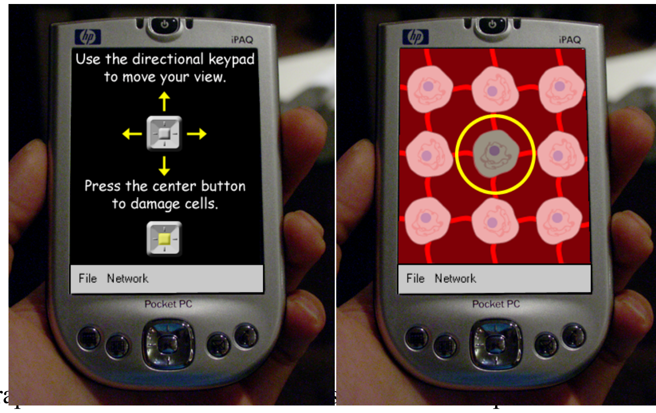


Figure 3. Photogra… …nterface on the left represents the “Simple” O-UI condition, as it displays a single, static image. The interface on the right represents the “Complex” O-UI condition, and displays a magnified region of the cancer growth simulation.

Input for positioning for both interfaces is provided by the directional pad (hardware buttons on the device). Visitors steer an “incision” around the grid of simulated cells, and can “operate” on cells within the incision (a rectangle that encompasses nine cells at a time). “Simple” users operate by pressing the center button of the directional control pad, which caused a circular region of tissue to be “excised.” “Complex” users, on the other hand, use a stylus to manually draw circles on the O-UI display to excise cells (to further increase the degree of visual attention required by the interface – recall the purpose is to try to explore a worst-case scenario for promoting the heads-down effect).

1.2.5.3 Setting and Procedure

The setting for both the formative and experimental phases of this research was the Exploratorium, a hands-on science museum located in San Francisco, CA. As one of the earliest hands-on science museums in the world, and as one of the museums at the forefront of exploring the use of mobile technology to improve visitor learning experiences, it was an ideal location to perform a Design-Based Research study. The lab used for the experiments was located in a small, controlled, (nearly) soundproof room located at the end of the museum at the opposite end of the building from the entrance. It was created for use by the in-house Visitor Research and Evaluation group. It was outfitted with state-of-the-art recording equipment, and a separate room that housed the computer system and hard drive array used to collect video and audio signals.

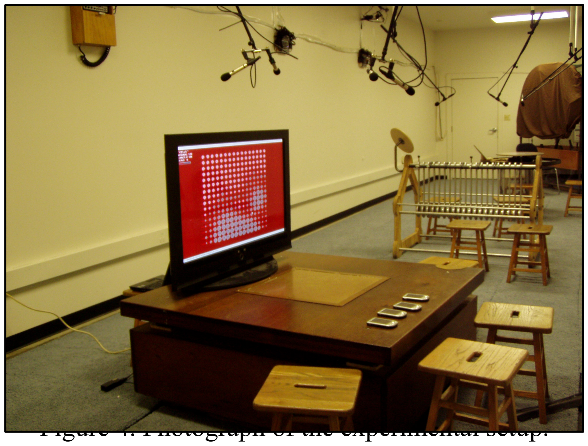


Figure 4. Photograph of the experimental setup.

The experiment took the form of a repeated-measures design with rotation: each group of visitors would be exposed to one of the conditions, then a questionnaire would be administered, and afterwards they would be exposed to the second condition. The order that the conditions were presented in was alternated, so that some groups would be exposed to the “Simple” condition first, whereas others would be exposed to the “Complex” condition first, to balance out any practice effects. The visitors were allowed to play as many rounds (defined by when the simulated patient would either die or be cured of cancer) as they desired. All sessions were videotaped by two cameras (one ceiling-mounted, and one mobile and on a tripod).

1.2.5.4 Participants

A total of 41 participants were recruited for the experiment, comprising about 11 groups of about three participants each that were exposed to each condition (the actual numbers vary slightly, since a few participants dropped out after the first condition). They were recruited via two methods: via a newsletter sent out to museum members, and from the floor of the museum itself. Four of the groups were mixed-age families, and 7 were groups of friends. The average age was about 26; the youngest participant was 10 years of age, and the oldest was 59, and the median age was 20. Most of those aged 18 and above had “some college” education. The gender ratio was roughly 45:55 female:male, with most groups having mixed gender composition.

1.2.6 Research Question 1: Establishing the Existence of the Heads-Down Effect

The heads-down phenomenon, as described by museum professionals, is characterized by both visual attention management behaviors (the tendency to stare fixedly at a handheld device for long periods) and by the consequences of those behaviors (a lack of awareness of the surrounding context). Evidence for each of these will be discussed in turn.

1.2.6.1 Visual Attention Management: Measures and Hypotheses

The primary evidence used to study visual attention behaviors comes from videotapes of the participants in the session. From the videotapes, the moment-to-moment gaze target of each participant was coded for analysis (the participant's O-UI, the shared display, or the participant's companions; all other targets were lumped into an "other" category). Using this data, several measures can be computed: the *Proportion* of total visual attention devoted to the various gaze targets, the average *Duration* of the gazes directed at each target, and the *Frequency* with which the participant shifted his or her gaze between targets. Given the prediction that participants in the "Complex" condition would suffer more from the heads-down phenomenon, the following specific hypotheses were put forth regarding individual attention management behaviors:

H1: *"Complex" participants will devote a larger* **Proportion** *of their gazes to their O-UI than "Simple" participants.*

H2: *"Complex" participants will gaze at their O-UIs for longer unbroken* **Durations** *than "Simple" participants.*

H3: *"Complex" participants will show a much lower* **Frequency** *of gaze shifts than "Simple" participants.*

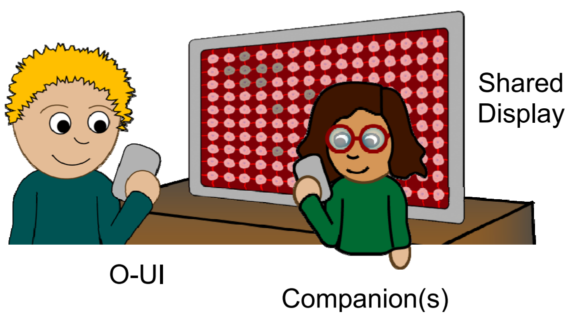


Figure 5. Illustration of the three main possible gaze targets for a participant in the experiment: the participant's O-UI, the Shared Display, and the participant's companions.

All of the hypotheses listed above refer to *individual* visual attention management, but because this is a collaborative activity, how the group as a whole managed its *joint attention* is also of interest. Joint attention management has been shown to be critical to the success of classroom-based collaborative learning exercises (Barron, 2000), the notion being that by attending to shared external artifacts (like the exhibit's shared display), learners can better coordinate their collaborative learning efforts. Joint attention also underlies the rationale for using any sort of shared display in collaborative work scenarios as well, from a chalkboard to a whiteboard to Single Display Groupware (Stewart, Bederson, & Druin, 1999). To gauge the degree of joint attention exhibited during this experiment, the individual-granularity visual attention data can be analyzed from a group-granularity perspective to produce a measure of *Gaze Synchronicity Degree*. The gaze targets for individual group members were sampled at 2-second intervals, and at each interval the proportion of group members simultaneously gazing at the shared display were recorded. These figures were then summed and divided by the length of the session, to provide an average *Gaze Synchronicity Degree* for the session. Once again, the prediction was that:

H4: *"Complex" groups will show a much lower* **Gaze Synchronicity Degree** *than "Simple" groups.*

1.2.6.2 Visual Attention Management: Results

The participants in the "Simple" and "Complex" conditions, while devoting about the same *Proportion* of time to gazing at their companions (around 2% of the play time), have diametrically opposed viewing habits regarding O-UIs (and, consequently, the Shared Display). Participants in the "Simple" condition spend only around 14% of their time gazing at the O-UIs, and the vast majority of their time, 83%, gazing at the Shared Display. On the other hand, participants in the "Complex" condition spend only 33% of their time looking at the Shared Display, reserving twice that amount, 65%, for their O-UIs (see Figure 6). The differences between the conditions are strongly significant for both the O-UI as target, $t(29) = 15.5$, $p < 0.0000001$ (one-tailed), as well as for the Shared Display as target, $t(29) = 14.8$, $p < 0.0000001$ (one-tailed), confirming prediction **H1.**

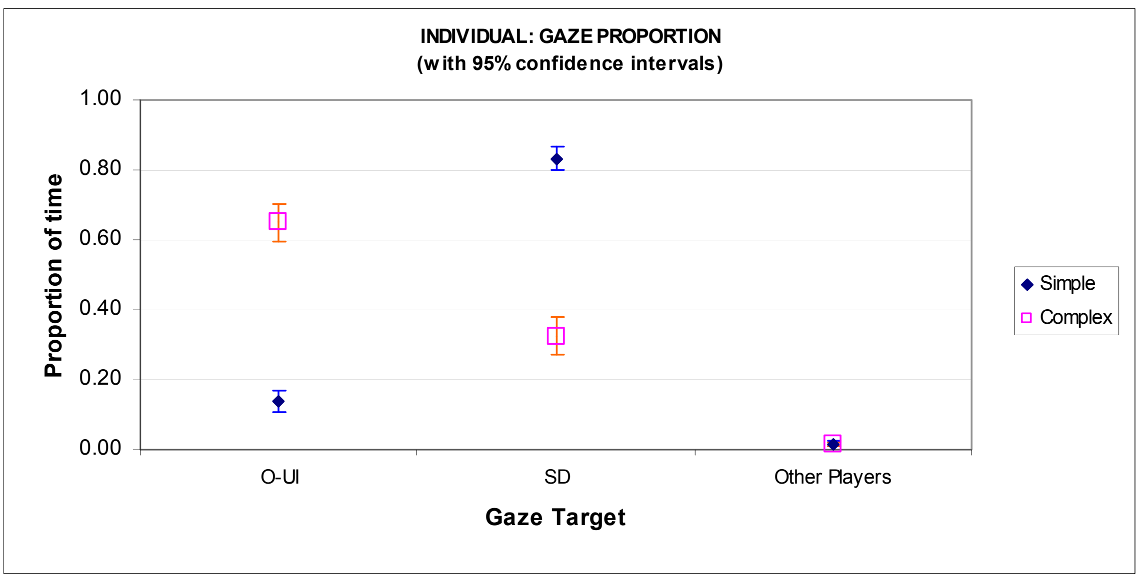


Figure 6. Plot of the *Gaze Proportions* that participants in the different conditions devoted to three different gaze targets, the Opportunistic User Interface (O-UI), the Shared Display (SD), and other members of the group, with 95% confidence intervals. (See Table 22 for more detailed data).

The "Complex" condition participants also stare at the O-UIs for significantly longer lengths of time, bearing out prediction **H2** (see Figure 7). Their gaze *Duration* when looking at their O-UIs averages 12.5 uninterrupted seconds (SD = 32.5), more than *four* times as long as they tend to look at the Shared Display (M = 2.8 s, SD =1.00). "Simple" participants show the opposite trend, gazing at their O-UIs for an average of 2.55 uninterrupted seconds (SD = 1.47), while gazing at the Shared Display around *eight* times as long (M = 20.6, SD = 24.8). The differences between conditions on O-UI gaze *Duration* are statistically significant, $t(29) = 1.81$, $p < 0.040$ (one-tailed), as are the differences between conditions on Shared Display gaze *Duration*, $t(29) = 4.25$, $p < 0.0001$ (one-tailed).

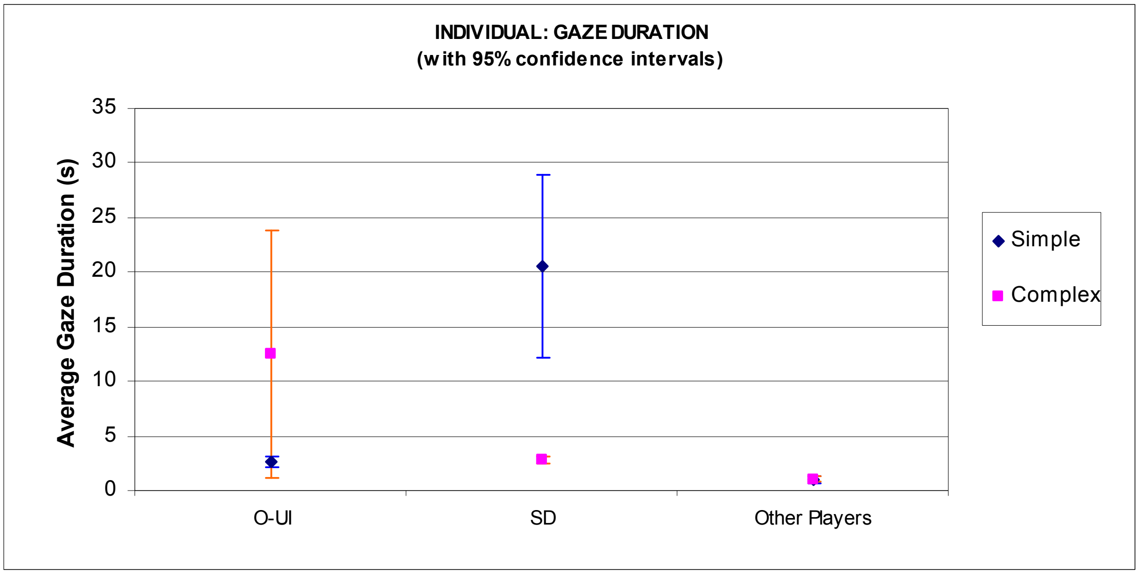


Figure 7. Plot of the average *Gaze Durations* that participants in the different conditions devoted to three different gaze targets, the Opportunistic User Interface (O-UI), the Shared Display (SD), and other members of the group, with 95% confidence intervals. (See Table 22 for more detailed data).

The third prediction, concerning Frequency, did not hold. "Complex" users exhibit significantly more *Frequent* gaze shifts, about 15 per minute (M = 15.0, SD = 5.87) in contrast to the roughly 10 gaze shifts per minute (M = 9.50, SD = 4.47) exhibited by the "Simple" participants, a different which was statistically significant, $t(29) = 6.15$, $p < 0.000001$, (two-tailed). This countermands prediction **H3**, which suggests that if players are suffering from the heads-down effect, they will shift their gaze less *Frequently*. A closer look at the gaze shits made by "Complex" users revealed that the higher Frequencies are the result of rapid fusillades of gaze shifting made by participants just prior to moving their incision rectangle on the Shared Display. At these times, visitors would shift their gaze back and forth between the O-UI and the Shared Display very rapidly, but would shift their gaze comparatively infrequently at other times. The countermanding of **H3** occurred because of a faulty assumption that the instances when participants shifted their gaze would be distributed relatively evenly across the session. Without an even distribution of gaze shifts, measuring the gaze shift frequency indicates little about the quality of a participant's monitoring practices.

The final prediction, **H4**, which was related to *group* visual attention management, was that "Complex" groups should demonstrate a lower gaze *Synchronicity Degree.* This prediction held very true: the average *Degree of Synchronicity* for "Simple" groups was 0.71 (SD = 0.19), compared to only 0.16 for "Complex" groups (SD = 0.10), $t(8) = 15.8$, $p < 3\text{x}10^{-16}$ (see Figure 8). To the extent that the gaze

Synchronicity Degree is an indicator of joint attention management, it seems to be showing that "Complex" groups do not engage in much joint attention management.

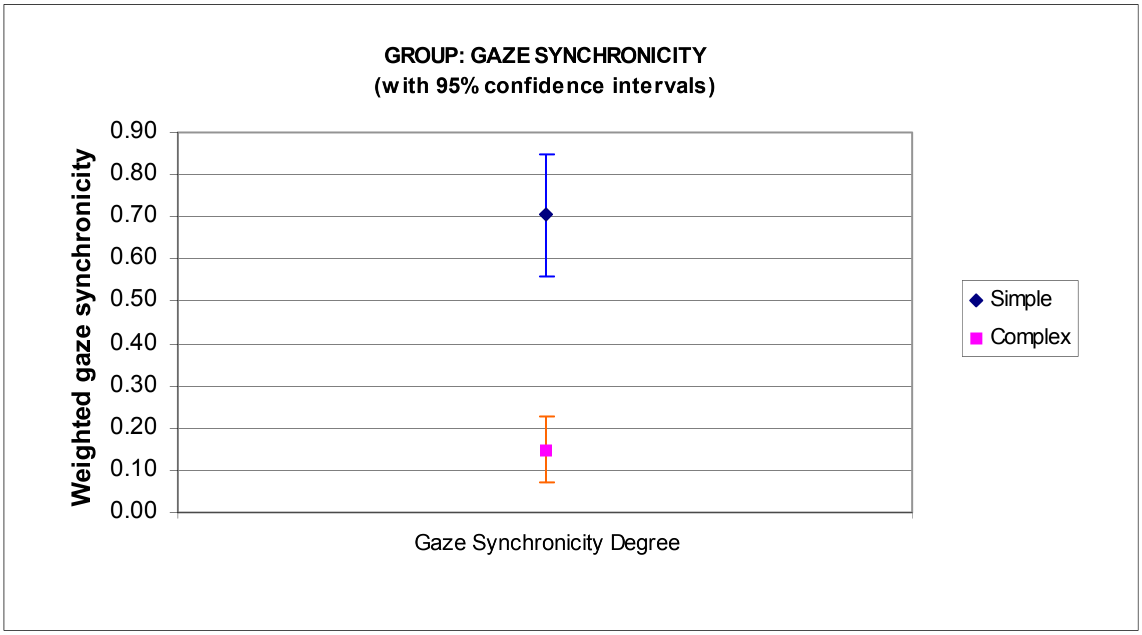


Figure 8. Plot of the group-granularity *Gaze Synchronicity Degree* measure, with 95% confidence intervals. (See Table 23 for more detailed data).

Taken together, the results show that the "Complex" participants do demonstrate visual attention management behaviors that are very much in line with what the heads-down phenomenon would predict. Despite the hoped-for potential of the dynamic graphics of the Shared Display or the actions of companions to be able to attract participants' visual attention, supplying visitors with graphically "complex" O-UIs does produce visual attention behaviors consistent with the heads-down phenomenon. The next step is to examine whether or not the deleterious secondary symptoms occur.

1.2.6.3 Awareness of Shared Context: Measures and Hypotheses

Gaze-related patterns of behavior are just the visible symptom, however; of the heads-down phenomenon. As described by museum practitioners, the secondary effect of the phenomenon is a marked lack of attention to other elements in the visitors' environment, which in this case includes both the visitor's companions and the Shared Display. If participants exhibit the heads-down phenomenon, one would also expect to see the following secondary symptoms that indicate a lack of awareness of companions:

H5: *Individual participants would report lower levels of* **Awareness** *of their partners' actions.*

H6: *Individual participants would engage in less* **Conversation** *with their partners.*

Participants must refer to the Shared Display in order to decide where they will move next, and so to the extent that the heads-down effect occurs, one would expect to see:

H7: *Individual participants would commit fewer* **Moves** *per unit time*

Assuming that the three secondary symptoms listed above are in fact caused by poor visual attention management behaviors, we would expect to see the following correlations:

H8: **a:** *Self-reported* **Awareness** *would correlate* **positively** *with more* **"heads-up"** *visual attention behaviors*

b: *Self-reported* **Awareness** *would correlate* **negatively** *with more* **"heads-down"** *visual attention behaviors*

H9: **a:** *Amount of* **Conversation** *would correlate* **positively** *with more* **"heads-up"** *visual attention behaviors*

b: *Amount of* **Conversation** *would correlate* **negatively** *with more* **"heads-down"** *visual attention behaviors*

H10: **a:** *Frequency of* **Moves** *would correlate* **positively** *with more* **"heads-up"** *visual attention behaviors*

b: *Frequency of* **Moves** *would correlate* **negatively** *with more* **"heads-down"** *visual attention behaviors*

1.2.6.4 Awareness of Shared Context: Results

Self-Reported Awareness

The self-reporting of participant awareness of the shared context was obtained by administering Likert-style questions to the participants only after they had been exposed to the first of the two experimental conditions, so that their responses would be a pure reflection of only one of the experimental conditions. On the questionnaire, two questions related to awareness. Question 7 (Q7): *"I was aware of how well my partners were doing at all times,"* and Question 9 (Q9): "*I was aware of how well our group was doing at all times."* Participants in both conditions responded fairly similarly (see Figure 9), although there was a significant difference between the way "Complex" participants answered Q7 and Q9, even though participants were expected to answer both questions similarly. This point will be referred to later, when collaboration is discussed.

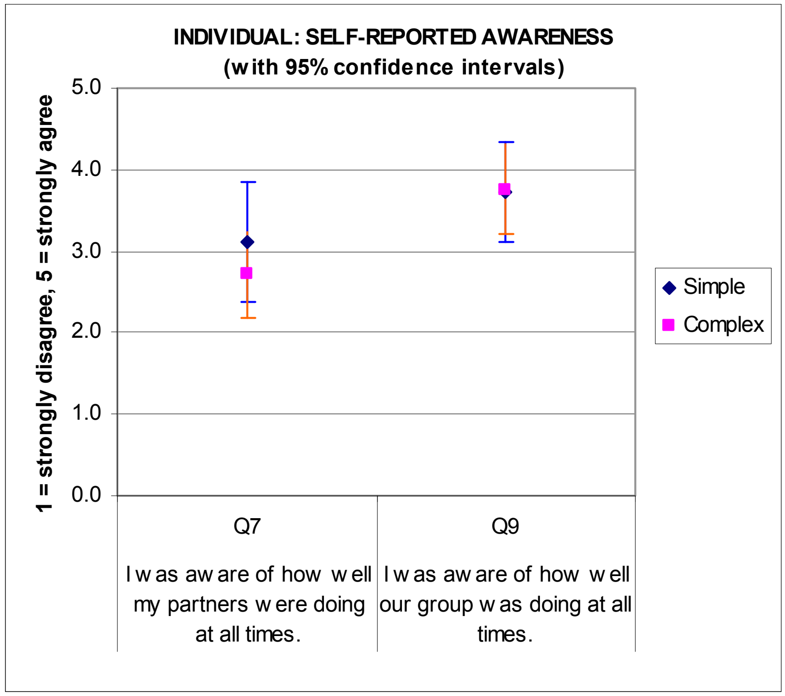


Figure 9. Plot of participant responses to Likert questions designed to elicit self-reports of awareness. (See Table 24 for more detailed data).

The responses given on Q7 and Q9 do correlate positively with "heads up" visual attention behaviors (namely, larger Proportions and Durations of gazes devoted to the Shared Display), as **H8a** predicted (see Figure 10). These positive correlations are significant for "Simple" participants on Q7, and for "Complex" participants on the relationship between Q9 and gaze *Proportion*. The evidence clearly suggests, though, that regardless of condition, paying *Proportionally* more attention to the Shared Display, and for longer Durations, results in better (self-reported) awareness of one's partners and group. Conversely, the responses on Q7 and Q9 tend to correlate negatively with "heads down" visual attention behaviors, as predicted by **H8b**, although the trends are only significant for "Complex" participants. It seems that "Complex" participants' reported awareness suffers more from long O-UI gaze *Durations* than "Simple" participants, which implies that "Complex" users could benefit from glancing up occasionally. Taking the evidence for predictions **H8a** and **H8b** together, it seems clear that visual attention management behaviors do affect participants' perception of their awareness of the shared context, despite the fact that there was no clear support for prediction **H5** ("Complex" participants would report lower -awareness). It may be a case of "Complex" participants "not knowing what they don't know," and overestimating their awareness. Looking at the less subjective measures for *Conversation* and *Move* frequencies should help establish if that is the case.

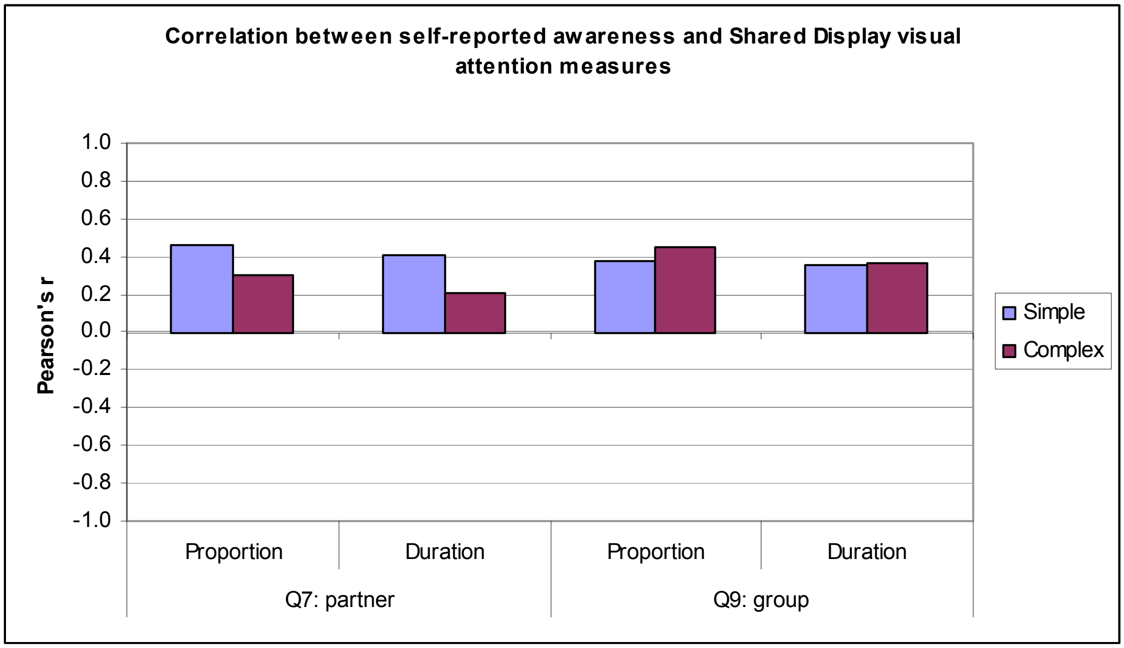


Figure 10. Bar chart depicting the correlation (Pearson's r) between Shared Display visual attention measures and the responses to Q7 (*"I was aware of how well my partners were doing at all times"*) and Q9 (*"I was aware of how well our group was doing at all times"*). **H8a** predicted positive correlations, which held significantly for "Simple" participants on Q7 for both the *Proportion* and *Duration* of gaze measures, and for "Complex" participants on Q9 for *Proportion*. (See Table 26 for more detailed data)

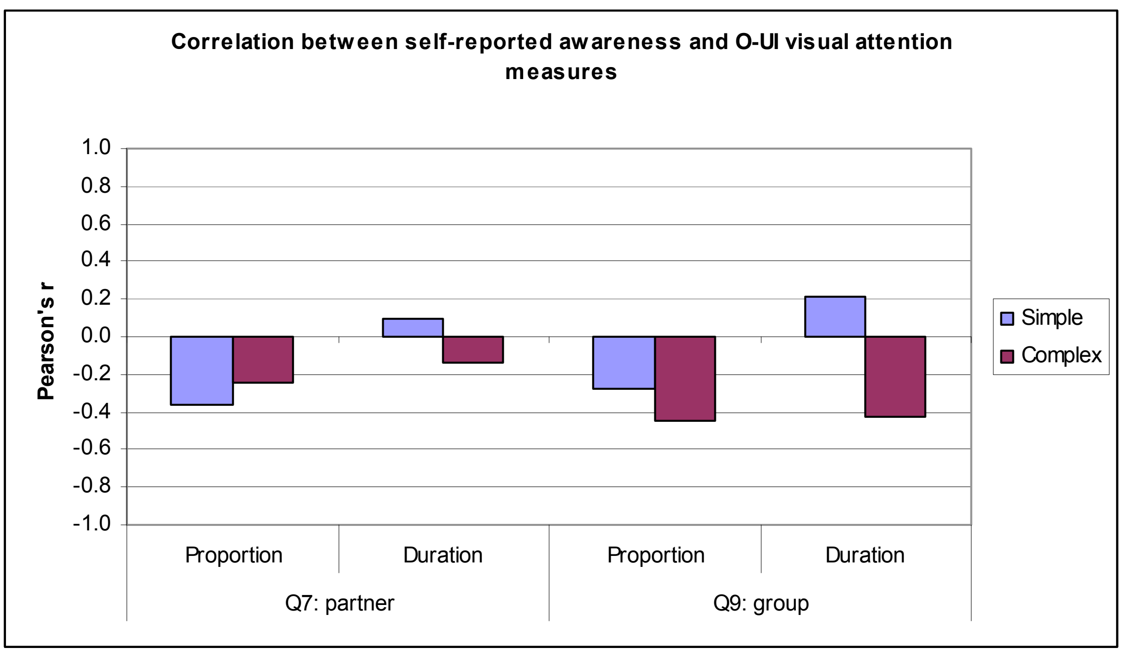


Figure 11. Bar chart depicting the correlation (Pearson's r) between O-UI visual attention measures and the responses to Q7 (*"I was aware of how well my partners were doing at all times"*) and Q9 (*"I was aware of how well our group was doing at all times"*). **H8b** predicted negative correlations, which mostly occurred, aside from "Simple" players and *Duration*. **H8b**'s prediction held significantly for "Complex" participants on Q9 for both the *Proportion* and *Duration* of gaze measures. (See Table 26 for more detailed data)

Conversation Frequency

Rather than relying solely on participant's recall of their awareness, other measures can be used to determine the degree of awareness exhibited by participants. One is to measure the frequency of *Conversation*, under the assumption that participants would be less likely to speak to their companions if their full attention is truly being commandeered by their O-UIs. **H6** predicts that "Complex" participants will demonstrate lower frequencies of *Conversation*, and a quick glance at Figure 12 shows this to be true, $t(30) = 4.32$, $p < 0.00008$, (one-tailed).

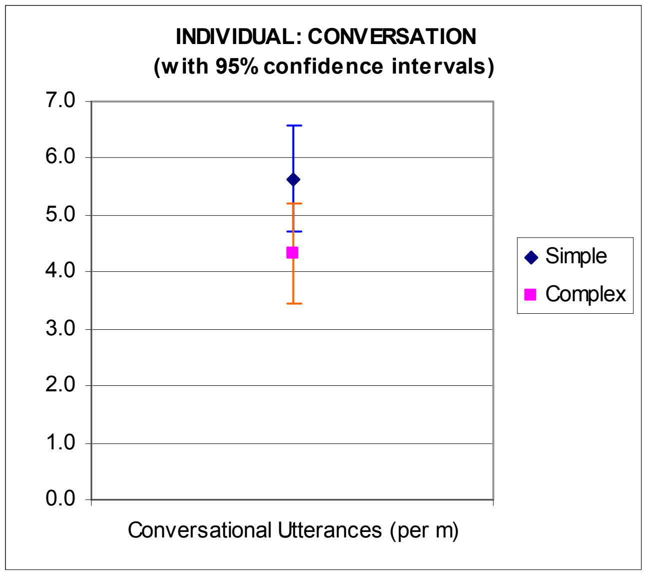


Figure 12. Plot of frequency of *Conversation*, with 95% confidence intervals. The lower level of *Conversation* made by "Complex" participants provides corroborating evidence that they are not attending to their companions as much as "Simple" participants are. (See Table 25 for more detailed data).

The implication is that "Complex" participants are not attending to their companions as much as "Simple" players are, but is this related to their visual gaze behaviors? **H9a** predicts that *Conversation* levels should correlate positively with the visual attention paid to the Shared Display, but this seems to be only mildly true for "Simple" participants, and "Complex" participants' conversations don't seem to be affected one way or another by attention to the Shared Display – their correlation coefficients are very close to zero (see Figure 13). **H9b** predicts that *Conversation* levels should correlate negatively with the visual attention paid to the O-UI, but this is only true for "Simple" participants and the *Proportion* of gaze they devote to O-UIs, and for "Complex" participants and the *Durations* of gaze they devote to the O-UIs. Taken together, there is only mild evidence that visual attention management is affecting the awareness visitors have of their companions – "Complex" participants may lose sight of conversing with their companions when they gaze at their O-UIs for long durations, but the mildness of the correlation doesn't allow for any strong conclusions to be reached.

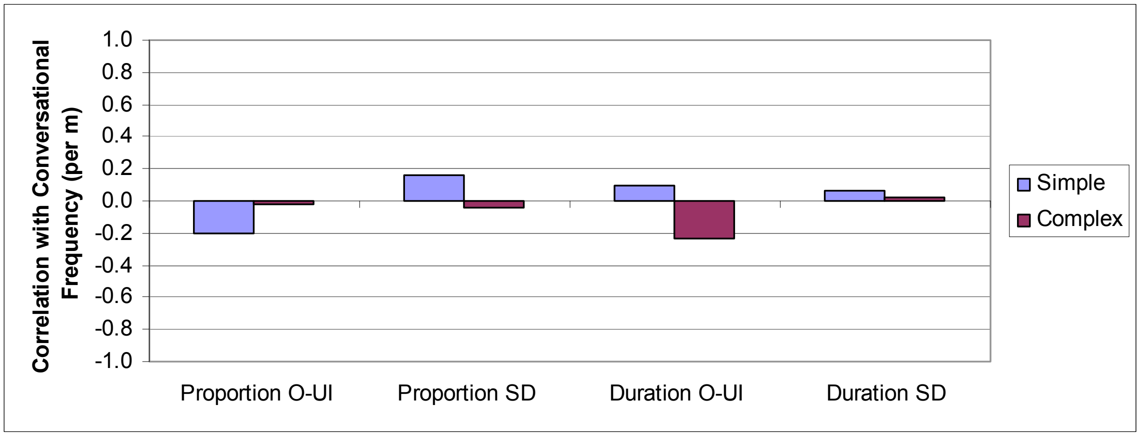


Figure 13. Bar chart depicting the correlation (Pearson's r) between visual attention measures and *Conversational* frequency. **H9a** predicted positive correlations with attention to the Shared Display, and **H9b** predicted negative correlations with attention to the O-UI. (See Table 27 for more detailed data).

Move Frequency

The second more objective measure of participants' awareness of their shared context is to determine the frequency of *Moves* the participants are making. Because participants have to attend to the Shared Display to move their incision rectangle, the Move frequency will tell us whether or not the poor visual attention management behaviors exhibited by "Complex" participants are nonetheless adequate to allow them to participate in the simulation as actively as "Simple" participants. **H7** predicts that "Complex" participants will demonstrate lower frequencies of *Moves*, which is in fact very significantly supported by the evidence, $t(30) = 7.96$, $p < 0.000000003$, (one-tailed) (see Figure 14).

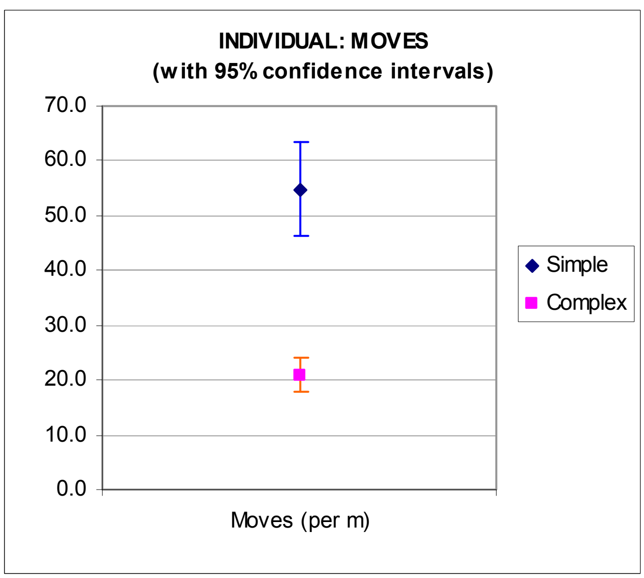


Figure 14. Plot of frequency of *Moves*, with 95% confidence intervals. The much lower number of *Moves* made by "Complex" participants provides corroborating evidence that they are not attending to the Shared Display as much as "Simple" participants are. (See Table 25 for more detailed data).

Once again, though, can the disparity in *Move* frequency be linked to visual attention management behaviors? **H10a** predicts that Moves will correlate positively with attention to the Shared Display. Figure 15 shows that the correlations trend in the direction **H10a** predicted, and they are significant for both conditions for the *Proportion* of gazes devoted to the Shared Display. This disproves any notion that "Complex" participants, despite attending visually more to their O-UIs, were still able to maintain levels of Shared Display awareness equivalent to that of "Simple" participants. The predictions made by **H10b**, that attention to the O-UI would correlate negatively with Moves, also hold.

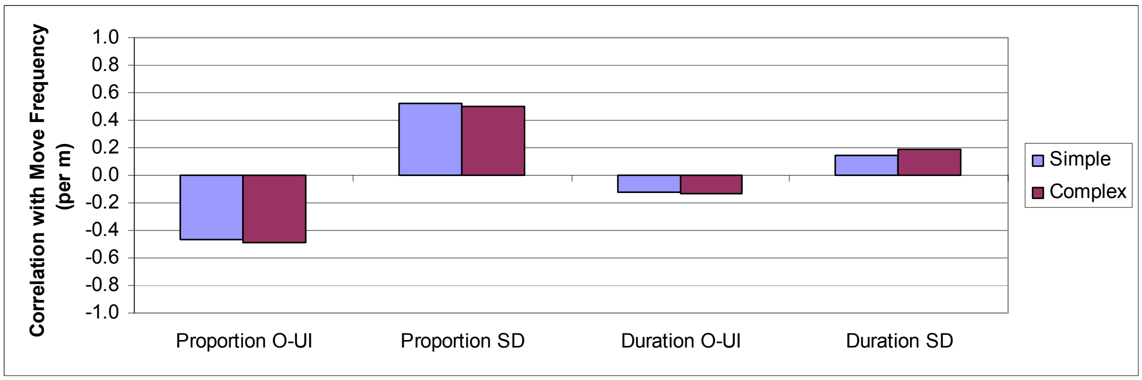


Figure 15. Bar chart depicting the correlation (Pearson's r) between visual attention measures and *Move* frequency. **H10a** predicted positive correlations with attention to the Shared Display, and **H10b** predicted negative correlations with attention to the O-UI. (See Table 28 for more detailed data).

1.2.6.5 Verdict on Research Question 1

Taking only the evidence from Section 1.2.6.2 Visual Attention Management: Results, it is clear that "Complex" O-Us promote externally-visible visual attention management behaviors in line with the described heads-down phenomenon. The predictions of **H1** (gaze *Proportion*), **H2** (gaze *Duration*), and **H4** (gaze *Synchronicity Degree*) all held, demonstrating that "Complex" participants did indeed gaze longer, and for longer unbroken durations, at their O-UIs, and that "Simple" participants were better able to establish joint attention. (**H3** was dismissed from consideration because it relied on a misleading measure, gaze *Frequency*). The next task was to look for secondary symptoms that would demonstrate whether or not the observed poor visual attention behaviors were in fact resulting in poorer awareness of the shared context. Section 1.2.6.4 Awareness of Shared Context: Results showed that while "Complex" participants reported about the same levels of *Awareness* (**H5**) as "Simple" participants, self-reported awareness was improved when participants visually attended more to the Shared Display. More objective measures showed that "Complex" participants were significantly less engaged with elements of the shared context, *Conversing* (**H6**) less with their companions, and interacting less with the Shared Display (as evidenced by *Move* frequency, **H7**). Although *Move* frequency was significantly tied to visual attention management behaviors in both conditions, Conversation did not seem to be for "Complex" participants. Still, the overwhelming weight of evidence confirms that **O-UI "complexity" does produce the heads-down effect**.

1.2.7 Research Question 2: Impact on Support for Collaborative Learning

1.2.7.1 Definition of Support for Collaborative Learning

There is no one definition for what good support of collaborative learning is, just as there is not one definition for what collaborative learning itself is. Collaborative learning will look very different depending on the context it emerges from, and so the measures taken to support collaborative learning under the different contexts will likely differ wildly. For this context, a shared museum exhibit, three broad categories of measures often used by educational researchers, museum researchers and computer-supported collaboration researchers to indicate the success of collaborative learning activities were chosen:

1. **Learners should have an awareness of joint goals**
2. **Learners should interact with one another**
3. **Learners should participate equitably in the activity**

One advantage of computer-based museum exhibits is that visitors' each and every interaction with the exhibit as they attempt to accomplish a task can be recorded. A fair amount can be surmised about a learner's understanding of the shared task goal by looking at his or her actions, especially when those actions have direct, score-able impact on the eventual success or failure of an endeavor, as when visitors eliminate more cancer cells than healthy cells with their surgery actions. For that reason, this experiment is using the following measure as an indicator of *individual*-level understanding of the joint goal:

M1: High *Individual Scores*

In a similar vein, if the groups are able to accomplish the aim of the activity (eliminating cancerous cells from the simulation), this betrays a certain amount of understanding of the joint goal of the activity, so the following measure can be used as an indicator of the *group's* level of understanding of the goal:

M2: Better group *Outcomes*

Finally, because this is a collaborative activity, a measure of whether or not the members of a group understood the *collaborative* goals of the activity can be measured by how successfully they were able to divide the task. The degree to which a group divided the task can be measured by the amount of the Shared Display that was occupied solely by a given individual, i.e., that individual's degree of territorial *Ownership*. Therefore, a measure that speaks to the group's understanding of the collaborative aspects of the shared task is:

M3: High degree of *Ownership*

In the case of this exhibit, the primary means by which learners interact is *Conversation*. The quantity of conversation is one issue (collaborative learning cannot take place if there is *no* interaction) but the *quality* of those interactions is perhaps more important. Collaborative learning cannot take place if the interactions are not *on task*, and so conversational utterances will be coded as being on- or off-task to establish whether or not collaboration is likely to be occurring. Because a major learning goal of this exhibit is for visitors to understand the underlying rules and processes of the simulation, coding the amount of *Tactical* or *Strategic* content in learner conversation can indicate something about the degree to which visitors are trying to come to understand those rules. Finally, the degree of *reciprocity* of communication was found to be a very strong predictor of whether or not the collaborative endeavor would succeed or fail, so the utterances will be coded for their *Interactional* properties. The following three measures, then, will be used to speak towards the quality of learner interactions:

M4: High proportion of *On-task Utterances*

M5: High proportion of *Tactical/Strategic Utterances*

M6: High proportion of *Interactional Utterances*

Finally, we come to one remaining category that is accepted by many collaborative learning researchers as a precondition for "good" collaborative learning: *equity*. Educational researchers have found that when some learners do not participate as much as others in a joint learning activity, they have lower individual learning gains (Cohen, 1994; Kapur & Kinzer, 2007; O'Donnell & Dansereau, 1992; Schellens, Keer, Valcke, & Wever, 2005). Perhaps for this reason, several CSCL researchers use measures of participation equity to judge whether a collaborative system is successful or not (Kapur & Kinzer, 2007; Schwartz, 1999), although the value of measuring participation equity also comes up in CSCW literature (DiMicco, Hollenbach, Pandolfo, & Bender, 2007; Morris, Morris, & Winograd, 2004; Rodden, Rogers, Halloran, & Taylor, 2003). Given that museums have the additional responsibility to serve a wide variety of visitors, supporting equity is especially important:

4. **Learners should participate equitably in the activity**

Specifically, equity in participation (which shows that no visitors are getting left out, from interacting with the activity or from interacting with one another) and equity in performance (which shows that the exhibit's activity is equally accessible to all group members) will be used as measures. A more successful collaborative exhibit should show:

M7: Lower *Participation Inequity*

M8: Lower *Conversation Inequity*

M9: Lower *Performance Inequity*

1.2.7.2 Evidence for Awareness of Joint Goals

Individual Scores: M1

One goal for the exhibit was for participants to come to understand that in order to eliminate cancer from a patient, doctors must be diligent about removing all of the cancerous cells they can, leaving no cells behind no matter how small or damaged, even if it means risking some collateral damage to do so.

By scoring individuals on their individual ability to perform this task, some insight is gained onto whether or not they understood this aspect of the shared task. The main scoring mechanisms used to evaluate **M1**, *Weighted Efficacy,* is structured so that it is more valuable to kill cancer cells than to damage them, and more valuable to kill or damage cancer cells than to kill or damage normal cells. For comparative purposes, though, the results for the *Unweighted Efficacy* score, which just values dead or damaged cells above dead or damaged normal cells, is also provided. Regardless, on both scores, "Complex" players significantly outperform "Simple" players (see Figure 16). On the *Weighted Efficacy* measure, the significance is $t(30) = 2.89$, $p < 0.004$ (two-tailed).. This indicates that "Complex" players seemed to have a better understanding of the cancer-cell removal goal.

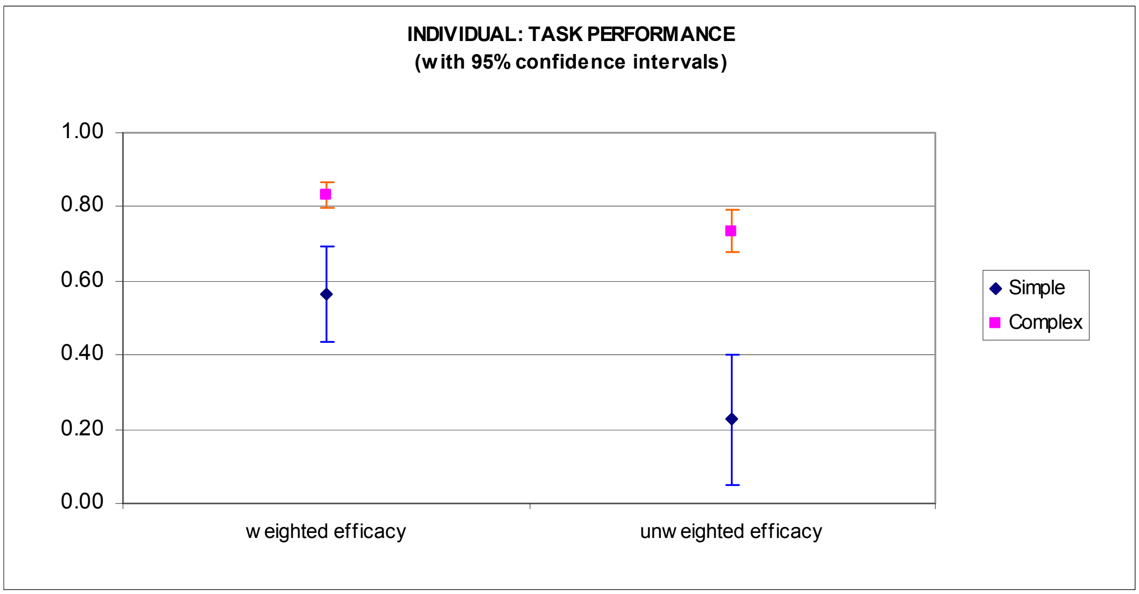


Figure 16. Plot of *Individual Scores*, with 95% confidence intervals. The *Weighted Efficiency* measure permits more collateral damage to healthy cells. (See Table 29 for more detailed data)

Ownership: M2

The guiding idea behind the *Ownership* measure, **M2**, is that in order to effectively collaborate on the joint task, players to avoid overlapping their partners and duplicating their work. Because "Simple" players were shown to visually attend more to the Shared Display, one might expect "Simple" participants to also demonstrate higher levels of *Ownership*. A check of the results, however, shows that the result is quite the opposite (see Figure 17). From this evidence, "Complex" players are significantly better at engaging in task division, $t(30) = 5.59$, $p < 0.000004$ (two-tailed). It seems that "Complex" players are either more aware of the collaborative aspects of the shared task, or more amenable to engaging in them: informal observations of participant behaviors suggest that rather than using the Shared Display to manage task division, "Simple" players were using it to be able to "cluster" or "herd" into the same places.

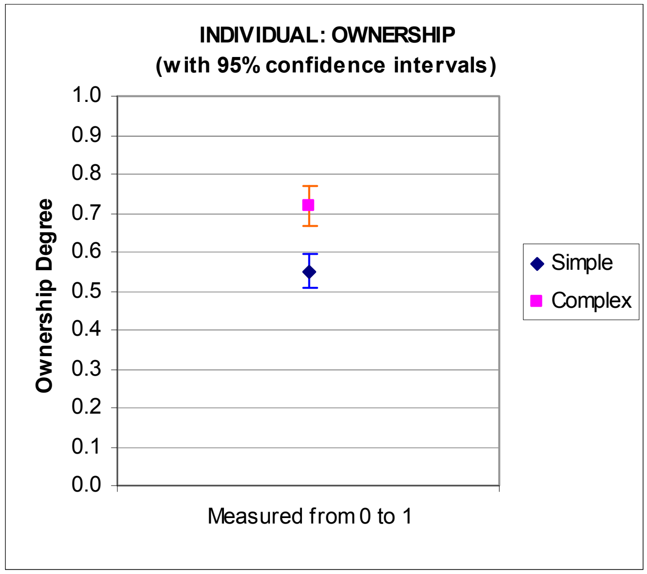


Figure 17. Plot of *Ownership Degree*, with 95% confidence intervals. (See Table 30 for more detailed data).

Outcomes: M3

The other prediction related to goal awareness is the group-level *Outcome* measure, **M3**. In this context, the group can succeed by eliminating the cancer, or fail by causing the patient to die as a result of too much collateral damage or allowing the cancer cells to take over. The *Outcome* is the percentage of episodes that a group participated in that ended in success. While “Simple” participants do see slightly more episodes ending in success (around 88%) than “Complex” users (around 81%) this difference is not significant (see Figure 18). This may be the result of a ceiling effect, however: 73% of “Simple” players and 71% of “Complex” players never lost a single session.

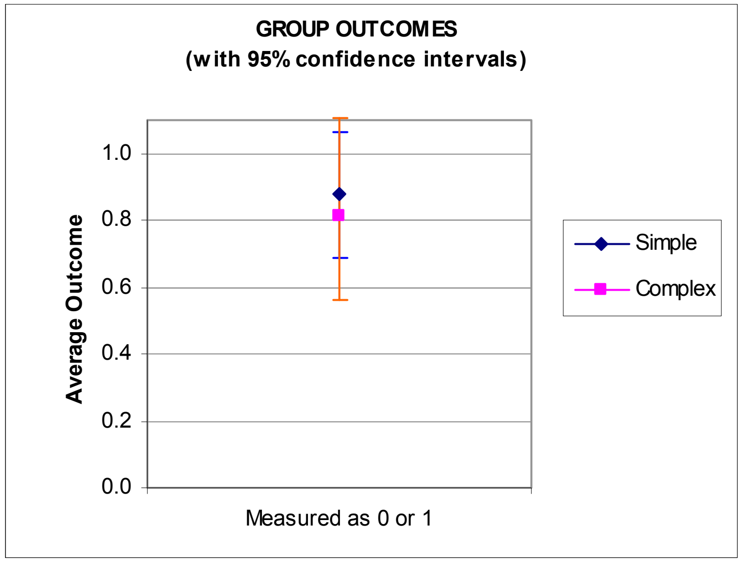


Figure 18. Plot of group *Outcomes*, with 95% confidence intervals. (See Table 31 for more detailed data).

Summary: Awareness of Joint Goals

Overall, "Complex" participants demonstrated higher awareness of joint goals, both by scoring significantly higher on individual measures of task performance (**M1**) and by demonstrating, if not more awareness of, at least adherence to, the collaborative aspects of the shared task (**M2**). The measure of group goal awareness, Outcome (**M3**), was essentially the same for both conditions, although a ceiling effect may have been involved.

1.2.7.3 Evidence for Interaction between Visitors

Although no predictions were made for the impact of O-UI "complexity" on the different coded utterance proportions, the results show that "Complex" participants demonstrate significantly higher *Proportions* of *On-task* (**M4**) and *Functional* (**M5**) utterances, while the proportion of *Interactional* (**M6**) utterances remains virtually the same (see Figure 19). Recall that the comparison being made is a paired t-test, meaning that the utterances made by one participant in one condition are compared against the utterances made by that same participant in the other condition, so the O-UI clearly has a strong impact on a person's *On-task* and *Functional* remarks. Although "Complex" users on the whole spoke less than "Simple" users, as established Section 1.2.6.4 Awareness of Shared Context: Results, nearly half (M = 0.49, SD = 0.17) of "Complex" participants' utterances were considered to be *On-task* (**M4**), compared to a little better than one third of the remarks made by "Simple" participants (M = 0.36, SD = 0.18), $t(30)$ = 4.04, $p < 0.0004$ (two-tailed). Likewise, the difference in Functional (i.e., Tactical/Strategic, **M5**) was strongly significant, $t(30) = 2.79$, $p < 0.0092$ (two-tailed).

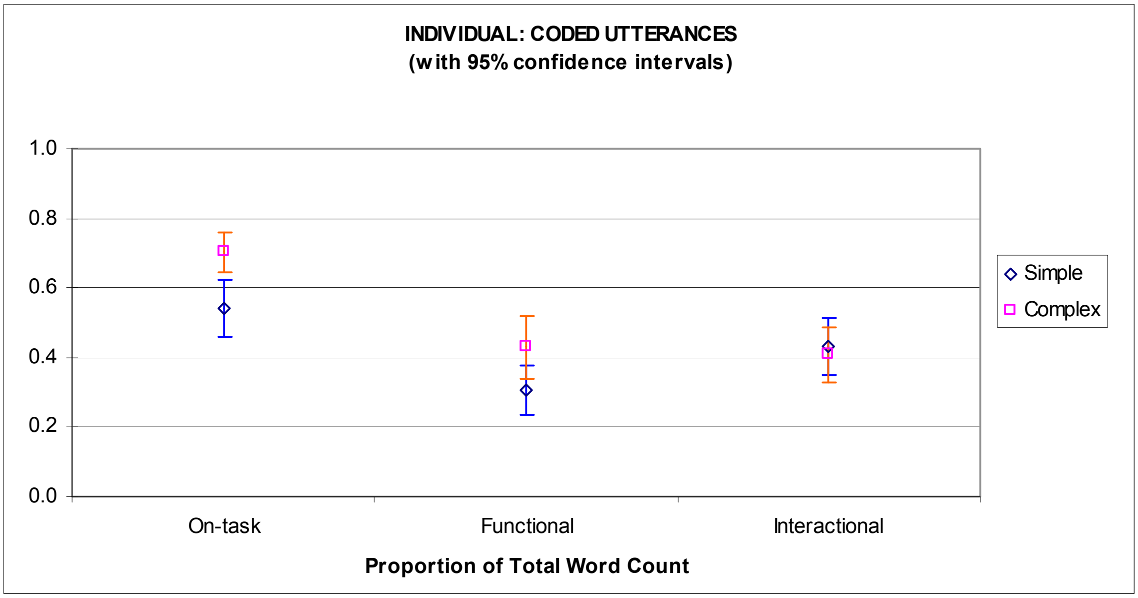


Figure 19. Plot of the proportions of *On-task*, *Functional* (i.e., *Tactical/Strategic*) and *Interactional* utterances, with 95% confidence intervals. (See Table 33 for more detailed data)..

1.2.7.4 Evidence for Equity

Inequity measures for *Participation* (**M7**), *Conversation* (**M8**) and *Performance* (**M9**) are all measured by taking the standard deviation across all of the group's members on a particular measure. The standard deviation indicates how unalike the group members are on that particular measure.

Participation Inequity: M7

Starting with *Participation* (**M7**), there are three different measures that Inequity can be computer for: the number of *Moves*, the number of *Damage* actions initiated by the participant, and the *Total Input* – a sum of the first two measures. A look at Figure 20 suggests that the difference in *Total Input Inequity* mostly derives from the *Inequity* of *Moves*. “Simple” participants show much more within group variation on *Moves*.

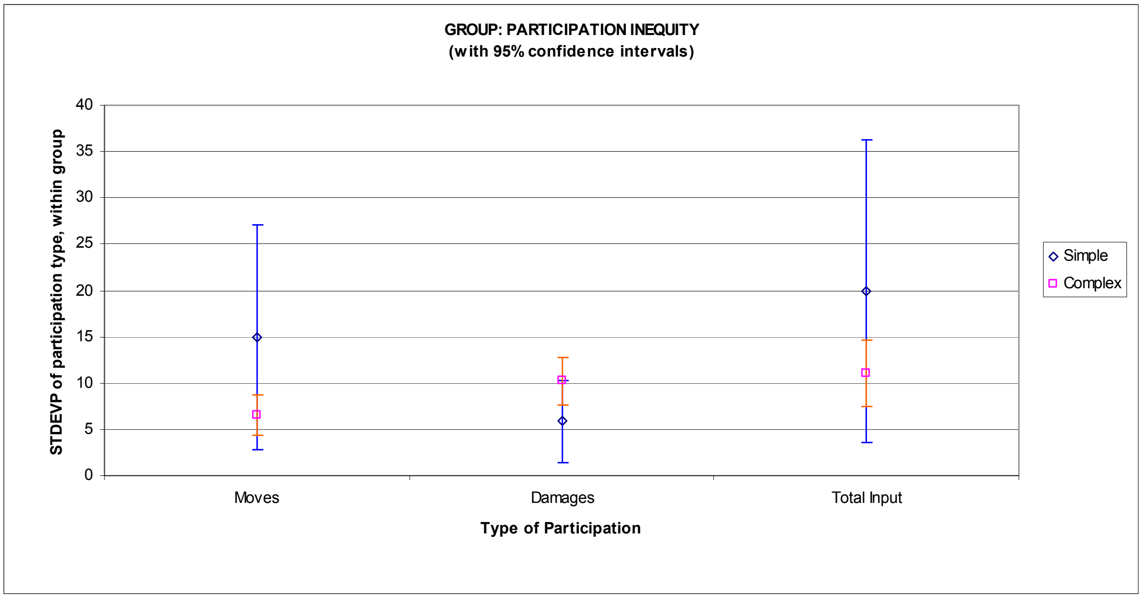


Figure 20. Plot of *Participation Inequity* (**M7**), for the *Total Input*, as well as the component measures *Moves* and *Damages*. Inequity is calculated by taking the within-group population standard deviation. (See Table 34 for more detailed data).

A clue for why "Simple" participants would show higher *Move Inequity* can be found in the following conversational exchange. It takes place between two older female participants (East and South) after switching from the "Complex" condition to the "Simple" condition:

West:	"More like a, more like a classic video game"
East:	"Yeah, this is more like a teenage boy thing cause you gotta have that eye - uh"
South:	"Yeah, pow pow pow" <Laughs>
East:	"But you gotta move around really fast, you know what I mean? They're good at that."
South	"Yeah I know."
East:	"I'm not as good at it."
East:	[addressed to West] "Ok, you got em! You guys could do this game by yourself, you could do…" …
East:	"You guys have that dexterity thing."
South:	<laughs>

What the transcript doesn't capture is that after East's next-to-last remark ("You guys could do this game by yourself…"), she put her handheld down and stopped interacting with the simulation, South soon following suit, allowing the younger group members to play on their own. This sort of behavior (a female participant putting the handheld down) was seen in several "Simple" sessions, but never in a "Complex" session. Further investigation demonstrated that gender was indeed a modifying factor on

Participation, but only for the "Simple" condition – the difference in *Participation* disappeared during the "Complex" condition. Keeping in mind that these are paired comparisons – the same participants in each condition, in the same group configurations – the gender-based difference is especially striking, and troubling. Museums have a mission to serve all visitors, so using an interface that encourages participation in one group, but discourages it in another, is a problem.

Conversation Inequity: M8

Unlike *Participation Inequity*, *Conversation Inequity* (**M8**) is nearly identical for the two conditions. The interpretation to make of Figure 21 is that within-group conversational patters hold true for all members of that group – if some members are chatty, all members are chatty, if one member tends to make a certain proportion of *On-task* remarks, all members make similar proportions of *On-task* remarks, and so on. It does not seem to be the case that O-UI "complexity" affects intra-group conversational equity.

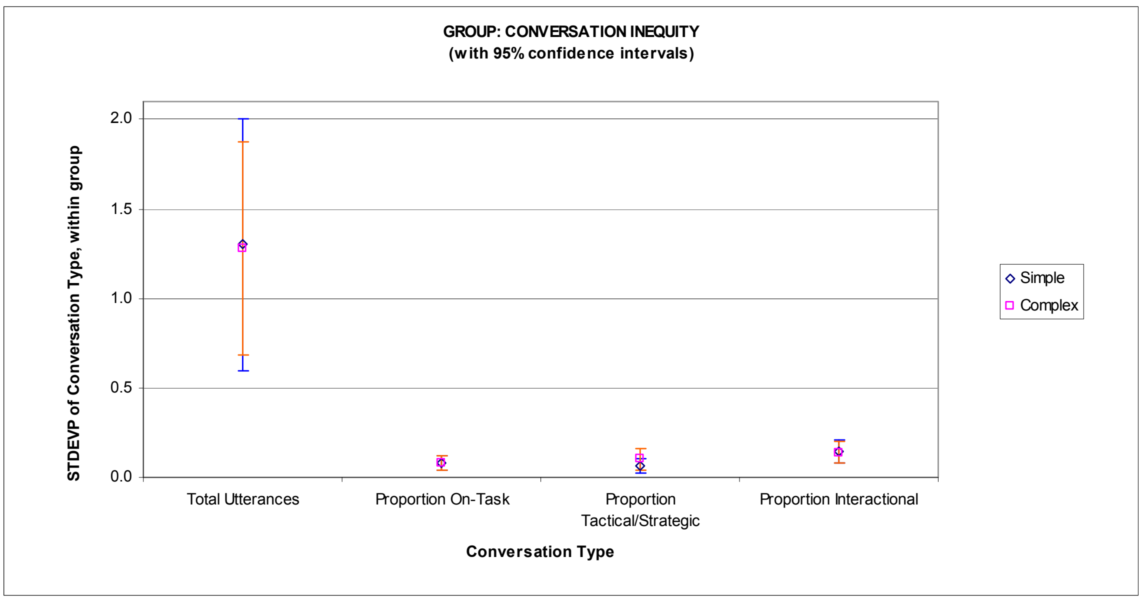


Figure 21.Plot of *Conversation Inequity* (**M8**), for the *Total Utterances*, as well as the proportional measures for *On-task*, *Tactical/Strategic*, and *Interactional* utterances. Inequity is calculated by taking the within-group population standard deviation. (See Table 34 for more detailed data).

Performance Inequity: M9

Performance Inequity (**M9**) can be assessed for *Individual Score* (both *Weighted* and *Unweighted*) and *Ownership* (see Figure 22). With respect to *Ownership*, it seems that the members within each group behave very similar to one another, and there is no effect related to O-UI "complexity." The "Simple" participants do show a greater amount of *Inequity* than "Complex" participants on the two *Efficacy* measures, and this difference is significant for the *Adjusted Weighted Efficacy*. The interpretation, then, is that the "Complex" O-UI condition encourages participants to "operate" somewhat more similarly to one another. Combined with the superior *Task Performance* exhibited by "Complex" participants, as described in Section 7.2.2.2 Evidence for Awareness of Joint Goals, one might conclude that the "Complex" O-UI supports both better and more equitable task execution.

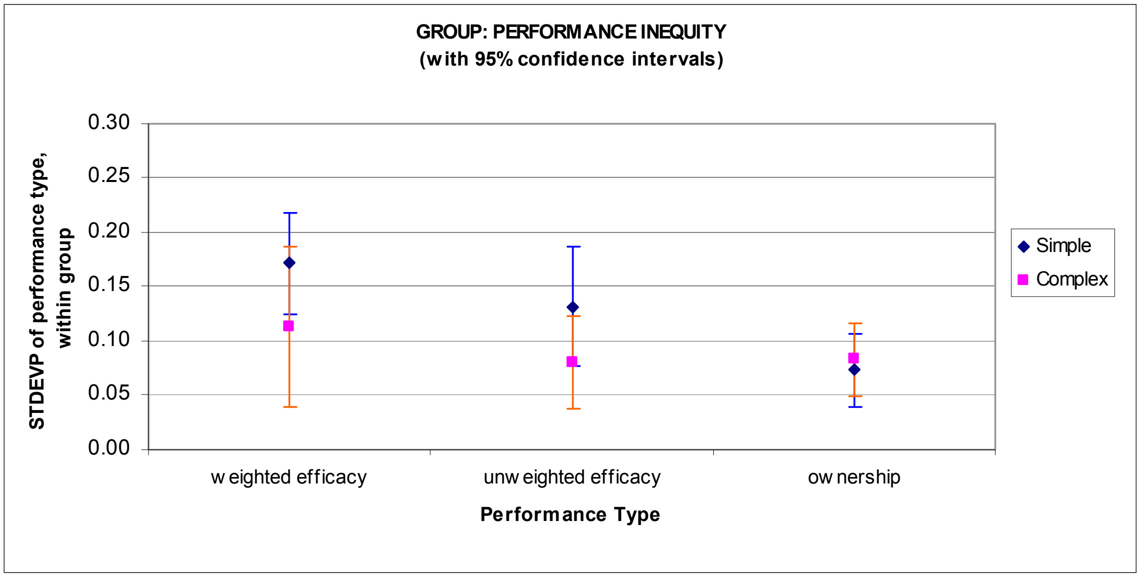


Figure 22. Plot of *Performance Inequity* (**M9**), for the *Weighted Efficacy*, *Unweighted Efficacy*, and *Ownership* measures. Inequity is calculated by taking the within-group population standard deviation. (See Table 36 for more detailed data).

1.2.8 Discussion

1.2.8.1 Weighing in on O-UI Complexity for Museum Exhibits

The "Complex" O-UIs promoted poor visual attention management, an effect known as the heads-down phenomenon, wherein visitors get so enmeshed with their O-UIs that they miss out on the shared context. Despite this shortcoming, the evidence from the preceding section implies that, even though the "Complex" participants suffered more from the heads-down effect, they went on to perform equivalently or better on each of the nine measures chosen to speak to the potential for the activity to support collaborative learning. The "Complex" participants performed significantly better on **M1** (Individual Scores), **M2** (Ownership), **M4** (On-task Utterances), **M5** (Tactical/Strategic Utterances), **M7** (on Moves component of Participation Inequity), **M9** (on Weighted Efficacy measure of Performance Inequity). "Complex" participants performed at about the same level on **M3** (Group Outcomes), M6 (Interactional Utterances), **M7** (on Damages component of Participation Inequity), **M8** (Conversation Inequity), and **M9** (on the Unweighted Efficacy and Ownership measures of Performance Inequity). Additionally, the "Simple" condition promoted more *Participation Inequity*, especially between participants of different genders.

Design Recommendation 1: O-UI-s for Museum Exhibits

The "Complex" condition was designed to embody an O-UI near the "worst case" end of the spectrum, in terms of demanding visitor attention, so it is probably safe to assume that making use of less "Complex" O-UI displays will not *increase* the heads-down phenomenon. That said, the heads-down phenomenon *does* negatively impact "Complex" participants' shared *Outcomes* (see Figure 23), whereas more heads-up visual attention behaviors (higher *Proportion* and *Duration* of gazes devoted to the Shared Display, and higher gaze *Synchronicity Degree*). For this reason, if designers choose to use a "Complex" O-UI, they may wish to experiment with mechanisms to help remind participants to attend to the Shared Display to help them maximize their performance on shared tasks. This forms the crux of the first **Design Recommendation** to emerge from this work:

DR1: *If designing O-UIs for museum exhibits, one can use "complex" output, but should incorporate mechanisms to remind or encourage users to direct attention to shared display periodically.*

The means by which a designer should go about directing O-UI users' attention is still very much an open question, which will be returned to in the discussion of future work.

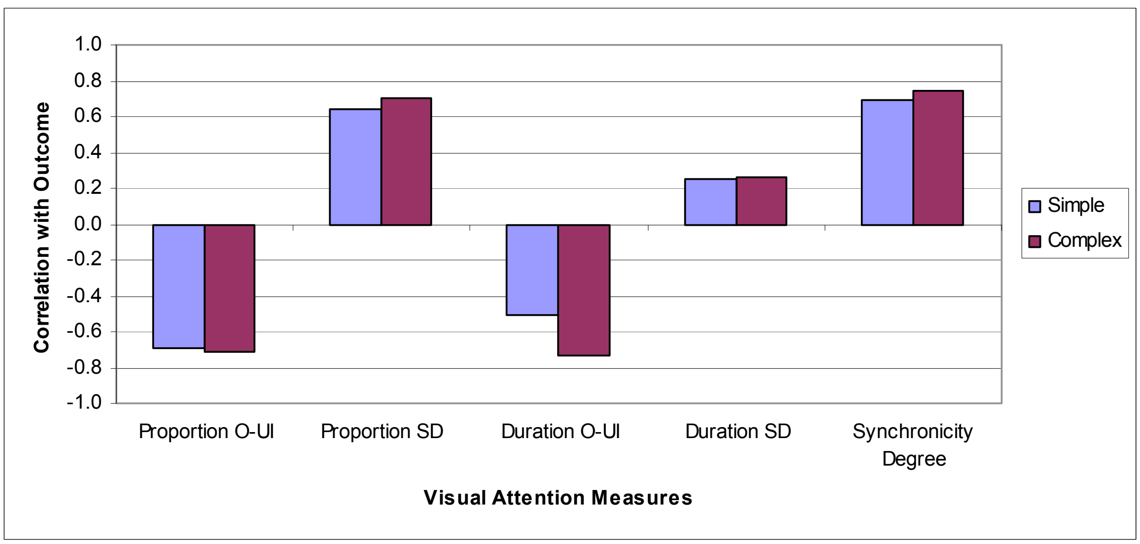


Figure 23. Plot of the correlation (Pearson r) of the group *Outcome* with group visual attention measures.

1.2.8.2 The "Simple" Condition and Competition

Although O-UI "complexity" does lead to attention management behaviors consistent with the predictions of museum professionals, it also seems better suited for supporting collaborative learning – an endeavor which, by all rights, should be reliant on good attention management behaviors. How can this be? The answer may lie, paradoxically, with the visual attention paid to the Shared Display. Participants in the "Simple" condition were observed to "herd" around the same locations on the Shared Display, which is why they scored significantly lower than "Complex" users on the *Ownership* measure of task division. It seemed that they would engage in a sort of emergent competition if they spent too much time monitoring one another's actions on the Shared Display – racing to be the first to get to the new metastasized tumor, and committing surgery actions with haste and a lack of care.

Comparisons of how different measures (*Ownership*, *Moves*, and different types of *Damage* actions) correlated with different degrees of *Gaze Synchronicity* show that "Simple" users did in fact engage in more *Moves*, display less *Ownership*, and get more careless as the degree of *Gaze Synchronicity* increased. "Simple" participants seemed to be affected by the amount of monitoring they were engaged in (and/or the degree to which they were being monitored by others), whereas "Complex" participants were not.

Increased monitoring clearly is associated with emergent competition in the "Simple" condition, but nor the "Complex" condition. Relating this finding back to larger themes found in Computer-Supported Collaborative Work (CSCW), and Human-Computer Interaction (HCI), the effect of supplying a "simpler" O-UI is the same as creating a tighter *coupling* between the users' outputs. As the O-UI is made "simpler" – here, by replacing dynamic visual elements with static elements – the users devote more and more of their visual attention to the shared display. Thus, as O-UI designs become "simpler," the overall experience begins to resemble that of a WYSIWIS (What You See Is What I See) system. WYSIWIS systems exhibit the tightest possible output coupling, wherein users are privy to the exact same output.

Researchers working in CSCW have long posited that tighter coupling of input and output encourages – enforces, even – collaboration. This research, though, shows that tight output coupling is not enough to ensure, or even promote, collaboration – despite experiencing a more WYSIWIS arrangement of output, "Simple" users actually performed significantly more poorly on measures of collaboration like task division, and even displayed non-collaborative, competitive behaviors.

Design Recommendation 2: Coupling and Multi-User Shared Displays

Other computer-supported collaboration researchers have found that Single-Display Groupware (which the "Simple" condition is an exemplar of) can cause competition (Benford et al., 2000; Birnholtz, Grossman, Mak, & Balakrishnan, 2007). These researchers responded by tightening the input coupling so as to further enforce collaboration, to mixed results – users preferred the freedom of loosely-coupled inputs. Because the only difference between the conditions in this research was that the "Simple" condition had tightly-coupled output, and the "Complex" condition's output was loosely coupled, so this work demonstrates that there is a third approach to ameliorating emergent competition: *loosening the output coupling*, the second **Design Recommendation** to emerge from this work:.

> **DR2:** *If designing for a synchronous, co-located collaborative activity that is centered around a shared display, to avoid emergent competition between group members, loosely-couple the output.*

1.2.9 Future Work

1.2.9.1 O-UI Visual Attention Management Mechanisms

Standard prompting "stick" approaches (e.g., pop-up prompts, passive indicators) have not met with much success in other multi-device applications. It may be that an incentivized "carrot" approach would be more successful, like information fission, when critical information is distributed across devices. If visitors need to inspect the shared display in order to complete their individual activities, they are likely to look at the shared display more often, and are also likely to inadvertently glean an awareness of their partner's actions as well.

1.2.9.2 Gender and Output Coupling

The "private" workspace offered by the loosely-coupled "Complex" O-UIs may have encouraged participation equity by increasing participation among group members who would otherwise shy away from "performing" in a public space. This can be confirmed by manipulating the degree of "privacy" in a MMUI setup, from keeping participants' contributions totally anonymous on one end of the spectrum, to making the actions (and the actor behind them) very prominently and publicly indicated on the public display.

1.2.9.3 Strategic Input Coupling

This work varied the degree of output coupling, but did not vary the input coupling. When the term "tight input coupling" is used, the assumption is that it is a low-level, operational coupling of input, as when two users are needed to press a key simultaneously. This sort of input coupling takes place at the *tactical* level. Collaborative learning researchers often make use of activity structures, like jigsawing, which assigns participants to play unique roles or hold unique abilities. Although it has not yet been labeled as such, this is really input *coupling* at a strategic level. A review of prior CSCL work using the "strategic input coupling" lens, suggests that tight input coupling at the *strategic* level may be a better way to encourage collaboration amongst SDG users than tight coupling at the tactical level.

The original design for the simulation used in these experiments called for visitor to take on different, complementary roles, and prior to leaving the Exploratorium, over 60 visitors participated in a formative Phase III of this study, which employed the multi-role version of the simulation. The data analysis for Phase III, which involved audio transcripts and video coding similar to that of Phase II, is ongoing, but should shed some light on the concept of strategic input coupling.

1.3 Dissertation Structure

This section will sketch out a brief overview of the dissertation, in a chapter-by-chapter manner, following the order in which the chapters are presented.

1.3.1 Context of Use: Museums

This dissertation opens with a rich description of the context in which the software will be used, with Chapter 2, the Context of Use: Museums. Design-Based Research stresses the importance of having a thorough understanding of the *context* of a proposed educational intervention, as one of the main drivers behind the inception of DBR was to create research that had strong *external* validity. Thus, Chapter 2 illustrates the historical-cultural, social, and physical aspects of the museum context that can affect the design of computer-based exhibits intended for science centers, drawing from theoretical and practical writings from the body of museum studies literature. The end of the chapter presents a series of recommendations for designing computer-based exhibits that emerge from this literature review, many of which were used to inform the design presented in Chapter 5.

1.4.2 Related Work: Computing Technology in Museums

Chapter 3 reviews the existing research that has been done on computer-based exhibits in museums, to get a sense of the different form factors that have been employed and whether or not they enabled collaborative use by groups of visitors. "Enabling collaboration" is examined here in terms of how well an exhibit design provided adequate *Access* for a group of visitors. Chapter 3's review attempts to suss out how a visitor group's *Access to Input Opportunities*, *Access to Output Opportunities*, and *Access to Companions* were or were not supported by the different exhibit form factors. The different form factors considered were single-user kiosks (Section 3.2.1), multi-user kiosks (Section 3.2.2.), single-user handheld device applications (Section 3.3.1), multi-user handheld device applications (Section 3.3.2), and Multi-Machine User Interfaces (MMUIs, Section 3.4). As with Chapter 2, the end of this chapter concludes with a summary of design guidelines for computer-based museum exhibits that emerged from the literature review, many of which were applied in the design of the activity (see Chapter 5). Something else to emerge from Chapter 3's review was a potential problem for the planned use of O-UIs in museums: the heads-down phenomenon.

1.4.3 Related Work: Supporting Collaborative Activities with Computer Technology

Chapter 4, the second Related Work chapter, pursued the issue of computer support for small-group, co-located collaboration by reviewing relevant research from the Computer-Supported Collaborative Work (CSCW) and Computer-Supported Collaborative Learning (CSCL) fields. The review starts with some of the theoretical constructs that underlie CSCW and CSCL research, like the concept of coupling. It then surveys different form factors used in CSCW and CSCL literature to support small-group, co-located collaborative activities, like Single-Display Groupware (SDG), Multi-Machine User Interfaces (MMUI), and networked handheld computers..

1.4.4 Design Rationale and Implementation

Chapter 5 describes the design of the computer-based museum exhibit that used as a testbed for this research. The idea is that museums could supply a single, shared display (as with Single Display Groupware), and visitors will be able to join the shared activity through their own personal computational devices: their O-UIs. The speculative design recommendations from Chapter 2 and the slightly more vetted design strategies from Chapter 3 informed the design of the exhibit, although not all of these recommendations and strategies are being experimentally evaluated by this research study. (The intent behind adopting various recommendations and strategies was to create software that would have as much external validity as possible, before setting up a more traditional experiment with high internal validity). Chapter 5 also presents a use scenario for the software, so readers can get a feel for how the software would be used in an ideal case.

1.4.5 Experimental Design and Methods

Chapter 6 provides a more specific definition of the research question and the independent and dependent variables of the experimental study. This research actually had two phases – a formative phase that refined the O-UI designs, and an experimental phase that pitted these refined designs, a "Simple" O-UI design and a "Complex" O-UI design, against one another. This multi-phase approach is typical of Design-Based Research, which is described in greater depth in Section 6.3. The methods used in both of these phases, and the settings for both, are also described in this chapter.

The first experimental research question addresses the heads-down phenomenon – can it occur in a MMUI-based museum exhibit? Does varying the "complexity" of O-UI design impact the degree of the heads-down phenomenon? Does the use of "complex" O-UIs negatively impact the ability of the exhibit to support collaborative learning processes? To assess these questions, dependent variables like division of attention, degree of engagement, and task performance are examined at two levels of granularity: the level of the individual, and the level of the group (see Section 6.2 for a more detailed explanation of the independent and dependent variables).

1.4.6 Results

Chapter 7 presents both the results from the second, experimental phase of this research, as well as two case studies drawn from the first, formative phase of this research. These case studies are presented not just to give the reader an idea of the type of work that was conducted during the formative phase, although they serve this purpose. These two case studies were selected in particular to address any questions the reader might have about the equivalence of the design of the "Simple" and "Complex" O-UIs. The desire was to construct two user interface designs that, although differing on their degree of "complexity," would impact the shared simulation in as similar a manner as possible, so that any differences were a result of true differences in usage, and not artifacts of the designs.

The second, experimental phase looked for two things: (1) whether or not the "complexity" of O-UI design impacted the severity of the heads-down phenomenon, and (2) whether or not "complex' O-UIs negatively impact a group's ability to engage in collaborative learning activities. The heads-down phenomenon was only qualitatively defined in prior research; to convert it into a quantitatively measurable phenomenon, ten hypotheses concerning the phenomenon's impact on measurable dependent variables were generated (see Section 7.2.1 Establishing the Existence of the Heads-Down Effect for a concise listing). Similarly, nine measures were selected to represent the what good collaborative learning at a museum exhibit would look like (see Section 7.2.2 Impact on Potential for Collaborative Learning for a concise listing).

1.4.7 Discussion and Future Work

The final chapter reviews the experimental results and generates two design recommendations, one for O-UI design, and one for synchronous, co-located collaborative software. Several future work directions are proposed.

CHAPTER 2

Context of Use: Museums

The software used in this research employs a Multi-Machine User Interface (MMUI) (Brad A. Myers, 2001), which means that the software is distributed across multiple different computational devices. These devices, in turn, may be used by multiple different users. The design of MMUIs, like the design of other ubiquitous computing systems, requires a broader understanding of the context of use than, say, the average single-user word processing program intended for use in a typical office setting (Dourish, 1995, 2004; Kray, Wasinger, & Kortuem, 2004). "Context" is a term with many implications, however, because it is essentially a blanket term for whatever intellectual framing one decides to give a scenario. If one is of a practical bent, one might be concerned with more tangible aspects of an application's context of use, like the technological infrastructure or the physical layout (Raptis, Tselios, Tselios, & Avouris, 2005). Others might be more concerned with the social or cultural contexts that help shape the way users interact with computing systems (Huh, Ackerman, Erickson, Harrison, & Sengers, 2007; Räsänen & Nyce, 2006).

The aim of this body of research is to produce and refine a multi-device educational intervention, while simultaneously expanding what it known about HCI features that can influence small-group learning in a museum setting. A Design-Based Research (DBR) methodology has been adopted to attain these aims, and as such, the research is in turn tightly bound to the *physical*, *social*, and *historical-cultural* contexts that the educational intervention is implemented within. The reasons why these three contexts are significant to Design-Based Research will be discussed in more detail in Section 6.3 of the Experimental Design and Methods Chapter. For now, one just needs to know that the purpose of this chapter is to acquaint the reader with the following three features of informal learning environments (specifically, science museums): (1) the *historical -cultural* context, because the legacies behind museum establishments affect their learning environments to this day, (2) the *social* context that can be found in typical science museums, as the composition and interaction of visitor groups greatly impact the styles of learning present in museums, and (3) the *physical* context, because it is important to understand the types of interactions already present in an environment before designing new interactions.

2.1 The Historical-Cultural Context for Learning in Museums

It is important to understand that, although they seem like stalwart cultural constants, museums can and do adapt to the changing needs of their audiences and of society at large. Without a conscious awareness of the evolving theories underlying exhibit design over the decades, designers run the risk of unintentionally aping conventions that do not serve their educational ends (Roberts, 1997), informed by culturally-received assumptions about "how museum exhibits should be." The pressure for exhibit designers working in new media (computer-based exhibits, for example) is especially high. Designers must be stalwart about *not* disregarding the unique learning affordances provided by a new media in an attempt to make their new media creations conform to old expectations for how a museum exhibit should look or behave[2]. Conversely, nor should a designer throw out the proverbial baby with the bathwater – museum professionals have, quite literally, had hundreds of years to experiment with what "works" in their context, and while technology may change quickly, the visitors themselves do not. As this section will illustrate, modern-day visitor behavior are often influenced by decades-old patterns of museum-going behavior.

This section will briefly trace through the history of changing attitudes towards learning in museums, to help lay bare the rationale behind certain exhibit design approaches (e.g., the reasons behind the endless rows of taxidermied flightless birds behind glass). The purpose of this history lesson isn't just for the reader's edification; museums, unlike classrooms, primarily consist of physical installations (exhibits) which persist long after the intellectual vogue that spawned them has passed on. The side effect is that their influence on the expectations of both learners and museum professionals can also persist in surprisingly tenacious manners.

By way of example, the early 20th century sculptures of native peoples in natural history museums persisted long after native groups began complaining that such presentations (especially when placed in the same building as stuffed African animals on the veldt) perpetuated out-of-date stereotypes and equated native peoples to exotic animal specimens (MacDonald & Alsford, 1995). And yet, the revamping and removal of these sculptures caused a tremendous amount of protest, from visitors and museum professionals alike, who felt that the dioramas were carrying out a legitimate educational mission[3] (Boyd, 1999). Many museums that have tried to alter the ways in which they structure exhibits have met with fierce resistance from critics and members of the general populace, who almost seem to take more umbrage at the idea of change itself than the actual changes being undertaken at museums (Miles, 2007; Traub, 2004).One doesn't need to weigh in on either side of this debate to acknowledge that expectations can be

---

[2] For example, Section 3.2.1 Single-User Kiosks describes how, in the earliest days of single-user kiosks in museums, designers would employ the kiosks to present text-based information in a serial, "click-next" manner. Essentially these kiosks were just digitally-displayed labels, differing from the printed labels mounted on the wall only by the fact that a much larger amount of text could be presented on the kiosk. It took designers a while to truly take advantage of the random-access nature of computers, and offer visitors more choice and control.

[3] The emotional invective found in these protests and counter-protests cannot be overstated – protestors use phrases like "subhuman chattel" to describe the diorama presentations, and counter-protestors inveigh against how "anti-intellectual moral relativism" is destroying our collective critical faculties.

powerful in shaping the judgments of visitors and museum professionals alike. Throughout this section, then, certain powerful themes (or memes) will be emphasized, and their relevance to the design of computer-based exhibits will be described.

2.1.1 Authenticity and the Object-Based Epistemology

Authenticity: *Traditionally, in museums, a characteristic that typically can only be possessed by physical objects or places, and encompasses some notion of "reality" (e.g., a "real sauropod femur"), or originality (e.g., "the original* Mona Lisa*"), and tends to promote an intangible feeling of awe or respect*

The Object: *Traditionally, in museums, a material item that has earned a reified status owing to its singularity or authenticity*

Museums have existed in one form or another since Greco-Roman times, but were primarily the private collections of wealthy individuals, like the *wunderkammen*[4] of Europe (Impey & MacGregor, 2001).This attitude changed in the enlightenment era of the 1700s and early 1800s, where museums were the sites of serious scholarly activities (like systematic ethnographic, botanic, and paleontological categorization). Because these collections primarily served as resources for researchers, very little effort was spent making them accessible or understandable to the general public, even after the bulk of active research relocated from museum halls into universities in the late 1800s (Conn, 1998).

Initially, museums and their curators believing that merely viewing the objects in their hierarchically-organized museum collections (see Figure 24) would bring visitors to the same level of understanding as the scholars arranging the collections (Conn, 1998). This authoritative, object-based epistemology places an inordinate amount of importance on bringing learners into contact with "real" or "authentic" objects, and carries with it an implicit desire, on the part of curators, to present these objects in a scientifically-sanctioned order or hierarchy. The design of science museum exhibits is affected by these ideas to this day, even though visitor research demonstrates that presenting material in this object-centric manner is not necessarily an effective strategy for helping visitors learn (Roberts, 1997).

[4] Literally, "cabinet of wonders" – a collection of odds and ends that struck the fancy of the (wealthy or royal) owner/collector .

Figure 24. Engraving of the Great Zoological Gallery of the British Museum (The British Museum, n.d.), showing the standard hierarchical/categorical organization style that dominated museums of the later 1800s.

2.1.1.1 Visitor Expectations for Objects and Museums

Regardless of what is “best” for their learning, a large proportion of visitors nonetheless *expect* to encounter a certain amount of “real” objects in science museums, and tend to respond negatively when museums break with traditional object-based presentations (Miles, 2007). For example, one group of researchers studying visitor responses to computer-based exhibits found that older visitors would avoid computer-based exhibits that had no objects, but would happily make use of computer-based exhibits when they were presented in conjunction with objects (Hornecker & Stifter, 2006a, 2006b; Moritsch, Hornecker, & Stifter, 2004). An interpretation of this contradictory behavior towards computers is that the visitors weren’t technophobes, but rather had an embedded expectation that exhibits must be structured around objects. This interpretation is supported by evidence from a large survey of science museum visitors that found that older visitors were significantly more likely than younger visitors to select “objects and artifacts” as a preferred interpretive strategy[5] (Korn, 1995). Such expectations were most likely instilled in visitors during their own childhood visits decades past, but unfortunately do not necessarily mesh well with many of the strategies modern researchers propose to help visitors learn from museums. Sometimes this conflict results in the abject failure of exhibitions, as the next example will illustrate.

[5] Not surprisingly, in the same study, younger visitors were significantly more likely to prefer “computer games” as an interpretive strategy as compared to older visitors.

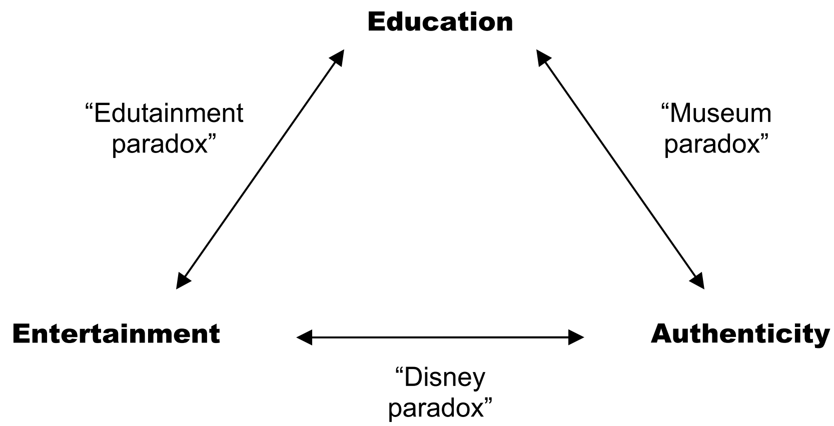


Figure 25. The three "paradoxes" faced by museum exhibit designers who wish their creations to be simultaneously educational, entertaining, and authentic, adapted from (Wolf, Lee, & Borchers, 2007).

The Brooklyn Children's Museum is the oldest children's museum in the world, and was home to a large collection of natural and cultural objects. In the wake of the success of institutions like the Exploratorium in San Francisco and the Boston Children's Museum, the Brooklyn Children's Museum decided to put most of its object collection in storage in order to present an array of object-free interactive exhibits based around more modern theories of learning. The experiment drew protest from visitors accustomed to the prior object-centered approach, and eventually fell into disuse and disrepair (McLean, 1999). Essentially, the Brooklyn Children's Museum fell prey to the "Museum Paradox" illustrated in Figure 25, by moving to far along the continuum between authenticity and education. Unlike classroom environments, where educators can have the "luxury" of a compulsory attitude towards learning, museums must try to both simultaneously educate and please their visitors. So museums are sometimes in a position of needing to cater to an object-based epistemology, even when it may be in conflict with instructional best practices.

2.1.1.2 Implications for Computer-Based Exhibits

The need to cater to a lingering object-based epistemology has ominous implications for the design computer-based exhibits, given the preceding. Computers present things (images, typically) that are patently unreal and inauthentic. Many museum theorists subscribe to the idea put forth in Walter Benjamin's seminal essay, "The work of art in the age of mechanical reproductions," that a recreation of a formerly unique, authentic object somehow lacks the original's meaning and power (Benjamin, 1969; Frost, 2002; Gurian, 1999). The experience of museum practitioners seems to bear this perspective out: visitors prefer to see "real" things, whether they be art objects or artifacts, and feel "cheated" if presented with a mere simulacrum when an original is in existence elsewhere (Gurian, 1999). A study done comparing visitor reactions to actual works of art and high-definition computer representations of those same works of art bears this out – visitors overwhelmingly prefer "real" objects to computer-based reproductions of them (B. Taylor, 2003).

Sometimes, though, visitors may be misled into thinking a simulacrum *is* an authentic object, especially in the more modern, profit-oriented era of museum exhibitions wherein a certain amount of faux-tableaux "Disney-fication" is *de rigueur* (Evans, Mull, & Poling, 2002; Heartney, 1997). The term "Disney-fication" is used by museum professionals to describe what happens when a museum may knowingly or unknowingly compromise elements of authenticity in favor of providing an immersive, entertaining experience, as the theme parks owned and operated by the Disney company are wont to do (see the "Disney paradox" in Figure 25).

Figure 26. Photograph of the Canada Pavilion at Disney World, showing totemic artwork from the Pacific Northwest alongside a building meant to evoke the stone fortified buildings found in Vieux-Québec over 2,000 miles away, taken from (The Walt Disney Company, n.d.). An uninformed visitor could easily get the impression that these items are contemporaneous and from the same region.

Children are especially susceptible to confusing the illusory for the real, although adults are less immune than one might think (Evans et al., 2002). The trouble with this confusion is that, when conditioned into a mindset that encourages the unquestioning acceptance of illusory items, visitors often make erroneous assumptions. For example, a display of animatronic dinosaurs might place dinosaurs from two completely different eras (e.g., the Cretaceous and Jurassic) alongside one another, causing visitors to assume that they were contemporaneous. Likewise, visitors may assume that the colors chosen for the robots' reptilian skin represents the "true" appearance of the long-dead creatures, and most troubling, may believe that there are no longer any open questions regarding dinosaurs' appearances (e.g., that they may have been covered with feathers). These sorts of forestallments of independent reasoning may occur because highly realistic representations, especially those that are three-dimensional and tangible, have a very high salience – visitors may forget that what they are viewing is only intended to be a representational object, and forget about its "real" referent (Callanan, Jipson, & Soennichsen, 2002).

Visitors are not the only group of people who have problems with inauthentic/virtual objects in museums. There also exists an almost prejudice against the use of computers in exhibits on the part of many museum staff members (Frost, 2002; Wilcove & Eisner, 2000). Museum leadership tends to exhibit a "knee-jerk reaction to defend the 'real'…one of the museums' most sacred cows"(Roberts, 1997). These expectations place constraints on exhibit designers who wish to use computer technology – they must find a way that respects the wishes of both visitors and museum staff.

Designers who wish to make use of computing technology while still acknowledging the visitor desire to experience authentic objects have a few options. One is to design technology that augments an authentic object instead of supplanting it: for example, an accurate 3D visualization of an authentic artifact that allows visitors to inspect it from angles not possible with a traditional glass-case display (Onda et al., 2004), or a visualization that allows visitors to get extra information about historical personages in a fresco (Alisi, Del Bimbo, & Valli, 2005). In both of these cases, the computer-based component of the exhibit is presented in conjunction with the actual artifact it references, so a visitor is able to satisfy his or her desire to experience the authentic object while gaining the extra knowledge that the computing technology makes available (Evans et al., 2002).

Figure 27. Photograph of a visitor obtaining more information on a character in *The Journey of the Magi* fresco by Benozzo Gozzoli (1421–1497), from (Alisi et al., 2005). The large query-able projection screen *augments* the original fresco, which is on a nearby wall, by allowing visitors to seek extra information not apparent by viewing the fresco alone.

This will be termed the augmentation approach because the role of the technology is to add to the experience a visitor would get from experiencing the authentic object alone. What little research that has been done on the computer-based augmentation of authentic objects in museums bears out its promise. When researchers compared the family conversations that occurred around real plants, physical models of plants, and computer representations of plants at a botanical garden, they found that each modality inspired different types of explanatory conversations that could be complementary to one another (Kevin Crowley & Eberbach, 2005). The computer augmentation approach also seemed to succeed in attracting the interest of older visitors who would otherwise avoid computer-based exhibits (Hornecker & Stifter, 2006a, 2006b).

Another way to satisfy visitor's desire for authentic experiences is to provide as "realistic" of a recreated experience as possible. This will be termed the *verisimilitude* approach. While this may not work when the reproduction is being substituted for the "real thing," cf. Benjamin, it may work satisfactorily when the "real thing" is not accessible (Gurian, 1999; Nicks, 2002). For example, using a Virtual Reality (VR) theater to present a recreated archaeological site may work because there is no other way for visitors to see what the site would have looked like in its full glory (Tanikawa et al., 2004), or because the site is too delicate to allow visitors to physically experience it (Muller, 2002).

Figure 28. Photograph of a virtual reality theater presenting a 3D recreation of the Mayan Copan ruins, from (Tanikawa et al., 2004). Because it is impossible for the museum visitors (who are in Tokyo) to view the Copan ruins any other way, the *verisimilitude* approach used here may be appropriate.

Verisimilitude can be dangerous, however, because visitors can have a hard time distinguishing the "authentic" aspects of the representation from the conjectural portions (the "Disney paradox", see Figure 25). This can be a particular problem for computer-based media, because a "digital surrogate" may very well contain less information than the original, but visitors may not be aware that they are missing it, and take the surrogate to be a full and accurate representation (Frost, 2002). Some designers have approached the problem of misleading visitors with verisimilitude by making the distinctions between real and surrogate very clear, like Paul Chemetov's "allusion, not illusion" strategy at the natural history galleries at the Jardin des Plantes in Paris (Blandin, 2002; Newhouse, 1999). When reconstructing the galleries of taxidermied animals, in lieu of placing, say, a polar bear on a reconstruction of an ice floe, he would place the specimen on a sheet of frosted glass. The *allusion* approach to grappling with authenticity in the face of virtuality would be very applicable to computer-based exhibits as well – rather than expend the effort to construct a to-the-blade-of-grass-accurate representation of a scene, designers could pick and choose which aspects of the virtual presentation will have emphasis. The allusion approach has the advantage of being already battle-tested in museum environments: it has been successfully used by conservators in museums for decades. For example, when pots are reconstructed from pot sherds, missing pieces are filled in with daubs of obviously modern modeling clay – allowing the filled-in areas to clearly allude to the shape and form of the original pot, but not causing visitors to believe they are seeing the pot in its entirety.

A fourth approach is perhaps applicable only to science museums, where, as often as not, the "object" being presented is a scientific phenomena. Because phenomena are actively unfolding processes, it may be possible to reproduce that process in a manner that captures the salient elements, even though the substrate is a digital (as opposed to a tangible) medium. For example, one of the few computer-based exhibits at the Exploratorium hands-on science museum in San Francisco is the *Satellite Orbit Simulation*. In this exhibit, a visitor can manipulate the velocity and direction of a satellite to be fired into orbit around the Earth. Although the visual appearance of the simulation is not remotely realistic, the effects of gravity on the projectile are. Because the phenomenon (Newtonian physics) is reproduced accurately, one might argue that the simulation is authentic, as it is faithful to the disciplinary conception of the phenomenon (Bain & Ellenbogen, 2002). This style of establishing authenticity will be termed the process approach. Computer-based exhibits might even be superior to authentic objects for the purpose of illustrating processes in a manner that is accessible for visitors. In the study of visitor responses to real versus virtual plants, visitors were found to make significantly more process-oriented remarks when interacting with a virtual representation of a plant than when they were interacting with real plants (Kevin Crowley & Eberbach, 2005).

2.1.2 Interpretation and Authority

Interpretation: *The process of coming to "know" or understand an object presented in a museum. Traditionally, in museums, interpretation is an act performed by curators for the benefit of visitors, wherein the curators select which aspects of a presented object will be salient for visitors*

Authority: *The power to make independent decisions and judgments. Traditionally, in museums, the institution possessed all of the power to make choices, to construct interpretations, to deem ideas worthy or unworthy for consideration*

A visit to a museum entails coming into contact with items (the "objects" of Section 2.1.1 Authenticity and the Object-Based Epistemology) and gleaning some form of meaning from them. Visitors predominantly come to museums to "learn something," whether that something be about the objects in the museum, or something latent about themselves (Csikszentmihalyi, 1993). How visitors engage in this learning, though, is via acts of *interpretation*: while viewing an object, reading a bit of text, speaking to a companion, or participating in an activity, a visitor is always making connections between what he or she already knows, and what he or she is currently experiencing. Sometimes these acts of interpretation are unidirectional, as when a visitor reads a text-based label and adds the content to his or her memory. Other interpretive acts involve more back-and-forth interaction, as when a visitor experimentally swings a pendulum leaking sand in varying arcs across a platform, or engages in a debate with a companion. Whether one-way or bidirectional, these interpretive acts can be seen as elements of a dialogue, especially to proponents of material culture theory who view objects as semaphores that can be queried via the senses (Hooper-Greenhill, 2000). A museum visit can be thought of as a series of dialogues: between visitors and their companions (Ash, 2003), between visitors and curators (Lord & Lord, 2002), and between visitors and the objects themselves (Taborsky, 1990). In museums over time, there have been many changes in what is "said" in these interpretive dialogues, and also in who has the authority to guide the course of the "conversation."

Museum exhibits were originally created by curators for other researchers when museums were still sites of active scholarly research, a practice that reinforced divisions between the “elite” educated class and the “unwashed masses” when the latter group was eventually permitted to enter the galleries[6]. As mentioned in Section 2.1.1 Authenticity and the Object-Based Epistemology, museum collections were originally laid out in an ontological fashion, and visitors were expected to absorb knowledge in the same ways the curating scholars did: via “interrogative” methods like inspection, comparison, and even handling. The responsibility was on the visitor to initiate querying “dialogues” with the objects. When museums first opened their doors to the public, the specimens were presented as they were to scholars (out in the open, available for handling), but museums very quickly realized that too much handling would degrade the specimens and tucked them away in glass cases, with only brief specimen labels (often in Latin) identifying each object (Conn, 1998). Visitors were becoming more and more removed from having direct relationships to the objects in the museums, impeding whatever slim chance they might have had to interpret the significance and meaning of the objects. Unlike the original curator/scholars, early 20th century museum visitors were unable to engage in “dialogic” relationships with the objects in museums wherein they would be free to touch, manipulate, and otherwise explore the objects that caught their fancy.

Relatively quickly, museum professionals began to recognize that regardless of the degree of intangible magic authentic objects might possess, when they were cosseted away in vitrines, visitors were limited to engaging in “*covert* interpretive acts” that relied primarily on making deductions from the relative arrangements of objects themselves (Evans et al., 2002). A more overt form of interpretation was needed to help visitors acquire the intended knowledge, and thus the interpretive label, familiar to all inveterate museum-goers, was born (Conn, 1998; Roberts, 1997). The field of Visitor Studies was born in the early 20th century to try to understand how to help visitors interpret museum exhibits, and a major area of focus was on the structure of label text (Ramsey, 1938; Robinson, 1928, 1930).

The addition of interpretive labels helped visitors interpret what they were seeing behind glass cases, but had the side effect of further damping down whatever autonomy and authority visitors might have brought to their museum experiences. The interpretive labels tended to be very “didactic-expository” (Witcomb, 2006) in nature, and “the anonymous voices of museum authority” (McLean, 1999) left little to no room for visitors to make their own interpretations of what they were seeing in a museum. The museum’s label was *the* scholarly-sanctioned presentation of a topic, and visitors were expected to merely absorb the imparted wisdom. Visitors were not expected to have any views or opinions of any value or validity, and certainly were not expected to question what they were being shown or interact with it in any way. By removing the element of choice from the visitor, and by presenting information without alternative perspectives, museums were controlling the dialogue: and were also unintentionally undercutting visitors’ intrinsic motivation to explore and, consequently, to learn (Csikszentmihalyi & Hermanson, 1995).

---

[6]Even when the general populace was permitted to enter, museums were closed on Saturdays, Sundays, and public holidays "to keep out the 'vulgar class,' such as 'sailors from the dockyards and the girls whom they might bring in with them.'" (McLean, 1999).

Figure 29. Photograph of a man reading an interpretive label posted in front of a printing press, from (Hornecker & Stifter, 2006b).

Over time, and with the contribution of many, many visitor studies of label text, museums began triangulating on the features that would keep visitors interested: focusing on an accessible (not arcane) “core idea,” including diverse perspectives, posing questions, and inviting visitors to draw their own conclusions (Nicks, 2002; Roberts, 1997; Spencer, 2002). Essentially, museums were trying to give back some choice and control, some authority, to visitors, to encourage them to participate in the interpretive process (Martin & Toon, 2005). Some of the earliest children’s museums, like the Franklin Institute in Philadelphia and the Boston Children’s Museum, had taken the process further in the early 20th century by including interactive, hands-on exhibit components that gave visitors physical control over exhibits. The success of these “interactives” at capturing and sustaining visitor interest inspired many institutions to follow suit. It wasn’t until the civil rights era, however, with its concomitant sea changes in educational theory and social justice, that the rationale behind interactive exhibits and accessible label copy was clearly articulated.

The overwhelming theme during the civil rights era was that of democratization; taking power from the elite and distributing it among the disenfranchised, or, as it played out in museums, wrenching authority from the grasp of a curators and placing it in the hands of visitors. Ideas about education shifted away from more passive, didactic approaches towards constructivist “active learning,” wherein learners were treated as capable agents in charge of their own learning processes (Bruner, 1961; Papert, 1980; Piaget, 1970, 1976). Interactive exhibits allowed museum visitors to engage with content areas at their own pace, allowed them to construct their own interpretations of the content areas, and were purposely designed to engage novices and experts alike (Oppenheimer, 1968). Label text began to reflect an amalgamation of views: educators, designers, representatives from previously excluded cultural groups, and even visitors themselves worked with curators to design the text, so that no one viewpoint was privileged above others. In this period of widespread authority-challenging, curators lost their authoritative grip on the interpretation (McLean, 1999; Roberts, 1997). Visitors were no longer just expected to receive knowledge; they were now expected to actively construct it.

Figure 30. Photograph of the floor of the Exploratorium in San Francisco, one of the very first museums structured around the new “hands-on”, active learning pedagogy. Image taken from (Fleck et al., 2002). Very few of the exhibits on the floor are devoted to “objects” in the traditional, specimen-oriented sense: most are purpose-built interactives designed to showcase various physical phenomena.

By the mid-to-late-20th century, museums began to adopt strategies for presenting objects and ancillary content to encourage more active, constructivist learning processes. Science museums were especially amenable to this shift away from traditional, passive transmission-learning models, and entire museums sprang up around a new “hands-on” pedagogy. The theorists began to catch up with the practitioners in the 1990s (Falk & Dierking, 1992, 1995; Hein, 1998), and over time, the new “hands-on” pedagogies (termed, variously, hands-on learning, discovery learning, or inquiry learning) began to coalesce into a collection of theoretical stances and practical guidelines united under the “inquiry learning” banner[7]. The focus of all of these efforts was to transfer the responsibility for interpretation to the museum visitor, and to give the visitor the sense of authority necessary to engage in his or her own interpretations.

2.1.2.1 Visitor Expectations for Interpretation and Authority

Section 2.1.1 Authenticity and the Object-Based Epistemology discussed how, owing to the longstanding role and importance of objects in museums, visitors still have certain embedded expectations for how objects should be employed within museum exhibits. Visitors likewise have attitudes regarding interpretation and authority that are holdovers from earlier eras of museum history. Despite the hard work

---

[7] This was spurred on, by part, by the establishment of the NSF-sponsored Institute for Inquiry at the Exploratorium in San Francisco, one of the original hands-on science museums, and perhaps the most influential in terms of research and theoretical output (Allen, 2004).

of those who revamped museum exhibits to place responsibility for interpretation in the hands of visitors, many visitors still expect to be told "facts" by "experts" when visiting museums (Martin & Toon, 2005), or need to be properly motivated to do the hard work of learning. Educational researchers have found that discovery learning (the strictest flavor of self-guided learning) just doesn't yield the promised results (Mayer, 2004); learners still need some sort of guidance.

Museum exhibit designers have intuitively recognized the continuing need for guidance when the authority for interpretation is placed in the hands of visitors. Without the benefit of any familiarity with learning theory, it is still very obvious when visitors "fail" to use exhibits productively – the single biggest signifier is when visitors walk away from an exhibit owing to an inability to suss out how to interact with it (Atkins, 2006). Research shows that visitors respond very well to procedural instructions (Atkins, 2006; Bell, Bareiss, & Beckwith, 1993-1994; K. Crowley & Callanan, 1998; Gelman, Massey, & McManus, 1991), perhaps because in the absence of an authoritative voice telling them what *facts* they should know, they still feel more comfortable with having the authoritative voice of the museum tell them what *actions* they must perform. The problem with a strictly procedural activity is that, even though the visitor may be more engaged with the exhibit, the interpretive dialogue is once again unidirectional, with the museum telling the visitor what to do, what to know.

Although an interactive exhibit with very clear procedures will attract and keep visitors, it will not necessarily foster the type of active, constructivist learning intended by inquiry learning advocates (Allen, 2004). When visitors follow a set procedure, they are adopting an epistemic frame given to them by the museum, in lieu of marshalling (and perhaps modifying) their own mental models. To help distinguish exhibits which *physically* allow for interactivity but do not promote active learning from those interactive exhibits that *do* promote constructivist knowledge-building, some museum theorists have begun to label the former "hands-on" exhibits, while the interactives that promote learning are termed "minds-on" exhibits (Hein, 1998; Witcomb, 2006).

Museum professionals, especially those at "hands-on" science museums, have long been at work to try to make their exhibits more "minds-on"; in other words, to promote more cognitive engagement (Witcomb, 2006). The Exploratorium uses the phrase "Active Prolonged Engagement," or APE, to describe the behavior that they see as a necessary precursor to constructivist learning at interactive exhibits (Humphrey, Gutwill, & Team, 2005). Researchers have found that one component of APE is that of play: when visitors are allowed to take a playful stance towards an exhibit, they are much more likely to remain engaged in a manner that promotes knowledge construction (Diamond, 1996). In order for visitors to play with an exhibit, the full range of interactions with it should be made as transparent as possible, so visitors are aware of what elements they have at hand to experiment with (Diamond, 1999; S. M. Taylor, 1991). Play, as well as APE, also seems to require a certain amount of open-endedness to the tasks at hand (Humphrey et al., 2005). Visitors need to be able to try different actions, and perceive different outcomes, to spend enough time (and cognitive effort) to learn in a constructive manner from an exhibit. At the same time, while visitors need to know *how* to manipulate things, and need a *range* of manipulations and outcomes to toy with, they may still need some guidance on how to order their thoughts & efforts to overcome the known problems with interpreting the results of their actions under pure discovery learning (Mayer, 2004). Exhibit designers, then, have to tiptoe along a divide between, on the one hand, providing too little guidance and thus leaving visitors confused about how they might proceed, and providing too much guidance, and thus removing opportunities for constructivist learning.

2.1.2.2 Implications for Computer-Based Exhibits

A review of the history of the visitor's ever-increasing role in assuming authority over the interpretation of museum exhibits shows that visitors need to be allowed to engage in some sort of interactive dialogue with exhibits in order to best learn from them. Computer-based exhibits, what with their near-infinite possibilities for dynamic display and the ability to accept wide ranges of inputs, are innately more "ready" to be conversational partners than static objects, and certainly turn the authority for controlling the dialogue over to the visitor in a way static objects cannot (Diamond, Bond, Schenker, Meier, & Twersky, 1995). They may even be superior to many of the non-static, physically interactive exhibits for dialogic purposes (e.g., the hinged flip-panels that pose a question on the surface, and reveal the answer when lifted). As we have seen with physically interactive exhibits, however, mere interactivity, or "hands-on"-ness, does not guarantee that visitors will learn from an exhibit. Taking cues from what we know about the recommended designs of non-computer exhibits, then, we can generate few practical recommendations for how to encourage visitors to engage in active interpretation while using computer-based exhibits.

First and foremost, it must be clear to visitors exactly what sort of input they can provide to a computer-based exhibit, and how they should go about providing that input. This design strategy will be dubbed *transparent manipulation*. At the time of this writing, it is safe to assume that the majority of museum visitors have had experience with desktop computers at one point or another in their lives, so standard desktop computer-based input methods (e.g., keyboards, mice, trackballs) should pose few problems (poorly designed software interfaces can always be an issue, of course). With newer devices, like handheld computers and cellular phones, making sure that the means of input is clear can still be something of a challenge; this will be covered in Section3.3 Handheld Devices. These potential input issues fall squarely in the camp of usability problems that Human-Computer Interaction (HCI) researchers investigate.

Section 2.1.2.1 Visitor Expectations for Interpretation and Authority described why one might not want an activity to be strictly procedural, lest visitors once again lose out on the opportunity to construct their own understanding of the material. A study on a computer-based sickle cell counseling activity in a museum didn't try to examine the different types of knowledge formed when visitors engaged more or less with the procedural elements of the activity, but the researchers did note that they viewed the visitors' overwhelming preference for procedural over general knowledge building activities as problematic (Bell et al., 1993-1994). A viable alternative would be to scale back the degree of procedurality, making the activity less linear and more open-ended, by providing visitors with more choices and a larger array of potential consequences for those choices (Fehrer & Rice, 1985). This design strategy will be dubbed *open-ended* *outcomes*. Exhibits with multiple outcomes have been shown to offer particular learning advantages to groups of visitors, because they allow for "observation and interaction [that] are sufficiently complex to foster group discussion" (Borun, 2002). Even without the presence of companion, an exhibit with *open-ended outcomes* should provide for a richer learning opportunity. When an exhibit has open-ended outcomes, visitors have more of an opportunity to engage in a bidirectional dialogue with the exhibit, and thus take charge of their interpretation of the presented content.

In order to transfer some of the initiative for that bidirectional dialogue onto the visitor, the visitor may still need some sort of guidance from the museum's exhibit designers to help them choose from among the input options, and to help them infer the meaning of the open-ended outcomes. Guidance is not the same thing as a procedure: the latter is omnipresent and non-negotiable, whereas the former (ideally) only appears when a learner is "stuck," and need not be obeyed. A well-used strategy in educational software research for providing guidance to learners is that of *scaffolding* (Pea, 2004). The notion is scaffolding is that another entity (a peer, teacher, or computer) will step in and help learners succeed in solving problems that would otherwise be too difficult; so that the learner can accomplish more than he or she would be able to solo, but is still doing the "hard work" of learning (Wood, Bruner, & Ross, 1976). Scaffolding can take many forms, from passive supports (like structuring a user interface to encourage the efficient entry of data) to more active approaches (like dynamic expert guidance that will appear only when a user reaches an impasse) (C. Quintana et al., 2004).

A final recommendation is more of a caution. Sherry Turkle is a sociologist and psychologist who has been researching people's relationships to computers and computer-enabled experiences since the early 1980s. She found that as computers became more and more capable of dialogue, people became more and more prone to forming relationships with them that were akin to the sorts of relationships they would form with other humans, often employing some of the emotions and social conventions present in human-human relationships (Turkle, 1984). Requiring that visitors form highly-dialogic relationships with computer-based exhibits in order to learn from them can lead to a bimodal problem space: one scenario is where visitors avoid or otherwise fail to form the necessary dialogic relationship, preventing the exhibit from aiding in learning. To illustrate, in the 1990s when Jane Margolis and Allan Fisher studied why female enrollment in Computer Science at Carnegie Mellon was so low, they discovered that one facet was that women are more prone to resist the "intimacy" required of an intensely dialogic human-computer relationship (Margolis & Fisher, 2002). There is a possibility that some visitors will avoid making full use of AV guides out of a reluctance to engage "intimately" with them. Signs of this were apparent in older visitors' reluctance to approach computer-based exhibits in Section 2.1.1.1 Visitor Expectations for Objects and Museums. On the opposite end of the problem space is a scenario wherein visitors form *too* strong of a dialogic relationship, to the point where visitors are so distracted by the computers that they miss out on other learning opportunities provided by the context. If visitors *do* form strong "relationships" with computer-based exhibits, these relationships can supplant relationships they might otherwise form with other entities: objects within the museum, for example, or with the other people they are attending the museum with. Some pilot studies of Audio-Visual guides seem to bear out these two problems, as will be described in Section 3.3.1.3 Access to Companions while engaged in Single-User Handheld Device Activities.

2.2 The Social Context for Learning in Museums

Section 2.1 The Historical-Cultural Context for Learning in Museums described how, throughout much of the last few centuries, visitors were expected to learn from the objects found in museums, and how the responsibility for interpreting those objects was alternately shifted from the visitor to the museum and back. With all of the attention given to the visitor-exhibit relationship over the years, much less consideration was given to understanding visitor-visitor relationships. Only in the latter portion of the 20th century did researchers begin considering the impact of a visitor's companions on his or her museum experiences (Diamond, 1986; Falk & Dierking, 1992; vom Lehn, Heath, & Hindmarsh, 2001). Social factors quickly gained a firm foothold in the agenda for museum research (Falk, Dierking, & Holland, 1995). Not unsurprisingly, once researchers began to look for social factors, they discovered that visitors come to museums for predominantly social reasons (Roberts, 1997), and furthermore, found that visitor learning depends on their social interactions to a large extent (Diamond, 1999). This section will review some of the results of that research. The audiences for different types of museums can be quite different – art and history museum audiences, for example, tend to skew quite a bit older than visitors to science museums – so as much as possible, studies set in a science center context will be used. Results describing the composition or behavior of audiences in other types of museums, when presented, will be highlighted as deriving from an alternate (and perhaps inapplicable) context.

2.2.1 Demographics at Science Museums

Visitors to science centers predominantly attend in groups, and mostly attend as multigenerational family groups (Borun, Chambers, & Cleghorn, 1996). In a survey-based study with 348 of science center visitors,[8], researchers found that 74% were members of adult-child groups, 21% were members of adult-only groups, and 2% were members of a tour group (Korn, 1995).[9] The vast majority of visitors, 97%, were in groups - only 3% of all science center visitors surveyed were attending alone.

There were nearly 300 science museums in the US at the beginning of the 21st century, and half of their visitors were under 18 in age (Scott. G. Paris & Hapgood, 2002). The average age of the children in the survey-based study mentioned above was around 9, and the average age of the adults was around 39, although it should be noted these figures were drawn from only two science centers, the St. Louis Science Center and the Miami Museum of Science (Korn, 1995). The average audience age can be very different from one institution to another, and from one month of the year to another. Slightly less than 50% of adult

---

[8] The number who completed the survey; there was a 15% refusal rate.

[9] The study was conducted in July, so the number of tour group attendees would likely be higher at other times of the year. In the author's prior research at the Ann Arbor Hands-On Museum, dramatic spikes in attendance occurred during the school-year months, September to June, excluding December, owing largely to an influx of school group tours (Lyons & Pasek, 2006).

visitors had had some college experience, illustrating that science centers may cater to a slightly under-educated audience vis-à-vis other types of institutions, like natural history museums (Korn, 1995). That said, exhibit designers and researchers in science center contexts have found that adolescents and "naïve" adults (those who have no area of specialty/expertise related to the specific content area of an exhibit) are very much alike in their misconceptions and responses (Borun, 2002; Fehrer & Rice, 1985), so overall education level may not be a significant factor.

Science centers also tend to have more diverse racial and economic profiles than other types of museums, and in a large-scale study of museum visitation habits, neither race nor income were predictive of museum visitation (Falk, Brooks, & Amin, 2001). Overall, slightly more adult women attend science centers than men (Korn, 1995), although this may be due to the large numbers of attending family groups – women are often the primary caregivers to children, and may be more likely to take the children on an outing than lone fathers would be. In this author's own work at the Ann Arbor Hands-On Museum, this pattern seemed to be true; moreover, women were far more likely to act as chaperones for school trips than men.

2.2.1.1 Expectations Regarding Visitor Demographics

There is no real reason to suspect that the demographics for science centers will change dramatically, although it is likely that, over the coming years, more and more non-native English speakers will make up larger percentages of the audience (Falk et al., 2001). Over time, more and more families have begun visiting museums in lieu of (or addition to?) entertainment opportunities like theme parks (Heartney, 1997), which has probably contributed to the exponential boom in the construction of science centers(Zucker, 1987). It is probably safe to assume that visitorship will continue to increase in numbers and diversity.

2.2.1.2 Implications for Computer-Based Exhibits

Science centers host a wide variety of visitors, and so computer-based exhibit designers should make an attempt to meet their various needs. It seems that the range in ages does not unduly affect the design of *content* for museum exhibits – non-expert adults and adolescents responded similarly – so the content for computer-based exhibits could probably be targeted at a middle-school level and still reach a much larger range of ages. This design strategy will be termed *naïveté knows no age*. In fact, this approach might very well avoid the so-called "numbskull factor," wherein visitors are too intimidated by the sophistication of content to feel comfortable even trying to learn from it (Martin & Toon, 2005).

Unfortunately, age may have an impact on the delivery mechanism for that content. Older adults (age 35 and above), when surveyed on their preferences regarding interpretive strategies, rated "low tech" methods (e.g., objects & artifacts, or live demonstrations) significantly higher than their younger counterparts, whereas younger visitors conversely rated "high tech" methods (e.g., computer games). higher than their older counterparts(Korn, 1995). Recall from the discussion of objects and authenticity in Section 2.1.1 Authenticity and the Object-Based Epistemology that older visitors would avoid computer-based exhibits that had no complementary artifacts or objects (Hornecker & Stifter, 2006a, 2006b; Moritsch et al., 2004). In addition to the proven-effective *augmentation* design strategy (placing computer-based exhibits in conjunction with "authentic" objects), it behooves designers of computer-based exhibits to further explore why older visitors may not wish to interact with computer-based exhibits. In the meantime, since younger visitors are known to be much more likely to try computer-based exhibits, rather than sharpen the divide along age lines, a meta-design strategy would be to *privilege the design preferences of older users*. Thus, when trying to decide between alternate designs, a designer should consider whether one of them will make the experience easier, more enjoyable, or more usable for older audience members.

Designers should take a similar "meta-design" tack with respect to female visitors, who make up a slight majority of the visitor population. Female computer users have always lagged behind male users in terms of being recognized by designers as a legitimate computer user population: witness the computer game industry's surprise at the success of atmospheric strategy games like *Myst* and its ilk, and the sharp growth in the audience for "casual" online gaming. Both of these trends were driven by high proportions of female (and often middle-aged) users, much to the surprise of designers who had long since internalized the notion that gamers were perpetually 14-year-old boys with a taste for guns and bosoms (Brightman, 2006).

An early gender-comparison study in museums reported observers noting that boys seemed more comfortable with technology-oriented exhibits than girls, and in general, "The [girls] were less interactive, less participatory," but interestingly, the data didn't necessarily reflect this in terms of the number of exhibits visited, staying time, etc. (Carlisle, 1985). Although it's unlikely that museums will begin showcasing *Grand Theft Auto*- style, violent and misogynistic computer-based exhibits, if designers believe that their primary audience is young boys, that assumption may lead them to choose content areas that traditionally appeal more to boys. Part of the gender difference observed in (Carlisle, 1985) may have derived from the fact that the technology-heavy exhibits tended to focus on topics like physics and mechanics, whereas the exhibits favored by the girls dealt with topics like the science behind music and art (and it was perhaps coincidence that they were not computer-based). We already know that even parents at museums will treat their own children differently on the basis of gender, with lengthier and more detailed explanations and interactions being reserved for boys (Kevin Crowley & Jacobs, 2002). It may behoove museums, then, to adopt the meta-strategy to *design for the female audience* from time to time, to ensure that the needs and interests of women and girls are not neglected[10].

2.2.2 Group Behaviors at Science Museums

In a study of 100 visitors to a science center, they state that social experiences are at the top of their list of reasons for attending science centers, and their primary expectation for interactive exhibits at museums is that they will "promote talking, communication, or doing things together" (Falk, Scott, Dierking, Rennie, & Cohen Jones, 2004). Groups that attend a science center together tend to more or less

---

[10] This recommendation is not meant to be an exclusive idea: of course "male" interests (inasmuch as interests can be categorized along gender lines) should also be considered. If women are a majority in the visitor population, though, it wouldn't be wise for computer-based exhibit designers to forget the fact.

remain in groups for large portions of their time, engaged simultaneously in the learning endeavor (Borun, 2002). Family groups have been seen to "shop around" at museum exhibits, roaming around until they find one that captures the group's mutual interests (Diamond, 1986; Scott G. Paris, 2002). This is probably true for non-family groups as well. Sometimes groups will temporarily split up to pursue independent investigations, but generally regroup before too long (Falk & Dierking, 2000). A commonly-seen pattern in science museums is that a group member will notice something of interest, and call over or bring over companions to ask questions of or share the experience with (Carlisle, 1985; McManus, 1994; vom Lehn et al., 2001). More rarely, a member of a group at a science center will prefer to explore exclusively on his or her own (Ellenbogen, 2002), but a preference for solo visits seems to typically be dependent on the type of museum. While more culturally-oriented museums can have upwards of 30% of their visitors attend solo (Ballantyne & Packer, 2005), the evidence suggests that visitors attend science centers in order to have shared group, and often family, experiences.

Family groups (who may have slightly different dynamics than other social groups, but on whom we have the most data) studied in a large science center were found to spend less than a minute at 57% of the exhibits they passed by (Diamond, 1986). Once at an exhibit together, though, groups tend to stay about as long as individual visitors do. A study of family groups found they would spend an average of 3-4 minutes per exhibit at a science museum (Borun et al., 1996), and a study of child-only groups found that they would linger between 1 and 2 minutes (Carlisle, 1985). (Without knowing the structure of the exhibits in both studies, it is hard to make a judgment about whether the difference in linger time arose from the type of group, or the style of exhibit). As is true for individual visitors, the longer a group lingers at an exhibit, the more they are likely to learn from it (Diamond, 1996; Scott. G. Paris & Hapgood, 2002), and thus linger time can serve as a rough measure of the potential for learning. Thus, a facile recommendation for computer-based exhibits would be to increase the linger time for a group, but doing this is not a simple task. First we must take a more nuanced look at what behaviors the visitors are engaged in while lingering at the exhibit.

2.2.2.1 Expectations Regarding Group Behaviors

Visitors can engage in a whole range of different interpersonal behaviors that researchers are just beginning to explore, and not all of them necessarily promote group learning. The design of the exhibits themselves seems to strongly affect how visitors will engage socially with them. Even when interviewed visitors entered a science center with the primary goal of engaging in social learning, they would sometimes report a low level of social learning, which the interviewers attributed to the preponderance of exhibits not designed for group use (Falk et al., 2004). To determine which design features of these solo exhibits are negatively impacting social learning, it can be helpful to turn from self-reporting surveys and interviews (the mainstays of visitor studies work) to ethnographic study methods.

One key theme that emerged from an ethnographic analysis (of hundreds of hours of videotapes of museum visitors) is that the physical access to exhibit components strongly affected visitor behavior (vom Lehn et al., 2001). Much like the flow of a stream is affected by boulders placed within it, visitors will avoid entering regions of an exhibit if they are occupied by strangers (vom Lehn et al., 2001). In one case study, it was a fellow group member who prevented his companion from accessing an exhibit. Although he clearly wanted his companion present (beckoning her over), he apparently only wanted her to be present as some sort of a witness or audience, and used his own body to block his companion's repeated attempts to access the exhibit's interactive components (vom Lehn et al., 2001). Both companions (and, as one might project, their relationship!) may have been better served if the exhibit design had allowed each an access point. Sometimes the obstruction is not physical, but in the realm of intention. Researchers working from observational data have found that children and adults often have different goals for the use of interactive exhibits, and this can cause conflicts over how the group should jointly make use of an exhibit (K. Crowley & Callanan, 1998; Falk & Dierking, 2000; Gelman et al., 1991).

Despite the foregoing examples, more often than not groups of visitors can and do interact in science centers in ways that are promotive of learning. Family groups often use exhibits to stimulate conversation, either by incorporating new information presented by the exhibits into the discussion or by bringing up prior knowledge (Borun, 2002). What's more, visitor conversations have been found to be overwhelmingly "on-task" (i.e., relevant to the content of the exhibit at hand, and not just "social talk," like lunch plans). In a study designed to more closely examine the types of conversation produced at science center exhibits, Sue Allen found that 83% of the remarks made at the exhibits visitors stopped at could be categorized as "learning talk" (Allen, 2002). A lengthier discussion of what learning conversations at exhibits "look like," and how they can be measured, will be had in Chapter 6, Experimental Design and Methods.

Perhaps one of the most important roles companions can play for one another in a museum is to help co-manage each other's attention, both when they establish a locus of joint attention, and when they work to maintain that joint focus. Establishing attentional foci occurs both within groups, as when a parent points out a relevant exhibit component to a child (K. Crowley & Callanan, 1998; Diamond, 1986), and across groups, as when visitors will investigate the areas pointed or gazed at by strangers after the strangers have left the area (vom Lehn et al., 2001). Maintaining that joint focus (usually only a within-group process) is recognized by researchers studying collaborative learning as a key precursor to being able to engage in productive group learning dialogues (Barron, 2000, 2003). "Through interaction with each other, visitors negotiate access to and participation in the exhibit, and it is through this interaction that they come to experience an exhibit in highly contingent and situationally relevant ways" (vom Lehn et al., 2001).

The maintenance of joint focus is not always an evenly-distributed task. Group members are found to assume different roles within the group that help or hinder the maintenance of joint focus, and in turn, the group learning (K. Crowley & Callanan, 1998; Gelman et al., 1991). Although roles are a part of group behavior, it is a rich enough topic that it will be discussed in Section 2.2.3 Identity and Roles in Museum Settings.

2.2.2.2 Implications for Computer-Based Exhibits

The dynamics involved when groups use museum exhibits are very complex, but a few recommendations can be pulled out from Section 2.2.2.1 Expectations Regarding Group Behaviors. Although a group is unlikely to approach an exhibit currently being used by stranger(s), visitors often watch what strangers are doing, only to approach the area of interest when it is vacated. They may even try to imitate what they have seen other visitors do. This gradual movement along the continuum from observer to participant (via observing the actions of others) is a key feature of "legitimate peripheral participation," a powerful but often overlooked learning phenomenon first studied in the context of apprenticeships (Lave & Wenger, 1991). It is not unsurprising to see that museum visitors also engage in peripheral participation (although, of course, the time scale is very compressed compared to that of a true apprenticeship).

Computer-based exhibit designers should not neglect this facet of the museum-going experience, and be sure to *support legitimate peripheral participation*. The most direct implementation of this strategy is to make sure that other visitors' interactions with a computer-based exhibit are made observable – that the input actions taken by other visitors and the resultant changes to the software's output are made apparent. This issue of *access to output* will be returned to in Chapter 3 on Computing Technology in Museums, where it will be used as a frame to organize the analysis of prior work on computer-based exhibits in museums. One of the most commonly-found implementations of the *support legitimate peripheral participation* design strategy is the use of a very large display screen, an approach that will be described in more detail in Section 3.2.2 Multi-User Kiosks: Large Display Kiosks.

Once a group has approached a computer-based exhibit, though, it is equally important that group members have *access to input*. It is clear from some of the case studies described in the previous section that the user(s) who control the access points to the exhibit also control and shape the entire group's interactions with the exhibit. By following a design strategy of *supporting equal access by all group members*, the group members will all have opportunity to engage with and contribute to the emerging learning activity.

Software usage can be extremely opaque, as anyone who has watched over the shoulder of a colleague performing an unfamiliar task can attest. After being given access to input and output, a group of visitors may need some sort of awareness support. As mentioned in the section above, a large part of group learning activities involves the establishment of a group center of focus. If an activity is digitally-mediated (e.g., via keyboards, trackballs, etc.) it may not be obvious what a given visitor is attending to from one moment to the next, so it is important to *make users' attentional foci apparent* when designing computer-based exhibits.

Sometimes, even when visitors have equitable access to exhibits and are aware of their companions' areas of interest, they still run into conflict because they may have different goals for the group activity. While software-based exhibits are hardly in the position to judge whose goals are superior, one thing software can do is "enforce" the "rules" selected by a group of visitors. This is one form of *social* *scaffolding*, wherein the social interactions of visitors are actively shaped and supported by software features. Although social scaffolding is not a design strategy yet employed in museums, nor in this current body of work, it will be discussed further in Chapter 8, Discussion and Future Work.

2.2.3 Identity and Roles in Museum Settings

Identity: *The understanding a visitor has of his or her existence as an individual person, and of his or her existence in relation to others*

Role: *A collection of behaviors used by a visitor when interacting with an exhibit or with companions, often used to explore or reaffirm the visitor's identity*

Museums have a vested interest in helping visitors learn, but have been plagued by a variation on the truism: "You can [help] all people [learn] some of the time, and some people [learn] all of the time, but you can't [help] all people [learn] all of the time." Certain exhibit designs work very well for some percentage of visitors, and not for others. Demographics cannot be relied upon: although some demographically-linked trends can be identified, on the whole demographics (e.g., race, age, income level) have been found to be a very poor predictor of learning in museums (Falk, 2006). To understand why, museums would need to understand more about the visitors' internal processes. Influenced by cognitive psychology and constructivist theories of learning, museum researchers recognized in the mid-to-late 20th century that museums functioned as sites of personal meaning-making. Only recently, however, has the notion of a visitor's identity (both internally-constructed and socially-constructed) been flagged as a prime influence on meaning making, and begun to be explored in detail (Falk, 2006; Rounds, 2006; Silverman, 1995).

2.2.3.1 Expectations Regarding Identity and Roles

Individual identity is often explained in terms of a visitor's motivations: why he or she is attending the museum, and what he or she hopes to get out of the experience. Labeled variously as "visitor strategies," "entrance narratives," or "personal context," the basic notion is the same, that a visitor has a particular sort of goal in mind when visiting a museum, and will wear some sort of "hat" that helps them attain that goal (Falk, 2006; Leinhardt & Knutson, 2004). Some researchers have tried to deduce general categories for these "hats," breaking down the roles visitors assume when visiting a museum into categories like explorer, facilitator, professional/hobbyist, experience seeker, and spiritual pilgrim (Falk, 2006). After conducting interviews with 52 visitors and coding for these categories, the vast majority (87%) were found to be members of either the explorer (people who come to museums to pursue their own interests) or facilitator (people who come to museums to support the interests of their companions; especially their children) categories, or some combination of the two (Falk, 2006). These results complement the predictions of theorists, who claim that museum visitors grapple with and employ two different facets of identity when using a museum exhibit – their individual identity ("who I am as a person") and their group identity ("who am I as a group member") (Silverman, 1995).

A person's individual identity cannot be identified by a single static role, but is instead a process that unfolds over time, that both is generated by and generates actions, that both conforms to and collides with structures like physical, social and cultural forces (Rounds, 2006). Under this perspective, a person's identity is perpetually under construction, and the need to do "identity work" to further that construction process is what drives visitor behaviors. Identity work is seen as the process of both confirming an existing sense of self (e.g., "I am a person who likes animals," or "I am anti-abortion") while simultaneously safely exploring alternative identities (e.g., "I might be a person who enjoys chemistry," or "I might not be opposed to stem cell research"), thus "laying the groundwork for future changes in identity" (Rounds, 2006). Via identity work, people explore both their value systems (i.e., what they believe) and their envisionings of themselves as actors in the world (i.e., what they think they do and/or are capable of doing). Viewed through the identity work lens, seemingly idiosyncratic visitor behavior, like the haphazard curiosity-driven browsing of exhibits, is actually a rationally optimal strategy for discovering and engaging with exhibits that will help visitors engage in identity construction work.

Visitors are sometimes drawn to exhibits dealing with content areas very different from their own areas of expertise (the lure of the "exotic"), which helps them explore roles or ideas very far from their current identity (Rounds, 2006). This approach may work very well in cultural museums, but it seems that science centers have problems with visitors being "put off" by esoteric science topics. After conducting lengthy focus groups with participants of varying levels of comfort with science, one group of researchers found that "People reported being comfortable with scientific material when it is framed as 'nature' (as opposed to 'science') or as a topic affecting their personal lives--for instance, workplace health hazards."(Martin & Toon, 2005). So science centers may need to be more cagey about what topics they present to visitors, or what presentation strategies are being used, in order to support individual identity work.

Visitors take on a group identity when they assume roles that are conditioned on the presence and identity of other group members. Sometimes this role is similar to the "facilitator" role defined by Falk's interview study. In another study of family behaviors at science centers, one member will often assume the role termed "leading learner," a person (usually an adult female) who helps unify the interests and efforts of the other members of the group (Borun, 2002). This is similar to the mediation behavior, wherein a parent helps highlight or channel exhibit information known to be of interest to a child (Kevin Crowley & Jacobs, 2002), described in Section 2.2.2 Group Behaviors at Science Museums.

Most commonly, though, group members adopt shifting roles with respect to one another, with the most common being the following reciprocal pairs: the explainer and the listener, the demonstrator and the observer (Carlisle, 1985). Group members will take turns adopting the more active roles (explainer and demonstrator) and reciprocating with the more passive roles (listener and observer). When seen within family groups, a parent may take the role of explainer more often, and, if anything, children may be slightly more oriented towards being demonstrators (K. Crowley & Callanan, 1998; Gelman et al., 1991; Schauble et al., 2002). For interactive exhibits that require both strategic thought and manipulation, parents of younger children are seen to take most of the “thinking work” for themselves, and guiding their children through the more manipulative tasks (Schauble et al., 2002). The unfortunate side effect of this is that parents are not always good at verbalizing what they are attempting to accomplish, or verbally interpreting the outcomes of their joint endeavors, leaving the strategic component of the task opaque for their children (Schauble et al., 2002).

2.2.3.2 Implications for Computer-Based Exhibits

People seldom have a chance, in everyday life, to challenge or explore their values, since many aspects of everyday life and many social structures (e.g., political parties, religious groups) are predicated on people remaining constant in their beliefs. Typically, people change beliefs only after some sort of confrontation, usually brought about by a difficult personal situation. Computer-based exhibits have the advantage of being able to immerse people in scenarios that can present visitors with ethical dilemmas that they would otherwise only face in the context of a more traumatic real-world personal experience. The immersion of visitors in a role-playing experience wherein they have to opportunity to confront and either confirm or change values will be dubbed the *values roleplay* design strategy.

Museums also offer visitors the opportunity to conduct in “identity work” by exploring wholly new roles or identities. Computer-based exhibits, and their unique potential for immersive roleplay can once again aid this process by allowing visitors to engage in *occupational roleplay*, where visitors can “try on” different professions. Via occupational roleplay, visitors can see what it’s like to be people they’re not currently, but might like to be. One unique character of immersive roleplaying experiences is that it exposes visitors not just to surface-level details about a role, but it exposes them to an abstracted version of the problem space faced by people in those roles (Gee, 2003). So, rather than coming away with just surface-level details about a role (e.g., “chemists wear lab coats”) visitors may come away with a deeper feel for that it means to be a chemist (e.g., “chemists must make very precise measurements”).

Visitors to science centers may also need special “help” to get engaged with certain scientific topics – unlike visitors to, say, cultural museums, science center visitors can be put off by exotic or esoteric topics. Thus, computer-based exhibit designers should try to present scientific concepts in a context that has *relevance* to visitors’ lived experiences. This can be done by relating a concept to a commonly-seen but seldom-understood phenomenon (e.g., using how refrigerators work as an entry point for discussing air pressure and temperature), by discussing themes dealing with nature (which seem perennially popular and accessible to visitors), or by emphasizing a concept’s impact on visitors’ personal lives (e.g., the impact of sun exposure on skin cancer rates).

Group identity should be supported by computer-based exhibits less by “giving [visitors] what they want,” as supporting individual identity requires, and more by “giving [visitors] what they need.” In general, visitors who tend to take more active roles learn more, and retain that learning over time, as compared to visitors whose purpose is to support or facilitate their companions’ visit (Falk, 2006). And yet, just as a conversation cannot consist of a group of people all talking at once, a group visit cannot consist of all demonstrators, or all explainers, and no observers or listeners. Computer-based exhibits should thus *support fluid role changes*, from explainer to listener, or from demonstrator to observer.

Ordinary interactive exhibits usually are constructed so observing visitors can see what demonstrating visitors are doing. Computer-based exhibits should be designed the same way, to *make user's actions apparent*: if visitors are to engage in roles like mediator or observer, a visitor must be privy to the actions of other visitors. Likewise, computer-based exhibits should *make outcomes apparent*, because the visitors engaged in guiding roles like mediator or demonstrator or explainer may not be good at articulating what is happening in the context of an interactive activity.

2.3 The Physical Context for Learning in Museums

The physical context has a powerful ability to shape the behavior, and thus the learning activities, of the people within it. There are three aspects of the physical environment of museums that are thought to, or have been shown to, have an impact on learning behaviors: the visual design, the spatial design, and the interaction design. The Section 2.3.1 will cover research on the visual design of exhibits, and how that appearance can encourage or discourage the attention of museum visitors. Section 2.3.2 will cover the physical arrangement of items in space, which can impact social learning behaviors by mediating and influencing the interactions between learners (Pea, 1993; vom Lehn et al., 2001).

The third aspect of the physical environment of museums, interaction design, which, as the name suggests, really only applies to interactive exhibits (as opposed to static exhibits like oil paintings in an art museum). Interaction design can be thought of as a specification for the types of actions and reactions permitted for both the visitor *and* the exhibit itself. The first "interactive" exhibits made use of simple, manually-operated manipulatives, like simple-flip-top question panels. Over time, however, interactive exhibits have not surprisingly become closely linked to current technological capabilities, embracing new digital media as computers became available for use in exhibits. Because the interaction design for *digitally* interactive exhibits falls squarely under the Human-Computer Interaction (HCI) discipline, it will be covered in more detail in the Related Work chapter, Chapter 3, Computing Technology in Museums. It is worth mentioning, though, that visitors consider interactive exhibits to be extremely important part of the museum-going experience, viewing them (correctly or not) as "the best way to learn" (Falk et al., 2004).

2.3.1 The Visual Design

Visitors are notoriously fickle about which exhibits they will engage with. In a large tracking study that spanned multiple types of museums, only a minority – about a third of all participants – were classified as "diligent" visitors, meaning those found to stop at more than 50% of the exhibits in an exhibition (Serrell, 1997). Over time during their visit, visitors become even more reluctant to stop at exhibits, entering a "cruising" mode after about 30 minutes when "museum fatigue" sets in (Falk, Koran, Dierking, & Dreblow, 1985). The visual design of an exhibit, then, becomes a very important factor in the success or failure of an exhibit in attracting visitor attention. The first thing a visitor notices about an exhibit, barring any visual disabilities the visitor may have, is its visual appearance. What he or she sees often determines whether or not the visitor will approach (and potentially learn from) an exhibit, or whether he or she will avoid the exhibit – a property known as the "attracting power" of the exhibit (Peart, 1984). There are several major visual elements that affect visitors' perceptions: the visual media of the exhibit, the overall aesthetic style, and the means by which text information is displayed.

2.3.1.1 Expectations for Visitor Responses to Visual Design

An exhibit's "visual media" can encompass a great range of properties, from abstract to concrete (e.g., text versus objects), from static to dynamic (e.g., photographs versus video), or from two-dimensional to three-dimensional (e.g., photographs versus objects). In general, visitors seem to be more attracted to more "concrete" exhibits, unsurprisingly preferring exhibits with dioramas and photographs to more abstract, text-only exhibits (Peart, 1984). Likewise, visitors were found more likely to linger (the "holding power" of an exhibit) in front of the more concrete exhibits (Peart, 1984). Perhaps unexpectedly, however, the dimensionality of visual representations did not seem to affect visitor responses – visitors were just as likely to be attracted and held by a diorama as they were to an exhibit that used a photograph of that same diorama (Peart, 1984).

Practitioner-oriented research in museums has identified several aesthetic approaches that can attract visitors regardless of their native interest levels: the use of bright colors, the use of unusual objects, and the use of unfamiliar scale (Kaynar, Pasek, & Lyons, 2004). The purpose of these varied techniques seems to be to incite surprise, or perhaps even cognitive dissonance, in visitors. That said, designers must take care not to overwhelm the visitor – too many stimuli crowded too close together will quickly lead to museum fatigue (Maximea, 2002b). What can initially attract attention can just as quickly repel it.

A very special visual component of museum exhibits is the text-based label. The "bread and butter" of traditional exhibits, text labels are usually responsible for the majority of information transmission. Over the years, some very precise recommendations have emerged for practitioners to avail themselves of: for example, main titles should contain only three to eight words, and should be two to three inches high, according to the *Manual of Museum Exhibitions* (Spencer, 2002). Label content and presentation has been one of the most highly-studied aspects of exhibit design over the years. The general consensus, though, is that text of different levels of importance should be of different sizes (titles being large, detailed text being smaller), the snippets of text should be easy to read, and brief (no big blocks of text), and the vocabulary should be aimed at a middle- or high-school age of reading comprehension (lower for museums more oriented to children, like science centers).

2.3.1.2 Implications for Computer-Based Exhibits

Visual exhibit design is much more of an art than a science, but if one was to adopt the going conventions for exhibit design for use by computer-based exhibits, there are a few ideas that transfer. When possible, *use imagery that is suggestive of more dimensions* (e.g., prefer 3D graphics over 2D graphics, or shaded 2D graphics over flat-color 2D graphics). For whatever reason, visitors tend to look longer at images with the illusion of greater depth. The *color schemes used should be bright*, to attract attention, but not so garish as to make continued viewing painful (or most importantly – they should allow visitors to clearly distinguish between interactive and non-interactive elements in the display). Digital displays offer different affordances for text than static, printed displays: most notable the fact that digital displays have a near-infinite capacity for displaying text-based content. That said, designers would probably be wise in reducing the *amount* initially presented to appeal to the younger part of the demographic. More information could certainly be made available, but presenting too much at first might "scare off" visitors. In a similar vein, when *delivering* text, it should probably be displayed in a series of smaller chunks, rather than as one big scrollable block of text.

2.3.2 The Spatial Design

No exhibit, much like no man, is an island. It exists in relation to the other exhibits in the local area (usually, a cohesive exhibition) as well as in relation to the larger context of the museum space itself. The earliest visitor studies work was devoted to discovering the behavior patterns of visitors when confronted by the panoply of exhibits available to them at a museum, and very quickly, these investigations began to acknowledge the critical role played by the spatial location of exhibits.

2.3.1.1 Expectations for Visitor Responses to Spatial Design

The early visitor studies revealed many interesting phenomena – for example, all other things being equal, North American visitors tend to explore the right side of galleries first, and follow a counter-clockwise path (Melton, 1935), and that when exhibits are indeed presented as "islands" in the middle of the room, they are neglected by visitors (Weiss & Boutourline, 1963). Such findings have been confirmed multiple times by other researchers (Serrell, 1997).

The manner in which exhibits are displayed relative to one another can also play a large role in visitor learning experiences. In a study of cross-exhibit effects, researchers at the Exploratorium found that only those exhibits with the clearest relationships to one another generated inter-exhibit talk (Allen, 2002). Absent any other orderings, visitors use extremely shallow visual cues to decide if exhibits are related or not, especially if these exhibits are not in close proximity to one another. Therefore, many exhibit designers spend a great deal of planning on the arrangement of exhibits *vis-à-vis* one another. They usually establish a physical arrangement of exhibits that mirrors, on the physical plane, the mental arrangement of concepts visitors should employ when constructing their mental models. Designers thus employ many different strategies like focal (radial) arrangements, hierarchical arrangements, sequential orderings, parallel orderings, two-dimensional matrices, and contextual approached like the "onion" or the "pizza" (Nicks, 2002).

The decision of which arrangement to use often has a lot to do with the path designers would like visitors to take through the exhibition, which in turn is usually based on the cognitive path designers would like visitors to take through the content area. For many historical exhibitions, for example, sequential, parallel, or matrix organizations are used, because these arrangements have an axis which can be used to represent time (and thus line up the artifacts in order of provenance). Science centers, though, typically lack any overarching narrative structure, and so they use only the loosest of arrangements, like the “pizza” structure wherein exhibits are scattered about like pepperoni, perhaps only loosely grouped into “slices” of common content areas.

The placement of exhibits *vis-à-vis* other exhibits is not just important to help visitors build mental models. The placement *vis-à-vis* other exhibits also dictates accessibility, depending on how much space is left on which sides of the exhibit. Some visitors would block others, physically, from using components of the exhibit, as when a male visitor used his body to block off his female companion’s access to the manipulative portions of an exhibit (vom Lehn et al., 2001). Thought must be given to how many visitors should be able to access the exhibit at a given time, and then ample space should be provided for the desired number to be able to reach the exhibit without blocking one another (Borun, 2002).

Visitors block other visitors in more subtle ways as well. Visitors are often loathe to enter the personal space of strangers, and will wait for them to vacate before engaging with the area of interest (vom Lehn et al., 2001). What is interesting, though, is that even though they will not enter into the space occupied by strangers, they will monitor what the strangers are paying attention to, and will often follow up on the pointed fingers and gazes of the subjects of their surveillance when they enter the vacated space. This means that exhibits should have roomy regions devoted to egress, and be placed so that visitors can covertly observe other groups while they use the exhibit.

2.3.1.2 Implications for Computer-Based Exhibits

An exhibit designer, though, seldom has much control over the spatial dimensions of a gallery, but he or she can control aspects of the design that will eventually dictate placement, like *orientation*. If an exhibit has a clear “back” and “front,” it will most likely be placed against a wall, which increases the chances of visitation. Because software-based exhibits lack authenticity, they can benefit from being placed near exhibits that do contain authentic objects (see Section 2.1.1.2 Implications for Computer-Based Exhibits). Regardless of where the exhibit gets placed, it should be designed so that the intended number of visitors find it *accessible*, so that they can all access the intended output and/or input opportunities. Planning for seating is one way to do this – providing seating has the added bonus of making it more likely that visitors will tarry and learn more from the exhibit. Finally, the exhibit should be designed so that visitors can easily observe how other groups of visitors are making use of the exhibit. Practically speaking, this will likely involve the use of a large or prominently-mounted display.

2.4 Summary: Recommendations for Multi-User Software-Based Exhibits Arising from the Context of Use

In this chapter we reviewed three different contextual perspectives relevant to the design of multi-user software-based museum exhibits: the historical-cultural context, the social context, and the physical context. The sources of information for these contexts were primarily drawn from both practical and theoretical museum studies literatures. This section summarizes the design strategies for multi-user software-based exhibits that emerged from the above contextual review. These strategies are not offered up as dictates: their purpose is to provide guidance rooted in the trends of thought and application found in museums. Some of the design strategies listed below are more strongly supported than others; although many may be empirically testable, the purpose of presenting them here is more to give a rough outline, a road map, to the sorts of design concerns that arise from a thorough understanding of the museum context as a problem space for software design. In Chapter 5 on Design Rationale and Implementation, these recommendations will be referred back to for the purpose of explaining and contextualizing design decisions.

- **Historical-Cultural Context Design Strategies**
    - *Authenticity*: Designers must accommodate the visitor (ad museum staff) expectation that authentic content will be presented. There are three ways authenticity could be established for software-based exhibits:
        - *Augmentation*: software is designed to accompany and complement an existing authentic object, perhaps by presenting information not immediately apparent from inspecting the authentic object itself
        - *Verisimilitude*: software is designed to reconstruct, as accurately as possible, an authentic object or context. This works best if an authentic real-world referent doesn't currently exist (e.g., recreations of ancient ruins)
        - *Allusion*: software uses varying levels of representational detail to indicate which portions are solidly supported by existing evidence (and/or real-world referents) and which portions are conjecture
        - *Process*: software is designed to reproduce, as accurately as possible, a process (e.g., a scientific phenomenon) by correctly representing the behaviors of all relevant elements of the process
    - *Interpretation*: visitors should be able to engage in a productive interpretive dialogue with the software
        - *Transparent manipulation*: The means of providing input to the software should be clear
        - *Open-ended outcomes*: Software should provide outcomes that depend on different visitor actions, so as to promote Active Prolonged Engagement
        - *Scaffolding Learning*: Software should provide just enough guidance to help visitors maintain a thread of inquiry and be mindful of the important issues at play, without dictating the course of the visitor's interpretive dialogue
- **Social Context Design Strategies**
    - *Support multiple demographics:* Science centers draw a wide range of visitors, and software should be designed to encourage their participation

    - *Naïveté knows no age*: designing activities around content targeted at middle-school (or slightly younger) audiences has been found to work just as well for non-specialist adults
    - *Privilege the design preferences of older users*. When trying to decide between alternate designs, a designer should consider whether one of them will make the computing experience easier, more enjoyable, or more usable for older audience members, who tend to otherwise avoid computer-based exhibits
    - *Design for female audience*: Adult women make up a slight majority in audience, but may be overlooked as a user population, so their needs and interests should be kept in mind.
  - *Social groups*: software should support use by small groups of visitors, since they make up the majority (97%) of all science center visitors
    - *Support legitimate peripheral participation*: visitors will not often approach an exhibit in use by strangers, but they will productively watch what strangers do
    - *Support equal access by all group members*: sometimes a member will attempt to commandeer interactions with an exhibit by dominating the exhibit's access point(s)
    - *Make users' attentional foci apparent*: group learning processes depend on knowing what aspects of the exhibit the other visitors are attending to
    - *Social Scaffolding:* use static and dynamic software elements to support and guide social interactions (especially goal attainment)
- **Identity and Role Design Strategies**
  - *Individual roles:* software should support visitors in their "identity work" by allowing them to take on roles that cause reflection
    - *Values roleplay*: software should allow visitors to make choices that cause them to examine their values in a "safe" manner
    - *Occupational roleplay*: software should support the ability for visitors to "try on" an occupational role (e.g., a meteorologist, or a park ranger) to understand the problem space faced by those professionals
    - *Relevance*: visitors respond better to content areas that relate to their lived experience of the world
  - *Group roles*: software should support the different types of roles visitors can assume relative to one another when in small groups
    - *Support fluid role changes*: visitors gain different things from accepting active (demonstrator or explainer) and passive (observer or listener) roles, so computer-based exhibits should allow visitors to switch between roles
    - *Make user's actions apparent*: if visitors are to engage in roles like mediator or observer, a visitor must be privy to the actions of other visitors

        - *Make outcomes apparent*: because visitors who assume the mediator, explainer or demonstrator roles might not be good at articulating what the outcomes of a group's interaction with an exhibit, computer-based exhibits should provide feedback to visitors regarding the outcomes

- **Physical Context Design Strategies**
    - *Visual Design*: software should be in keeping with the dominant visual styles present within the museum
        - *High dimensionality:* imagery that looks to have extra dimensions (e.g., 3D graphics) will innately interest visitors more
        - *Bright but distinct color schemes*: software should use color schemes that attract the eye, but colors should not be so jarring that interactive elements are indistinguishable from non-interactive elements
        - *Informational Content*
            - *Amount*: software has the capability of providing near-infinite depth of content, but initial exposure should always target the lowest common denominator: children
            - *Delivery*: software should space out the delivery of informational content into easily-digestible chunks
    - *Spatial Design*
        - *Orientation*: one-sided exhibits increase the likelihood of being placed against a well-trafficked wall (and power drops)
        - *Relational placement*: software exhibits should be placed in the same area as other exhibits with similar content, especially if some of those exhibits contain authentic objects (see *Historical-Cultural* > *Authenticity* > *Augmentation* recommendation above).
        - *Accessible*: software exhibits should provide output and/or input opportunities, and seating and for all group members
        - Observable: software exhibits should allow other visitors to visually "eavesdrop" on how the exhibit gets used by others

CHAPTER 3

Related Work:
Computing Technology in Museums

The purpose of this chapter is to familiarize the reader with the types of computer technology already in use in museums and elsewhere to support co-located small-group learning activities. Computing technology was present in museums as early as the 1970s, but usually only in the form of bespoke, single-purpose units[11]. In a sense, they were artifactual objects in and of themselves, and didn't really become seen as "mere" vehicles for digital media experiences until the advent of multimedia laserdiscs in the 1980s (Diamond, 1989). Since that time, of course, there has been a great proliferation in the forms computers have taken in museums. The literature on technology use in museums is sparse and tends to be practice-oriented, so the next chapter will review the more research-based Computer-Supported Collaborative Learning (CSCL) and Computer-Supported Collaborative Work (CSCW) bodies of literature to round out the reader's understanding of the theoretical concepts that can be brought to bear on the challenge of face-to-face small-group cooperative activities.

3.1 Introduction to Computing Technology in Museums

The largest body of research on computer-based experiences for visitors in museums covers the design of single-user audio-visual (A/V) guide devices, which will be covered in Section 3.3.1 Single-User Handheld Device Activities in Museums. Perhaps the next-largest body of research covers collections management databases, and software to make digital collections records available to visitors (Cameron, 2003). Since the area of interest here is science museums (which usually have no collections, per se) these collections-based computer systems will not be reviewed.

The last category of software found in museums is software-based exhibits, of which there are a wide variety of different implementations, but in terms of published records, the documentation tends to

---

[11] While conducting the research for this dissertation at the Exploratorium in the summer of 2007, the author happened to have a conversation with a staff member who had been a regular visitor to the museum during his childhood in the 1970s. He recalled that, shortly after the arcade game *Pong* came out in 1972, the Exploratorium acquired a unit and placed it on the floor of the museum – the first computer-based exhibit ever placed in the museum. He said the game remained there until after the creation of *The Computer History Museum*, which offered to trade a version of Lunar Lander for the *Pong* game. *Lunar Lander* was still on the floor as of 2007.

take the form of anecdotal stories or exhibit reviews rather than formal research reports. Many of these software-based exhibits appear anecdotally in museum practitioner publications (e.g., *Curator* often contains exhibit reviews), so these sources will be used when more empirical sources are not available. The two most commonly-seen examples of software-based exhibits involve interactive displays: in the small-scale, these take the form of single-user kiosks; and when writ large, they take the form of multi-user kiosks – essentially, kiosks with a large shared display. A specialized category, and one that is just beginning to be explored, is the Multi-Machine User Interface (MMUI) approach to museum exhibits, wherein input and output for the software is distributed across multiple devices. All of these categories will be reviewed in this chapter.

3.1.1 Organizing Principles: Access

It is difficult to organize these disparate bodies of literature into a cohesive narrative, seeing as the researchers in CSCL and CSCW and the practitioners in museums are often grounded in different bodies of theory. The input/output (I/O) modality will be used to frame the discussion of the different computer systems described in this chapter, much as was done in (Stewart et al., 1999). An I/O categorization perspective is very flexible, in that (a) all HCI scenarios can be described and categorized in terms of the I/O particulars, and (b) this categorization does not rely on any particular theoretical perspective. So, for example, even though a CSCL application designed for picture-sorting may be based on different theories of human behavior than a CSCW application designed for picture sorting, and both of these may in turn differ from a museum exhibit designed to allow visitors to sort through images of artifacts, if all three share an IO configuration (e.g., a touch-sensitive tabletop), a thoughtful comparison of their designs may yield broadly applicable insights.

Apart from I/O, a third organizing principle needs to be considered when categorizing collaborative computer systems. A particular I/O configuration may work perfectly well for a scenario wherein a user is not expected to converse with his or her companions, and yet fail spectacularly when applied to a scenario where constant conversation is the norm. For this reason, the degree of interaction a user can have with his or her companions really must also be specified.[12] So, the organizing theme for this chapter will be a trifold description of a user's access: (1) access to input opportunities, (2) access to output opportunities, and (3) access to his or her companions.

These three themes relate to the three contexts for learning presented in Chapter 2. The accessibility of input and output is very much a part of the physical context (see Section 2.3 The Physical Context for Learning in Museums). The access to one's companions is very much a part of the social context (see Section 2.2 The Social Context for Learning in Museums). All three themes (input, output, and companion access) are subject to the historical-cultural context of museums. As explained in Section 2.1

[12] There may be a cross-cultural issue here as well. Although the author has not done a formal study of this facet, the impression is that much of the CSCL and CSCW research produced in Japan is predicated on the assumption that collaborators will not, in fact, communicate directly with one another vocally, even if they are expected to be in the same room at the same time.

The Historical-Cultural Context for Learning in Museums, people's behavioral norms are affected by their socially-constructed expectations for an environment. For example, one group of researchers have found that older visitors tend to avoid computer-based exhibits, preferring instead to only interact with more traditional object-based exhibits (Hornecker & Stifter, 2006a, 2006b; Moritsch et al., 2004). This is not a simple case of technophobia, however, as the researchers fond that the same older visitors can be persuaded to interact with a computer if it is presented in conjunction with an object. This example shows that it is probably the expectations of visitors (specifically, the expectation that exhibits are about objects – see Section 2.1.1 Authenticity and the Object-Based Epistemology) that is shaping their behavior. It is likely, then, that the manners by which visitors take advantage of access opportunities to input, output, and their companions are likewise influenced by culturally and historically shaped expectations.

Each of the form-factor categories covered in this chapter (single- and multi-user kiosks, single- and multi-user handheld device activities, and MMUI systems) has its own tradeoffs with respect to a user's access to input, output, and to his or her companions. These access tradeoffs naturally affect the complexity of the interfaces that can built around these different form factors, and in turn impact the richness of learning activities that can be supported by those interfaces.

3.2 Kiosks in Museums

Kiosks, in one form or another, are the most common form of software-based exhibits found in museums today, although fewer research publications have been devoted to them than to more esoteric hardware form factors. Kiosk exhibits were originally created from standard personal computers (or "microcomputers," as they were known at the time), and later came to incorporate more elaborate input devices and larger displays as technology improved. We will first explore the original form-factor, single-user kiosks, and then talk about how the addition of large displays altered the kiosk's access parameters in a separate section, Section 3.2.2 Multi-User Kiosks: Large Display Kiosks.

3.2.1 Single-User Kiosks in Museums

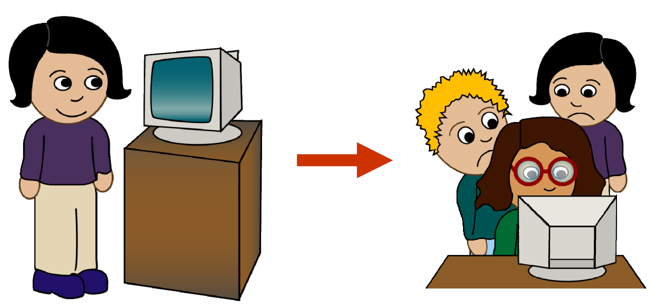


Figure 31. Illustration of the typical presentation of a single-user kiosk in a museum. The image on the right shows how groups of visitors may run into space limitations when trying to simultaneously use a single-user kiosk.

The conventional form computer-based exhibits take in a museum setting is that of a single-user kiosk (see Figure 31). These exhibits, much like the A/V guides that will be covered in Section 3.3.1

Single-User Handheld Device Activities in Museums, are an outgrowth of an older presentation technology: video displays. With the original video kiosks, visitors could only passively watch a looping video, but later video-disk based displays would allow visitors to choose which video segment to view. For a long time, this narrative approach, wherein the visitor provided little to no input to the system, affected the way computer-based kiosks were designed.

Early design approaches to single-user kiosks favored a linear, narrative presentation over a random-access, visitor directed experience, the rationale being that an average visitor might experience cognitive overload if presented with a wealth of options (Cooper, 1993; Yamada, Hong, & Sugita, 1995). This fear proved not to be true – in fact, some researchers found that the degree of open-endedness correlated positively with the "staying power" of exhibits (Sandifer, 2003). For example, the creators of one of the largest early exhibitions to heavily use single-user kiosks, the Smithsonian's *Information Age* exhibition at the American History Museum, were concerned about a particular single-user kiosk because it was much more open-ended than any of the others in the exhibition (Allison & Gwaltney, 1991). When the exhibition opened they were surprised to find that the single-user kiosk that they were afraid would prove too complicated was in fact the one most popular with visitors.

3.2.1.1 Input to Single-User Kiosks

Since the 1990s, single-user kiosks have trended towards allowing visitors to provide a much richer array of input. As far as input mechanisms go, touch-screens have proved very popular at many institutions, such as London's Museum of Science (Gammon, 1999), but specialized buttons, trackballs, mice, and even keyboards are also quite common. The only trouble with these latter forms of input is that they are designed around use by a single person at a time, as represented by the tangential intersection of the *users* and *input* categories in the Venn diagram in Figure 32. Although more than one user *can* use a touchscreen at a time, the form factor sometimes makes this inconvenient by not allowing all of the group members to get close enough to the screen to touch it (see Section 3.2.1.2 Output from Single-User Kiosks). Moreover, the technological limitations on the number of fingers that can be recognized at any one time force any attempts at joint use into a serial (as opposed to a synchronous) use paradigm. Experiments with mouse turn-taking in classrooms (Inkpen, McGrenere, Booth, & Klawe, 1997) and practical experience with touchscreens in museums (Gammon, 1999) show that when multiple users share a means of input to a computer (especially younger users), conflict will often ensue.

3.2.1.2 Output from Single-User Kiosks

Although designed to be used by a single visitor at a time, when on the floor these types of exhibits are quite frequently used by small groups (Allison & Gwaltney, 1991; Bell et al., 1993-1994; Lyons & Pasek, 2006). The trouble with the form factors of these exhibits is that the output from the kiosks is usually displayed on a standard computer screen, which doesn't practically allow too many visitors to gather around and have a clear view of it (Gammon, 1999). The consequences are illustrated in Figure 31, which depicts the degree of user access to output opportunities. The main user – the person providing input

– of the single-user kiosk presumably has complete access to the kiosk's output, but the other users may or may not have a clear view of the output, represented by the incomplete overlap of the *users* and *output* categories in the Venn diagram in Figure 32. (One solution to this problem is to provide a larger output screen, an approach covered in the later section, Section 3.2.2 Multi-User Kiosks: Large Display Kiosks).

3.2.1.3 Access to Companions while using a Single-User Kiosk

The form factor of the single-user kiosk has the potential to allow companions to interact in a manner that promotes learning, in that those who do have a view of the output screen can converse with one another about the activity. Parent-child dyads, especially those where the child is still somewhat too young to use the kiosk alone, have been observed to engage in tutorial-style dialogues while engaged with a single-user kiosk (Lyons & Pasek, 2006). This usually works well because the child's head does not impede the parent's view of the output screen, even when seated on the parent's lap. Larger groups, or groups with older members, however, generally run into view obstruction problems.

Another known issue with the social aspects of single-user kiosk usage is the effect of scarcity on group members' social behaviors. As seen in many classroom studies where learners were asked to share a single computer terminal, a not insignificant number of the groups devolved into unproductive squabbling over access to the output screen and access to the input device (Inkpen et al., 1997; S.D. Scott, Mandryk, & Inkpen, 2003). This behavior has been observed by professionals and researchers in museums as well, and it occurs especially frequently among siblings, who are likely to interfere with one another's efforts if feeling "shut out" of I/O access opportunities (Gammon, 1999). Seeing as a major portion of museum visitors are nuclear families (Korn, 1995), this is a nontrivial problem.

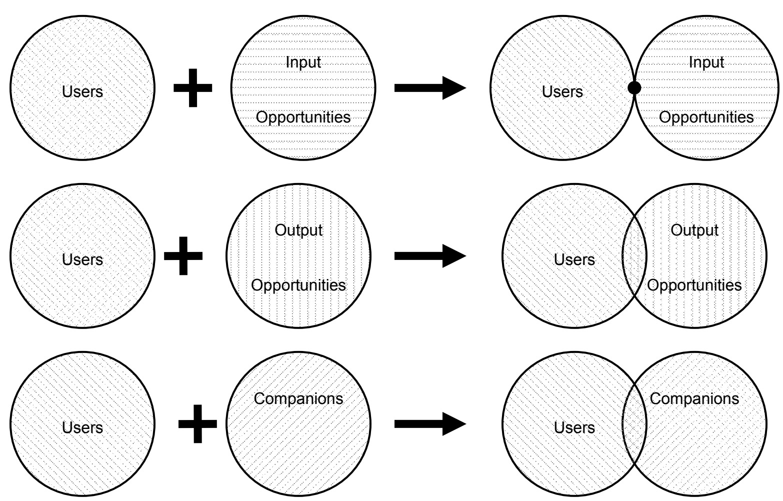


Figure 32. Venn diagram illustrating the access users have to input opportunities, output opportunities, and interaction with companions while using a canonical single-user kiosk. Typically, only one user will be able to provide input (illustrated by the black dot placed at the intersection of User and Input Opportunities). Depending on the size of the output display, several users may have access to Output Opportunities. A user's interaction with his or her companions is usually limited to the number of users who have access to output (i.e., those that remain abreast of the software's state).

3.2.1.4 Single-User Kiosks: a Summary of Group Learning Opportunities

This section has taken a somewhat critical tone towards single-user kiosks, largely because of the lack of support for small-group learning, but in truth, single-user kiosks can be highly successful when used by a single visitor. The ability of a single visitor to have complete control over the input to the system, and complete access to the output of that system, provides the opportunity for that visitor to engage in a highly rewarding dialogue with the software. For example, in a large-scale study of a single-user kiosk centered around a genetic counseling activity, participants who used the kiosk showed significant gains on their understanding of the topic (Bell et al., 1993-1994). When these gains were compared against those of participants who were given the same task, but used informationally-equivalent paper pamphlets in lieu of the kiosk, the kiosk group was the only one to show significant gains. This suggests that when designed well, the interactive dialogue between visitor and content material that is made possible by single-user kiosks can help museum-goers learn (see Section 2.1.2 Interpretation and Authority, for a discussion on the importance of dialogue in museum visitor learning experiences).

The term "dialogue" is used in museum literature, but when translated to a software-based activity, it becomes "interactivity." It seems that the major wrinkle with the use of single-user kiosks for museum education just their inability to support small groups of learners. Research on small groups' use of shared single-user computers in a classroom context has shown that those users who were afforded the greatest interactivity with the system had the greatest learning gains, whereas the more passive users learned less (Inkpen et al., 1997; Mevarech, 1994). There is no reason to believe that this finding wouldn't hold for museum contexts.

Figure 33. Image of a single-user kiosk exhibit in the Wired Worlds exhibit , taken from (Sweeney, 2001). Notice that in this case, the exhibit designers have placed two identical kiosks next to one another, to better support synchronous use by multiple visitors (in this case, dyads).

It is worth mentioning, though, that museums are not insensitive to the need to provide interaction opportunities to visitors, although practitioners most often detect it through the lens of 'throughput'. In other words, if practitioners notice that queues are forming around a particular kiosk, they may provide multiple identical kiosks right next to each other to allow more visitors to have a chance to use the kiosk's activity. Usually, though, this is not thought of as a way to improve group learning experiences – rather, it is an opportunity to move more visitors through the exhibition. One exception is the natural selection simulation at the *Wired Worlds* exhibit (Sweeney, 2001), which appears to allow creatures designed on different kiosks to be released into the same simulated environment (see Figure 33). Space and cost concerns often prevent museums from providing enough kiosks to approach a 1:1 visitor-to-computer ratio, however. An alternative to reaching a 1:1 ratio with desktops would be to provide visitors with smaller, lower-cost computational devices, which will be discussed in Section 3.3 Handheld Devices. Yet another approach to engage more visitors simultaneously is to simply make the display larger, which we will discuss next in Section 3.2.2 Multi-User Kiosks: Large Display Kiosks.

3.2.2 Multi-User Kiosks: Large Display Kiosks in Museums

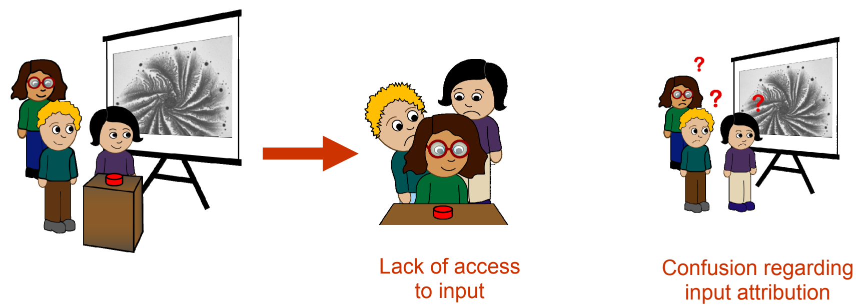


Figure 34. Illustration of the typical presentation of a large-display system in a museum. The images on the right shows how groups of visitors may run into input access problems. The first scenario, if insufficient input opportunities are provided, means that not all visitors will be able to interact with the exhibit. The second scenario, wherein input is aggregated (e.g., noise levels, or movement), may give visitors trouble determining whose individual input actions resulted in changes in the system.

Section 3.2.1 Single-User Kiosks pointed out one major shortcoming of such systems in supporting small group learning: the small size of the kiosk screen. That was not a failure of design, but rather, a lack of opportunity. It is easy to forget that in the early days of software-based exhibits, or "multimedia" exhibits as they were termed in the 1980s, large computer displays were simply not available – and even if they were, the GPU-less computers of that era may not have been powerful enough to drive graphics for them. So quite naturally, as larger graphics displays became available, museums began to incorporate larger displays into exhibits. For example, in one of the early attempts to support small-group use of a kiosk in the early 90s, the video signal for a standard 14" single-user kiosk was split and routed to a 20" monitor mounted directly above it so a user's companions could also see the screen (Diamond et al., 1995). That basic approach – take an activity that could be designed for a single-user kiosk and upgrade it with a much bigger screen – dominates many of the multi-user kiosks found in museums to this day, as we will see next, in Section 3.2.2.1 Input to Multi-User Kiosks.

3.2.2.1 Input to Multi-User Kiosks

Figure 35 The image on the left is a multi-user display in the *Wired Worlds* exhibit , taken from (Sweeney, 2001). Notice the single-user touchpad used for input, which is used to select “email messages” to be sent. The “email message” transmission paths would then be illustrated on the globe on the large shared display. The image on the right is the *Shannon Portal,* designed for use as an exhibit in Shannon Airport, Ireland, taken from (Ciolfi et al., 2007). The large shared screen is in the background, controlled by the touchscreen mounted in the dais in the foreground.

Multi-user kiosks tend to be single-user when it comes to input. A typical multi-user kiosk will usually have only one input “station,” usually a cluster of controls mounted on a table, which allows a visitor to provide input to the display (see Figure 35). Like single-user kiosks controls, these input devices can take many forms: hardware buttons, joysticks, trackballs, keyboards, or touchscreens. As with single-user kiosks, only one person at a time can practically provide input via these means, regardless of how large the screen is. In (Meisner, Craubner, Dech, & Eales, 1999), a room-sized “interactive theater” that presented data related to global warming could be controlled only via a single touchscreen panel. Similarly, a large-display exhibit placed in an airport was controlled via a single touchscreen (Ciolfi et al., 2007).

Many multi-user kiosks offer input methods not commonly seen with single-user kiosks, probably because the input technology evolved as the large display technology did. In one example, a visitor could use a digital pen to act as a “brush” to try their hand at painting in different artistic styles on a large digital canvas (Onda et al., 2004). In another, cameras were set up to detect pointing motions from a user standing at a special spot in front of a large display showing a work of art. The pointing motions are processed to determine where on the display the user is pointing, to provide additional information (Alisi et al., 2005). In both of these cases, though, only one user could paint with the brush, or point to areas of interest, at a time.

In one instance, researchers found it necessary to remove visitors from the multi-user context entirely in order to provide them with input opportunities. The multi-user “kiosk,” in this case, was a Virtual Reality (VR) cave that presented a 3D recreation of Mayan ruins (Tanikawa et al., 2004). The trouble with VR caves is that only one user can control the simulated movement through the 3D environment, and all other users merely view the 3D environment from the perspective of the controlling user. Finding the experience to be too passive for some visitor groups, they constructed a small computer lab immediately outside the VR cave so that each visitor could use a desktop computer to individually navigate through the same 3D virtual recreation.

Figure 36. The *Energy Everywhere* exhibit from the BP-funded Energy gallery at the Science Museum of London, taken from (Viney, 2005). At this exhibit, visitors provide input by moving in different ways (e.g., jumping or waving arms) when prompted to do so by the large display. Because the motion sensors detect don't distinguish between the movements of multiple visitors, it can be hard for a visitor to understand which of their motions results in the changes to the display on the large shared screen.

Designers are not necessarily being shortsighted by limiting input to just a single source. It is more difficult to design and build activities that can interpret multiple concurrent inputs and still present feedback to the users in a coherent manner (Patterson, 1991). One of the most basic rules about user interface design for museum settings is that, when a visitor provides an input of some sort, he or she should receive an instant responding output to acknowledge the input action (Gammon, 1999). When this rule is violated, visitors typically respond either by repeating their input attempt (sometimes by trying the same action faster or with greater force – which often can damage equipment like touch-sensitive screens) or by just giving up entirely. When a system needs to receive multiple inputs from multiple users, however, it can be very hard to unequivocally indicate to each user that his or her input event has been accepted. An example of this is the *Energy Everywhere* exhibit shown in Figure 36, where the motions of visitors are detected to provide input to the interactive display. Because the motion input is aggregated to produce a result (the animations on-screen become more numerous and pronounced with increased overall activity), it is difficult if not impossible for a particular user to be sure of how much he or she is contributing to the overall effect. This is a problem when it comes to helping people learn – a key aspect of any theory of learning is that a learner needs to receive feedback on his or her actions in order to acquire understanding. Feedback on the performance of individuals is especially crucial in group learning scenarios, lest some of the would-be learners become “free-riders,” “bystanders,” or “social loafers” and thus reduce their own opportunity to learn (Hudson & Bruckman, 2004; Slavin, 1992)

Apart from hampering individual learning, a lack of individualized feedback in a group scenario can make the group task itself harder to accomplish. When input is aggregated, learners are only receiving feedback on the group’s actions, and not on their own contributions to the group effort, which can make improved coordination difficult (Slavin, 1992). For example, in the *Shannon Portal* in Figure 35, computer-vision cameras interpret visitor movements to determine where on the large shared screen a fisheye magnification lens will appear (Ciolfi et al., 2007). Because these movements are averaged, if a small group of visitors is standing in a line in front of the display (the best way for all to have an unobstructed view), the lens will just appear in the middle. Knowing which of their companions should move so that the lens will move to a desired location, and in which direction, may not be obvious to visitors – indeed, they may not be aware of the connection between body and lens location at all. The visitors would need to move as one unit to attain a shared viewing goal, presuming they had a shared goal and weren’t moving to try to maximize the satisfaction their own individual interests. Rather than encouraging a multi-partner dialogue, wherein each visitor has his or her own unique connection to the shared display that in turn enriches their dialogue with one another, this setup “flattens” the space of interactional possibilities, and forces visitors to act as if they were one visitor, or risk not accomplishing much at all. Even in museums, a strong emphasis is placed on ensuring that visitors can engage in a “dialogue” (albeit a hazily-defined activity) with the objects and people present in the museum (see Section 2.1.2 Interpretation and Authority).

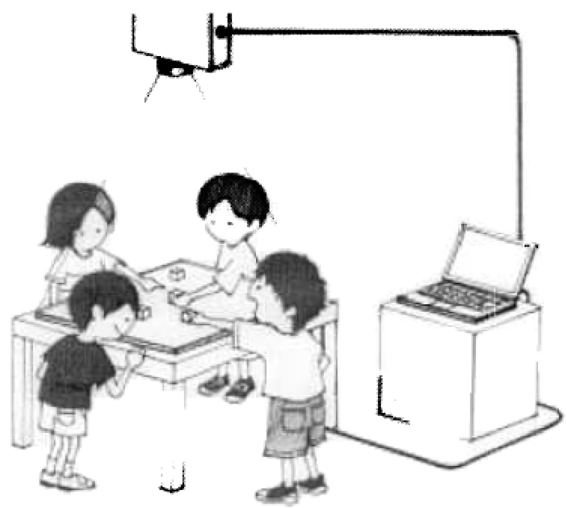


Figure 37.Illustration of a tabletop-based multi-user kiosk, taken from (Kusunoki, Sugimoto, & Hashizume, 2002). Visitors use RFID-equipped objects (similar to the tokens from a game like *Monopoly*) to provide input to the system (in this case, an interactive museum guide).

Designers should indicate, in some way, how each user is affecting the shared display so as to improve visitors’ chances of engaging in the sort of dialogic process that leads to learning. One way would be to represent each user and his or her actions on the shared screen, perhaps through a unique cursor or avatar. In an exhibit on dinosaurs shown on a large projected screen, users were issued different-colored laser pointers, which could be used to “paint” recreations of dinosaurs (Macedonia, 2003). A more tangible approach can be found in the system described in (Kusunoki et al., 2002), where visitors use RFID-augmented physical tokens to interact with a tabletop display (see Figure 37). In this case, the display is an interactive guide to the museum, and visitors can use their tokens to get extra information on exhibits, or play a game where they “hunt” for the animals represented by the exhibits. In this case, the tools used to provide input and the tools used to represent user actions are one and the same, reified into physical objects held by the users themselves, so there is never any confusion about how to tie effects in the output to individuals.

3.2.2.2 Output from Multi-User Kiosks

The term “multi-user kiosk” has been applied rather liberally to mean *any* interactive display that was designed to be shared by multiple simultaneous visitors, and so there is a wide variety of different output displays employed for this purpose in museums. Many have already been touched on – the earliest were just slightly-larger than normal, prominently-mounted computer screens (Diamond et al., 1995). More modern versions by-and-large haven’t strayed too far from that idea, with the *Energy Everywhere* exhibit making use of a large front-projected image(Viney, 2005), and the *Shannon Portal* utilizing a rear-projection screen (Ciolfi et al., 2007). A casual visitor to science museums in the 6 or 7 year span prior to this writing could testify that large screens, whether they be projections or large-format plasma or LCD screens, are the most favored for multi-user kiosks – just about every museum has at least one [13].

There are a few other, much higher-cost, variations on large shared displays that derive more from ideas of theater than from traditional single-user kiosks. Some are structured as theaters, like the “interactive theater” in (Meisner et al., 1999) or the “Digital Earth Theater” that provided laser pens to the audience in (Macedonia, 2003). Unlike all of the prior examples, which only show two-dimensional imagery, the virtual reality cave used in (Tanikawa et al., 2004) immerses visitors in a three-dimensional image space. Although not mentioned explicitly by the authors, 3D caves are limited by the fact that the three-dimensional “perspective” is established with respect to a single point, usually the person with a handheld “joystick” controller. So visitors are not free to truly explore the 3D space – they must experience another person’s choice of vantage points (which is why Tanikawa et al. ended up providing single-user desktops that allowed individuals to fully explore the 3D space shown in the cave).

[13] Documentation could not be found in the literature on this, but in the author’s experiences visiting national and international science museums it is generally true. In the past 7 years the author has attended: the Ann Arbor Hands-On Museum in Ann Arbor, MI; the Exploratorium in San Francisco, CA; the Museum of Science and Industry, and the Field Museum in Chicago, IL; the St. Louis Science Museum in St. Louis, MO; the Smithsonian, Air and Flight Museum, and National Museum of Natural History in Washington D.C.; the American Museum of Natural History in New York, NY; COSI in Toledo, OH; the Ontario Science Center in Toronto, ON, CA; Montreal Science Centre, Montreal, QC, CA; TELUS World of Science, Vancouver, BC, CA; the Science Museum, London, England; Explore-At-Bristol, Bristol, England; Osaka Science Museum, Osaka, Japan; and the Eksperimentarium, Copenhagen, Denmark.

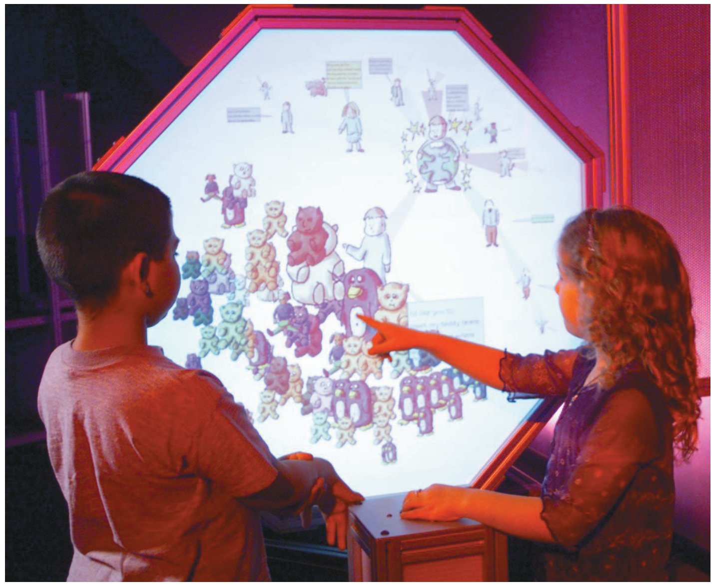


Figure 38. Image of a large display system, the *Hyperbolic Reader*, from the *XFR* exhibit at The Tech museum in San Jose, from (Back et al., 2001; Harrison, Minneman, & Balsamo, 2001). Notice that only one visitor at a time can control the joystick. Even though the main user's attention is on the image she is pointing to on the shared display, her other hand is still completely (possessively?) covering the joystick used to control the cursor. Notice the second visitor's arms – one placed on the control podium – in a mixed gesture, both reaching and restraining.

3.2.2.3 Access to Companions while using Multi-User Kiosks

No studies have been done to explicitly prove that visitor communication is improved when a group of visitors are using a kiosk with a large display in lieu of a kiosk with a small display, perhaps because the results are fairly obvious. As we will see in Chapter 4, on collaborative technologies for work and learning, this correlation between screen size and improved communication has been proven in more formal environments like classrooms and workplaces (K. O'Hara, Perry, Churchill, & Russell, 2003). As far as a museum setting goes, however, the increased screen size of multi-user kiosks certainly makes it much easier for all members of a small group to see the output, which can help them more fully participate in conversation. This is known as conversational "grounding" – when all parties understand what it being referred to in a discussion, collaborative discussion is improved (Clark & Wilkes-Gibbs, 1986). Perhaps another influencing factor is that the manner in which people can gather around larger screens make it easier for them to see each others' faces. We know from communication research that being able to make eye contact is a very important component of conversation (Argyle, 1975). The *Hyperbolic Reader* exhibit depicted in Figure 38 demonstrates how, with a large shared display, visitors are able to position their bodies so as to see their companions' expressions and gestures. It is also worth noticing, however, that the *Hyperbolic Reader* only has one input device – the joystick under the hand of the girl in Figure 38. It can be dangerous to ascribe too much meaning to the body language expressed in a single photograph, but it does appear that the girl on the right was, at the moment the photo was taken, in a more dominant role than her companion. Although there haven't been any documented cases of visitors coming to blows over the controls to a multi-user kiosk, one might reasonably assume that the types of conflicts often seen with single-user kiosks (and especially among siblings) would transfer to a multi-user kiosk scenario wherein input was limited to one person.

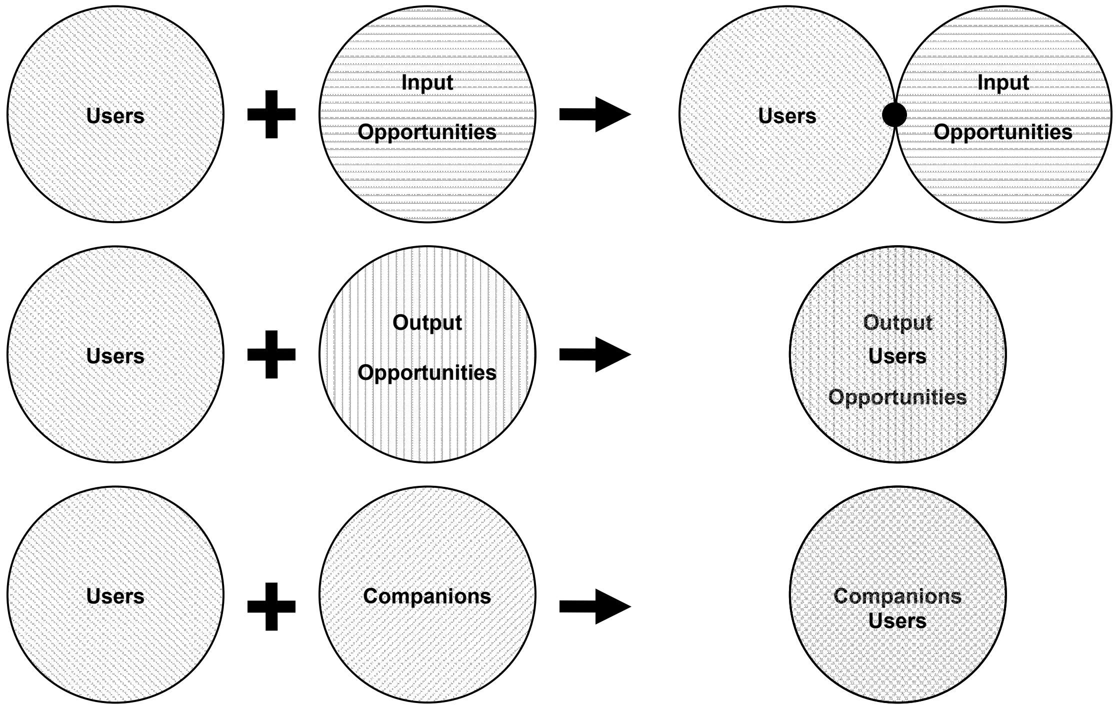


Figure 39. Venn diagram illustrating the access users have to input opportunities, output opportunities, and interaction with companions while using a canonical multi-user kiosk exhibit. As with kiosks, only one user will typically be able to provide input (illustrated by the black dot placed at the intersection of User and Input Opportunities). Assuming that the output display is large enough, all users will have access to Output Opportunities. A user's interaction with his or her Companions is usually unlimited

3.2.2.4 Multi-User Kiosks: a Summary of Group Learning Opportunities

The group learning opportunities for multi-user kiosk are a bit brighter than those for single-user kiosks, if only because the added visibility of the larger displays allows all visitors in a group to be privy to the software's output, as represented by the complete overlap of Users and Output Opportunities in the Venn diagram in Figure 39. Similarly, the larger displays also allow visitors to arrange themselves in space such that they can communicate more easily with one another, and are more likely to do so, as represented by the complete overlap of Users and Companions in the Venn diagram. Another effect of the larger screen is that it expands the social use of the output screen, from a work space to a performance space (Reeves, Benford, O'Malley, & Fraser, 2005). This can increase the number of spectators, who by watching others make use of the software, are engaged in a learning process known as *legitimate peripheral participation* (Lave & Wenger, 1991). This concept will be discussed further in Section 6.3.1 Inception of DBR.

Users of multi-user kiosks may find that, unlike access to output, input access can still a bit tricky to get at. The most common paradigm for multi-user exhibits is to expand the screen size of a single-user kiosk without altering its controls, leading to the same scenario that happens so often with single-user kiosks, where only one user is really able to provide input at a time. This is represented by the tangential intersection of the User and Input categories in the diagram in Figure 39. Although a few creative input mechanisms that allow *every* visitor to provide input have been tried, like using laser pens (Macedonia, 2003) or motion sensors (Viney, 2005), by and large multi-user kiosks still limit the input to one locus of control[14].

The lack of input access problem seen with many multi-user kiosks is probably less a problem with a museum's ability to provide input devices (arcade style buttons are cheap and have been battle-tested for use in museums for years), but rather, a problem with designing activities that support inputs from multiple users in a meaningful and educationally beneficial way. Merely making the screen of a single-user software activity bigger doesn't magically transform it into a multi-user activity. Some of these concerns surrounding input and activity design will be covered in Section 3.3.2 Multi-user Handheld Device Activities in Museums, and again in Section 3.4 Multi-Machine User Interfaces in Museums.

3.3 Handheld Devices in Museums

"Handheld devices" is a blanket term applied here to all mobile computational devices that museum visitors could hold in their hand while at a museum. The most common forms of handheld device found in active us on museum floors are Personal Data Assistants (PDAs) that have been outfitted to act as Audio/Visual guides, a paradigm that will be discussed in Section 3.3.1 Single-User Handheld Device Activities in Museums. Museums are also increasingly becoming interested in allowing visitors to use mobile phones in museums. Sometimes this takes the form of special-purpose promotional exercises, as when Motorola supplied a cache of their new Razr phones to the Chicago Museum of Science and Industry to simultaneously test the phone's potential to act as an A/V guide and to promote their new model (Motorola, 2006). Others are trying to find ways for visitors to use their *own* mobile phones or PDAs (Bruns, Brombach, Zeidler, & Bimber, 2007; Haneef & Ganz, 2002). One form this can take is for museums to post special phone numbers on exhibits that, when dialed, provide audio guide commentary, as the *Science Now, Science Everywhere* program at the Liberty Science Center in New Jersey has done (Bressler, 2005). Other mobile devices, like clamshell PDAs or tablet computers have also been experimented with (Cheverst, Davies, Mitchell, Friday, & Efstratiou, 2000; Hsi, 2003; Tomlinson, 2005), but are generally not well-received by visitors owing to their weight, so this section will largely focus on PDA and mobile phone applications.

[14] This is not strictly true: in Section 3.4 Multi-Machine User Interfaces in Museums several exhibits will be described that make use of either multiple touch screens or handhelds to allow each visitor in a small group to provide input to a shared display. Although the presence of a large shared display makes them candidates for being considered multi-user kiosks, there are enough usability differences (especially those concerning public and private interactions) that they seemed more natural to lump in with MMUIs.

3.3.1 Single-User Handheld Device Activities in Museums

Figure 40.Illustration of the typical presentation of an audio/visual (A/V) guide in a museum: the visitor wears headphones that are jacked into a handheld computer device that has display capabilities (e.g., an LCD screen). The image on the right illustrates that while A/V guides may successfully be used by all members of a group of museum visitors, the form factor discourages interactions between members.

The primary single-user application of handheld devices in museums is in the form of Audio/Visual (A/V) guides, which have inherited many features of the audio guides that came before them. Audio guides have been used in museums since 1957, when portable reel-to-reel players were used to present a prerecorded audio track that lead visitors through a tour of President Franklin D. Roosevelt's home (Acoustiguide, n.d.). With respect to content, the audio track is very similar to what a live docent might say to visitors during a museum tour. The advent of portable audio delivery devices just allowed museums to adapt this pre-existing content to a much more scalable delivery mechanism. Similarly, when handheld computer devices with both audio and visual capabilities became available, museums latched onto the increased modality and began employing them as A/V guides [see (Raptis et al., 2005) for a review of many of the more well-researched A/V guide systems].

As a natural outgrowth of the audio guides that came before, many A/V guides provide essentially the same auditory output as regular audio guides, but are augmented with the addition of visual output in the form of extra text content, images, or video clips. The extra capabilities of these devices meant that a lot of boots-on-the-ground style work needed to be done just to make them functional. Thus most of the published work on A/V guides is more in the realm of proofs-of-concept (Benta, 2005; L.-D. Chou, Lee, Lee, & Chang, 2004; L.-D. Chou, Wu, Ho, Lee , & Chen, 2004; S.-C. Chou, Hsieh, Gandon, & Sadeh, 2005; Fujimoto & Matsuo, 2005; Koshizuka & Sakamura, 2000; Kusunoki et al., 2002; Yamaguchi, Kaji, & Kusunoki, 2005) or prototypes of technological solutions (Abowd et al., 1997; Bruns et al., 2007; Haneef & Ganz, 2002; Huang, Chuang, Chang, & Sandnes, 2007; Okuma, Kourogi, Sakata, & Kurata, 2007; Schwieren & Vossen, 2007; Y. Wang, Yang, Liu, Wang, & Meng, 2007).

Not all research has been into functionality, however – some has been devoted to the usability of A/V guides (Bellotti et al., 2002; Fleck et al., 2002; Hsi, 2002, 2003; Walter, 1996; Wessel & Mayr, 2007). The Exploratorium[15] played an early leading role in inciting interest in the usability of handhelds in museums. The institution hosted two forums that assembled leading researchers and developers to share

[15] The Exploratorium is a hands-on science museum in San Francisco, and is the place chosen to be the site of the *in situ* research study presented in this work.

their understandings of how to design mobile experiences to meet visitor needs (Exploratorium, 2001, 2005). The first of these forums reflected the speculative nature of the enterprise in the early 2000s, as the content was more focused on the many *projected* advantages of mobile devices in museums. For example, the forum attendees generated a list of advantages that included: increased individualization of content, support for differently-abled visitors, support for inter-exhibit meaning-making, extension of learning beyond the visit experience, catalyzation of socialization, extension of sensing abilities, and the expansion of interaction capabilities with exhibits. The second forum, which took place four years later in 2005, reflects a maturation of the field. Many of the broader speculative ideas generated at the 2001 forum had been reframed into questions with more of a traditional HCI focus, like how to design layouts and graphics for output to small screens, how to design activities for different user groups, and how to design handheld applications that allow for socialization.

The wide range of potential applications and their concomitant HCI questions reflect the explosion of interaction possibilities handheld computer devices now made possible. Compared to audio guides, visitors could now receive output in all sorts of different modalities (audio, text, images, and video). Perhaps most significantly, visitors could do something not possible with the original audio guides: provide input. All kinds of input, as the next section will illustrate.

3.3.1.1 Input to Single-User Handheld Device Activities

One of the largest differences between audio and A/V guides is in the degree of input visitors can provide. The earliest audio tours forced visitors to not only move from area to area in a prescribed serial order, but often also forced them to follow it on a prescribed timeline. A/V guides, however, take full advantage of the random-access nature of the underlying technology and allow visitors to choose which content elements they will access, and to choose the times of those accesses.

The means of input are also widely varied. Some systems make use of the hardware buttons on the handheld devices (Aoki et al., 2002; Fujimoto & Matsuo, 2005; Grinter et al., 2002), some make use of touch-sensitive displays by presenting the users with onscreen buttons triggered by a stylus (Aoki et al., 2002; Cheverst et al., 2000; Fleck et al., 2002; Hsi, 2002, 2003; Yamaguchi et al., 2005) or fingertip (Bellotti et al., 2002), some allow visitors to select user interface options by tilting devices equipped with accelerometers (Mantyjarvi, Paternò, Salvador, & Santoro, 2006), many use positional data determined by proximity to wifi or bluetooth hotspots, or active RFID (L.-D. Chou, Lee et al., 2004; L.-D. Chou, Wu et al., 2004; Huang et al., 2007; E. Klopfer, Perry, Squire, & Jan, 2005b; Koshizuka & Sakamura, 2000; Mantyjarvi et al., 2006; Okuma et al., 2007; Schwieren & Vossen, 2007; Y. Wang et al., 2007) some devices can detect glyphs or barcodes with cameras (Koshizuka & Sakamura, 2000; Kenton O'Hara et al., 2007; Wagner, Schmalstieg, & Billinghurst, 2006), and some camera-equipped handhelds even engage in limited image recognition (Albertini, Brunelli, Stock, & Zancanaro, 2005; Bruns et al., 2007). Regardless of the method, each user has access to the means of input, either through hands-on manipulation or movement through space, and so the Venn diagram that relates *users* to *input* shows a complete overlap (see Figure 42).

The majority of these different input modalities are used in the service of determining which object the user wishes to obtain more information on. As one might infer, there are a whole host of usability issues involved in getting each of these different input modalities to work satisfactorily, which is reflected in the literature: large numbers of publications present proofs-of-concepts and technological solutions, and a relatively much smaller number grapple with HCI issues. Without many comparative experiments to report on, wherein different input modalities would be directly and empirically contrasted, all that can be reported on is the zeitgeist. The general trend seems to be to move away from using forms of input that require active manipulation (e.g., using a stylus to provide input), and instead tie input to more passive visitor behaviors (e.g., moving from place to place, or pointing the device at an object of interest). When active manipulation is used (usually in the form of on-screen buttons), attempts have been made to simplify the process as much as possible, by replacing complex visual targets with simpler ones (Fleck et al., 2002), and by making input controls easily triggered by fingers (as opposed to styli) (Bellotti et al., 2002).

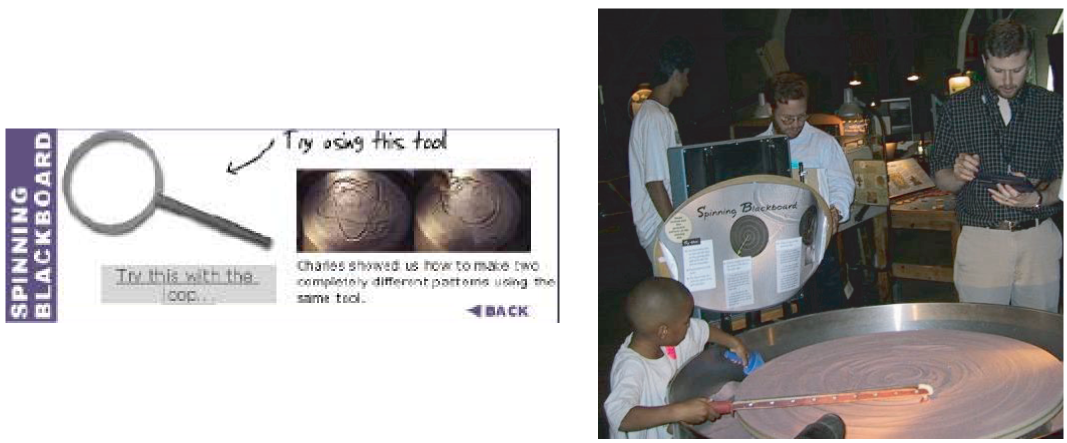


Figure 41. Image on the left depicts the mobile web content presented on the Exploratorium's A/V guides, taken from (Hsi, 2002). This screen suggests ways to use the Spinning Blackboard exhibit by presenting both text suggestions and a demonstrational video. The photograph on the right shows visitors using the Spinning Blackboard exhibit, taken from (Fleck et al., 2002). The visitor on the right is in the middle of consulting his A/V guide.

3.3.1.2 Output from Single-User Handheld Device Activities

Each visitor who has a handheld device (whether it be an A/V guide or other) has complete access to its output,[16] which is illustrated by the complete overlap of the *users* and *output* categories in the Venn diagram in Figure 42. Modern handheld devices support as wide an array of output as regular computers: text, still images, video, and sound are all possible modalities. Many designers for these A/V guides, then, pursue similar design strategies as they would if designing output for a desktop application. In the reams of research done on the educational potential of multimedia *desktop*-based educational applications, advanced multimedia (when properly employed) is found to help with aspects of learning from motivation to

[16] As with the other examples, the potential visual or auditory impairments of visitors are not taken into account. Accessibility for impaired users is a valid line of investigation, but out of the scope of the current research, which pertains to usability for normally-gifted users.

comprehension to retention (Najjar, 1998). In museums, the presence of advanced multimedia has been shown to cause increases in learning measures like recall and visual recognition tasks, as well as in self-reported satisfaction (Bellotti et al., 2002). So, as much as the technology allows, A/V guide designers seek to deliver high-quality video, images, and audio to visitors. Unfortunately, they may be victims of their own success – the "souped-up" A/V guides are *so* successful at engaging visitors, (a) visitors are sometimes reluctant to use them (as will be described shortly), and (b) when visitors do make use of them, their relationships with other visitors suffer (as will be described in Section 3.3.1.3 Access to Companions while engaged in Single-User Handheld Device Activities).

Sherry Hsi of the Exploratorium conducted a study of an A/V guide dubbed the *Electronic Guidebook* (Hsi, 2002, 2003). The aims of the guidebook were similar to those of many other proposed A/V guides: to provide visitors with sophisticated audio and visual content designed to augment visitors' learning experiences with existing physical exhibits in the museum. She discovered (via observational and interview data) that visitors had a reluctance to engage with the handheld device while at the museum (Hsi, 2004; Hsi et al., 2004), a finding that paralleled the observations of an earlier electronic guidebook trial at the Exploratorium (Fleck et al., 2002). Many visitors made comments to the effect that, by using the device, they felt like they were at a remove from fully experiencing the museum – that using the handheld guide interfered with their ability to relate to the exhibits on the floor. While one plausible explanation for that feeling of interference could be ascribed to sheer physical constraints (many of the Exploratorium's exhibits require hands-on manipulation, so the guidebooks directly competed for visitors' "hand time") a similar study of A/V guides in an aquarium setting (where hand manipulation is impossible – all exhibit elements are quite literally behind glass) supported and refined the finding that many visitors found A/V guides to be distracting.

In the aquarium, some visitors appropriated the use of the A/V guides to manually reduce the amount of visual distraction: they chose to listen to just the audio component of videos, but not watch the videos themselves. The researchers found a significant correlation between visitors who only used audio and the length of the A/V guide's use (Bellotti et al., 2002). In other words, when visitors made use of the AV guides in manners that reduced the degree of visual distraction they provided, they were prone to use the devices longer. This is a significant finding because length of use has been one of the "gold standards" used to measure the degree of visitors' satisfaction with informal learning experiences since the inception of visitor studies in the late 1920s (Conn, 1998; McLean, 1999; Roberts, 1997). Another discovery was that self-reported user "enjoyability" correlated with age: apparently older visitors (e.g., ages 25-45) found the use of electronic guidebooks less satisfactory than younger users (Bellotti et al., 2002). So a tentative conclusion that can be reached from this is that some visitors, and especially older visitors, may shy away from engaging fully with handheld AV guides to avoid "missing out" on developing relationships with other elements of the museum context (see Section 2.1.2 Interpretation and Authority for a more detailed discussion of visitors' dialogic relationships with museum elements).

3.3.1.3 Access to Companions while engaged in Single-User Handheld Device Activities

The worries of visitors about "missing out" appear to not be unfounded: those visitors who engaged with AV guides were observed to be much more isolated than the average visitor (Bellotti et al., 2002). At the Exploratorium, a large majority of visitors reported experiencing feelings of isolation, with some also reporting that because of their usage of the device, they were prevented from interacting with their human companions (Hsi, 2003). To quote one visitor: "I didn't really notice other people; I wasn't paying to anybody except for reading the screen" (Hsi, 2002). An earlier study noted that some visitors got "lost in hyperreality," and ceased paying attention to exhibits in favor of the handheld device (Fleck et al., 2002). The ability of handheld devices to usurp visitors' attention to a deleterious degree has been reported in other forums, to the degree where it's been given a name in museum practitioner circles: the "heads-down phenomenon" (Exploratorium, 2005; Wessel & Mayr, 2007).

A telling example is that of the electronic guide to the ancient Roman baths in Bath, England (Walter, 1996). While using A/V guides to the site, many visitors were so intent upon the visualizations of the Roman baths on the screens of their handheld devices that they literally did not look at the actual Roman bath ruins that they were walking past. In this case, the visitor *may* have had a superior individual learning experience by engaging as deeply as he or she did with the handheld device. For example – when comparing aquarium visitors who used AV guides against those that did not on visual recognition and declarative knowledge tasks, AV guide users had significantly superior performance (Bellotti et al., 2002). But if the visitors miss the opportunity to engage with the nominal purpose of their visit (e.g., the bath ruins themselves), and to engage with their human companions, one might make the argument that there is little purpose in attending a physical museum, as opposed to a virtual one, and little purpose in attending with other people, as opposed to *au sole*.

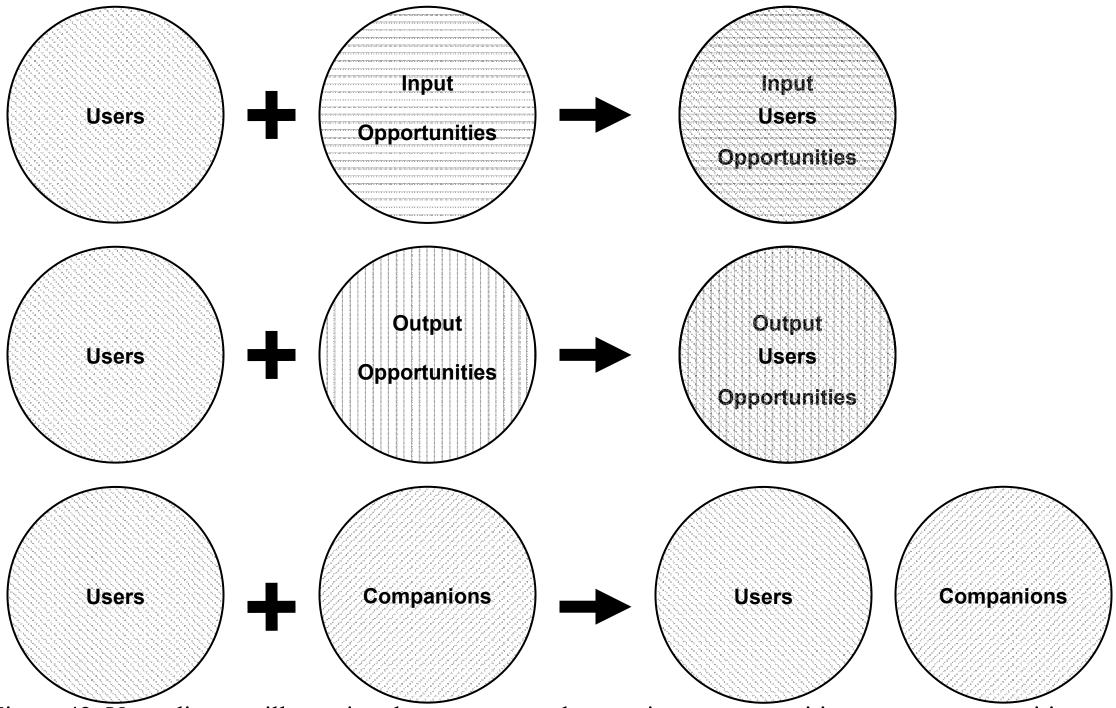


Figure 42. Venn diagram illustrating the access users have to input opportunities, output opportunities, and interaction with companions while using a canonical A/V guide. Users typically have complete access to input and output opportunities, owing to the 1:1 ratio of users to devices. In the canonical A/V guide condition (where users may be wearing headphones), however, users have virtually no interaction with their companions.

3.3.1.4 Single-User Handheld Device Activities: a Summary of Learning Opportunities

The individual choice and control that single-user A/V guides provide should, theoretically, be superior in promoting learning vis-à-vis the older, choice- and control-free audio guides. Both educational literature and literature on informal learning environments (see Section 2.1.2 Interpretation and Authority for a relevant summary) offer both theoretical and empirical support for the importance of a give-and-take dialogue between a learner and an informational resource, whether that resource be another person (Collins, 2006; Sawyer, 2006), a book (Palincsar & Ladewski, 2006), a digital resource (E.A. Davis, 2003; Edelson, Gordin, & Pea, 1999; Ketelhut, 2007; Linn, Davis, & Bell, 2004; Chris Quintana, Shin, Norris, & Soloway, 2006), or even a static object (Hooper-Greenhill, 2000; Scott. G. Paris & Hapgood, 2002; Taborsky, 1990). A/V guides certainly have the affordances to allow for such dialogic learning. By distributing them in a 1:1 fashion to group members, A/V guides also support equal opportunities for learning across all group members – no one in the group is excluded from input and output opportunities (see Figure 42). The problem with A/V guides is not one of a *lack* of opportunity for learning – it is a much more subtle problem emerging from the *success* of the devices at encouraging highly-dialogic relationships.

User interface designers, then, are given a tricky task: to design A/V guide interfaces that encourage visitors to become engaged in a dialogic relationship with the device to a degree that promotes learning, but not to such a degree that visitors become isolated from the "real" aspects of their museum visit. As some researchers have noted, the visitor's attention is the critical resource (Fleck et al., 2002), and an ideal solution would be one that encourages a visitor to distribute his or her attention across all elements of their museum visit context: the physical elements of the museum, their human companions, and their handheld device. This distribution of attention has been termed "back-and-forthing," referencing the way visitors will shift their visual attention back and forth between the handheld device and other elements (Exploratorium, 2005).

No one has explicitly studied which design factors might encourage back-and-forthing, although some recommendations have emerged from studies focused on other elements of A/V guide design. These recommendations tend to fall into two different camps, which will be emphasized because they will become relevant to the experiment design presented in Chapter 6. The first revolves around different types of *simplifications* that can be made to the device's user interfaces. To encourage back-and-forthing, some researchers recommend removing the audio component (Bellotti et al., 2002; Hsi, 2002, 2003; Wessel & Mayr, 2007), while others recommend creating "simple" graphical user interfaces that, presumably, are unlikely to capture a visitor's attention for too long (Bellotti et al., 2002; Fleck et al., 2002; Yatani et al., 2004). The other camp involves changes that can be made to the *activity structure* – what tasks the users are engaged in, and how those tasks are configured – to encourage if not outright require that users interact with other elements of the context. Rather than the individualistic, self-guided activity structure typically found with A/V guides, the activity associated with the use of handheld devices may require visitors to actively integrate the objects and people present with them in the museum into their digitally-guided experiences, an approach that will now be reviewed.

3.3.2 Multi-user Handheld Device Activities in Museums

Figure 43. Illustration of a multi-user handheld device activity in a museum. At first glance, it is very similar to single-user handheld device activities, but visitors do not wear headphones, and visitors are free to converse with one another.

Section 3.2.2 Multi-User Kiosks: Large Display Kiosks brought up the point that merely allowing more people to use what as originally designed as a single-user activity does not transform it into a multi-user activity, at least not one that necessarily carries with it benefits for collaborative learning. The same hold true for handheld device activities. One approach, which doesn't go very far to alleviate the isolation problems seen with single-user A/V guides, allows visitors to use their A/V guide to leave a "message" at an exhibit that can later be read by another visitor (L.-D. Chou, Lee et al., 2004). A slightly more proactive version allows visitors to "find" their companions current locations in the museum (S.-C. Chou et al., 2005). Neither strategy really encourages ongoing, active group learning, however. Section 3.3.1.4 Single-User Handheld Device Activities: a Summary of Learning Opportunities presented two alternate strategies for supporting small groups in their use of handheld devices: (1) simplifying the user interface so the visitors will have more attention to give to their companions, and (2) designing the handheld-mediated activities to explicitly be used by small groups. This section will discuss the latter approach.

Some researchers have tried to adapt single-user A/V guides to more explicitly support small group use. To make A/V guide experiences less socially isolating, the *Sotto Voce* system developed at Xerox PARC was designed to allow visitors to "listen in" on the audio their companions are listening to (Aoki et al., 2002; Grinter et al., 2002; Woodruff, Szymanski, Aoki, & Hurst, 2001). The idea behind this system is that, although visitors may not interact with each other directly, they can still find out what their companions are "up to" by listening in on their currently-playing audio clips. This allows what would otherwise be an isolating individual activity to be shared by companions.

A final approach to redesigning A/V guides to be used by groups was devised by a group of researchers who worked on the PEACH project (Personal Experience with Active Cultural Heritage - a large-scale multifaceted investigation of the use of technology in cultural museums in Italy (Stock & Zancanaro, 2007). They proposed an A/V system that would encourage visitor-visitor communication by delivering different information to each participant (Kruppa, Lum, Niu, & Weinel, 2005). Users provide their personal informational preferences ahead of time, and when a group reached an exhibit, each would be delivered information about that exhibit that is tailored to their particular interests. The underlying theory is that because the visitors are each privy to different facets of information, they may feel a desire to share their "special" information with their companions, or discover what their companions are learning about. This is a form of "jigsawing," a strategy often used in education to encourage collaborative learning (Aronson, Blaney, Stephan, Sikes, & Snapp, 1978; Cuthbert, 1999; Dillenbourg, 1999; Slavin, 1992). By providing differentiated information to participants, they are empowered as "experts" in their areas, and motivated to help broaden their companions' understanding.

Other researchers take a more proactive stance towards structuring small-group interactions with handheld devices by structuring their interactions in the form of a game (E. Klopfer, Perry et al., 2005b; Kusunoki et al., 2002; Thom-Santelli, Boehner, Gay, & Hembrooke, 2006; Wagner et al., 2006; Yatani et al., 2004). This is not an unreasonable approach: museum visitors are, after all, in search of fun and social experiences (Falk & Dierking, 1992). Moreover, many game theory researchers would claim that any multi-agent interaction, even between friends and family, can, at its root, be broken down into a game-like structure complete with its own "Players-Actions-Payoffs" tuple (Rasmusen, 2006). The key, then, is in devising actions and rewards that will engage a group of visitors. Of course, because a museum visit is under time constraints not seen when, say, a family sits down to a game of *Axis and Allies*[17], designers must also be careful to devise a user interface that allows visitors to *easily* become engaged in the game-like

[17] A strategy board game with complicated rules developed by Nova Game Designs, now distributed by Hasbro.

activity structure. This means that designers need to be sensitive to how the devices' input and output will fit into the structure of the game.

3.3.2.1 Input to Multi-User Handheld Device Activities

Physically, many multi-user activities make use of some of the same types of inputs used by single-user activities presented in Section 3.3.1.1 Input to Single-User Handheld Device Activities. All of the systems reviewed in this section made use of touch-screen interfaces, although some augmented this form of input by other means, like RFID tags mounted on exhibits as inputs (E. Klopfer, Perry et al., 2005b; Yatani et al., 2004), or glyphs (Wagner et al., 2006). The system that proposed to deliver different information to different visitors must have its users enter their preferences as input at some point, but the manner of doing so was not discussed, nor was the manner by which the devices would discover proximity to exhibits (Kruppa et al., 2005).

Figure 44. Photograph of the main user interface of the collaborative quiz game, from (Yatani et al., 2004). Each square in the grid corresponds to a quiz question for an exhibit. When answered correctly, the square disappears to reveal a portion of the picture hidden underneath (a). Questions that haven't been answered yet are in white (b), and questions that were answered incorrectly (and thus are no longer accessible) are in grey (c).

The more interesting facet is how, in the game-oriented activities, the inputs are used. In (Yatani et al., 2004), visitors had a grid of squares on their handheld screen, each of which corresponded to an exhibit (see Figure 44). By scanning RFID tags posted next to the exhibits, the squares could be "unlocked" to expose quiz-style multiple-choice questions relating to the exhibit. Similarly, in the *Mystery at the Museum* activity, visitors would use RFID tags and their device's proximity to wifi beacons to discover clues, collect virtual "items," and extract text-based "interviews" from virtual suspects while trying to solve a mystery (E. Klopfer, Perry et al., 2005b). Unlike traditional A/V guides, in both of the aforementioned cases the visitors are not providing input in order to browse information; rather, the input opportunities have been converted into collectable items. With *Mystery at the Museum*, special abilities were sometimes encoded into some of the RFID tags as well – by scanning them visitors would summon up a representation of a control panel for a virtual piece of equipment (e.g., an electron scanning microscope) that would allow users to further examine the virtual objects they had "collected."

Judging by the overwhelming prevalence of "item collection" and "power ups" as themes in popular games (e.g., card games like *Magic: The Gathering* and *Pokemon*, and legions of video games – perhaps best exemplified by the many power-ups and coin boxes in the *Super Mario Bros.* franchise) this is quite likely a fine strategy for encouraging participation. Also, because the input opportunities are localized to specific exhibits, the strategy *may* encourage visitors to engage in more "back-and-forthing" (see Section 3.3.1.4 Single-User Handheld Device Activities: a Summary of Learning Opportunities) as they pay some attention to those exhibits, somewhat counteracting the heads-down phenomenon.

Figure 45. The photograph on the left shows visitors using camera-equipped handhelds to view an Augmented Reality (AR) game, from (Wagner et al., 2006). The image on the right is a screen shot from one of the AR handhelds showing how various cultural objects are rendered in front of the glyphs.

One game is in fact specifically designed to *require* that visitors look at their handhelds, and only their handhelds – the ultimate in "heads-down." The handhelds are equipped with cameras, which are used to detect black-and-white glyphs mounted on the walls of a gallery (see Figure 45). There are no other objects in the gallery – visitors must use their handheld screens to see the Augmented Reality (AR) "objects," rendered in 3D, that are associated with the glyphs.

3.3.2.2 Output from Multi-User Handheld Device Activities

The output modalities used for multi-user handheld device activities, much like their input modalities, are not too dissimilar from single-user handheld device activities. Text, imagery, and animation are all present. Sound is used to deliver information in only two cases, with the "sanctioned eavesdropping" of the aforementioned *Sotto Voce* system, but it seems to be conspicuously absent from the other systems reviewed here, aside from the culturally-appropriate background music used in (Thom-Santelli et al., 2006). None of the authors state so explicitly, but the probable reason for the lack of audio output is to avoid interfering with verbal communication between visitors. The collaborative quiz game goes so far as to supply pairs of visitors with headphones – not to listen to audio recordings, but to be able to listen, *exclusively*, to one another via a radio transmitter (Yatani et al., 2004).

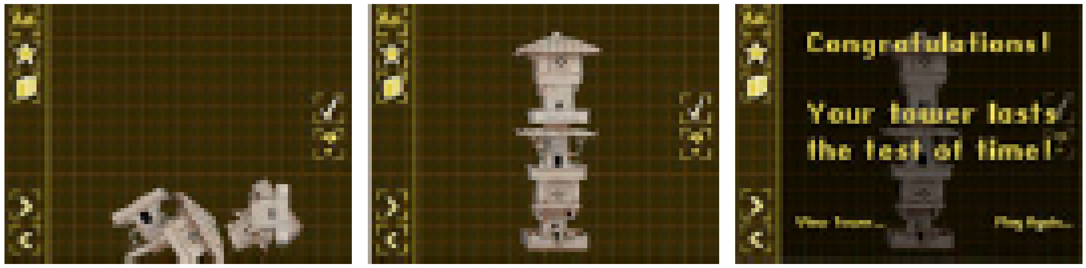


Figure 46. A series of images representing one of the exhibit-relevant "mini-games" from (Thom-Santelli et al., 2006). The activity begins with a jumble of pieces, as shown on the left, that the visitor must assemble to form a re-creation of a model present in the museum, shown in the center.

The visual output for the multi-user handheld activities generally tends towards less dynamic modalities (e.g., static images in lieu of animations, text in lieu of spoken audio), and more "simplistic" user interfaces. The authors, once again, do not make any mention of whether or not they employed less-complex interfaces as a deliberate design strategy, but it seems sensible to suppose that with less of a visitor's attention drawn by his or her device, there is more attention to be given to his or her companions. In the case of the collaborative quiz game, the visual output even *encourages* a visitor to pay attention to his or her companion. By looking at the grid on his or her handheld display, a user can assess the group's progress by inspecting the number of untried, transparent, and permanently opaque squares – it may look simple, but it is in fact an implicit "group status" screen (Yatani et al., 2004).

It may be telling that one application to use highly-dynamic interfaces, in the form of interactive "mini-games" (see Figure 46), was also the one study that reported instances where the visitors would get lost in the heads-down phenomenon, forsaking the actual exhibits for their associated handheld-based activities (Thom-Santelli et al., 2006). The other application that employed a highly-dynamic interface, the AR exhibit from (Wagner et al., 2006), required visitors to give all of their visual attention to the handheld devices – the exhibit's "objects" were rendered *only* on the handheld screens (see Figure 45).

It is also worth noting that, unlike the single-user activities of Section 3.3.1, not all users of a multi-user activity have complete 1:1 access to the output of the activity. If the activity is designed to be distributed across multiple devices, visitors may each have access to only a portion of the output made available to the entire group. How they go about trying to get complete access in these scenarios will be discussed in the next section, as it involves interactions with companions.

3.3.2.3 Access to Companions while Engaged in Multi-User Handheld Device Activities

All of the systems presented here were designed to allow, if not encourage or require, visitors to have connections with one another. Some approaches were more successful at engendering visitor-visitor dialogue than others, however. Those that had audio narrations, such as the *Sotto Voce* system (Grinter et al., 2002; Woodruff et al., 2001) and the AR exhibit (Wagner et al., 2006), made it hard for visitors to converse with one another while listening to the audio tour information, as the visitors had to wait for the audio to cease before conversing. The *Sotto Voce* system deliberately made use of single-ear headphones, and the AR system just made use of the PDAs' native speakers, to try to ensure that headphones wouldn't isolate visitors from one another. The lesson here is that the presence of audio narration, and not just the presence of headphones, impedes collaborative interactions. The difficulty seems to derive from audio narration's serial (as opposed to random-access) nature, and its propensity to "commandeer" the main communication channel (audio) that visitors use to communicate with one another.

In the one system that *only* presented content via audio, researchers also found that the visitors would engage in "asymmetric adoption" of the technology – some visitors, especially the tech-savvy ones, would "lead" the activity by being the person to select which audio tracks to listen to, while their partner would assume a passive, listening-only role (Grinter et al., 2002; Woodruff et al., 2001). The problem with this is that people tend to learn more when they are actively engaged in their learning, as opposed to passively engaged (O'Donnell & Dansereau, 1992). So what we learn from this is that, even when a handheld-based activity provides "symmetry of action" (Dillenbourg, 1999), meaning that all learners have an equal opportunity to participate, visitors may still need to be encouraged (or required) to enact that symmetry of action.

Yatani and his colleagues more actively encouraged their visitor dyads to work together by going beyond enabling, by *enforcing* a joint outcome: if the visitors answered a multiple-choice question correctly, the corresponding square would be made transparent on both visitors' handheld displays (Yatani et al., 2004). Conversely, if either visitor answered incorrectly, the square would become permanently opaque (see Figure 44). A pilot test was performed with 25 pairs of visitors, and they indeed found that some of the pairs were motivated to discuss the questions with each other, and to discuss the larger strategic picture (e.g., reviewing which questions have been tackled, and which should be approached next). Because most of the data reported on was qualitative in nature, however, and derived from written questionnaires, it is difficult to discern how and to what extent the user interface design may or may not have encouraged cooperative behaviors.

Klopfer and his colleagues, like Yatani et al., also chose to adopt a joint-outcome problem-solving approach with their *Mystery at the Museum* game (E. Klopfer, Perry et al., 2005b). In their activity, groups of six visitors were organized into two teams of three, and were asked to compete to try to solve a "theft" at an art museum. Visitors used handhelds that were able to read RFID tags posted throughout the museum to collect "clues," and were able to use their proximity to Wifi hotspots to trigger the ability to "interview" virtual characters that were "present" in different rooms of the museum. Visitors had to use the clues and interviews to help narrow down which character of a fictional band of thieves was responsible for the crime, an activity very similar to that found in the classic educational game, *Where in the World is Carmen Sandiego?* (Brøderbund, 1985). Further encouraging interaction, each player would assume a role (e.g., a detective, or a biologist) which would limit the virtual equipment they were allowed to use, and would alter the output they received when "interviewing" a virtual character. Each visitor, then, would be able to collect unique information (e.g., electron scan results, or interview text), which they could then "beam" to their companions' devices via the IrDA ports. In terms of the activity design, then, the *Mystery at the*

Museum game can be thought of a jigsawing exercise, wherein each collaborator brings a unique ability or area of expertise to the shared task (Aronson et al., 1978).

Klopfer et al. performed a pilot study with around 40 participants total, and found via interview and survey data that the Mystery at the Museum game was successful at encouraging groups to both talk and work together (E. Klopfer, Perry et al., 2005b). Each player felt that they played a significant part in solving the mystery in the game, probably owing to the unique contribution their role allowed them to bring. The researchers did not explicitly study the effect of different user interface elements on collaborative behaviors, although they did note that players would beam information to one another when trying to solve particularly difficult problems. Groups would also often "stick together," even though the activity was designed to allow people to roam at will, because they reported it easier to work together that way, probably because they didn't have to spend as much time bringing each other up to speed. It seems that the players were trying to engage in a "grounding" process, wherein all members of the group would be made aware of the referents that the conversation would center around, a key component to effective collaborative talk (Clark & Wilkes-Gibbs, 1986). This also seemed to happen during the AR game in (Wagner et al., 2006). Visitors reported in interviews that it was difficult to collaborate using the AR PDAs because they had a hard time showing each other the AR representations on their PDA screens. Clearly, being able to share referents was an important part of the collaboration process in both of these game-like activities. Unfortunately, it is not clear that either of the information-sharing strategies of these two activities were adequate to the task – how to support grounding during a collaborative multi-user handheld-based activity is still an open question.

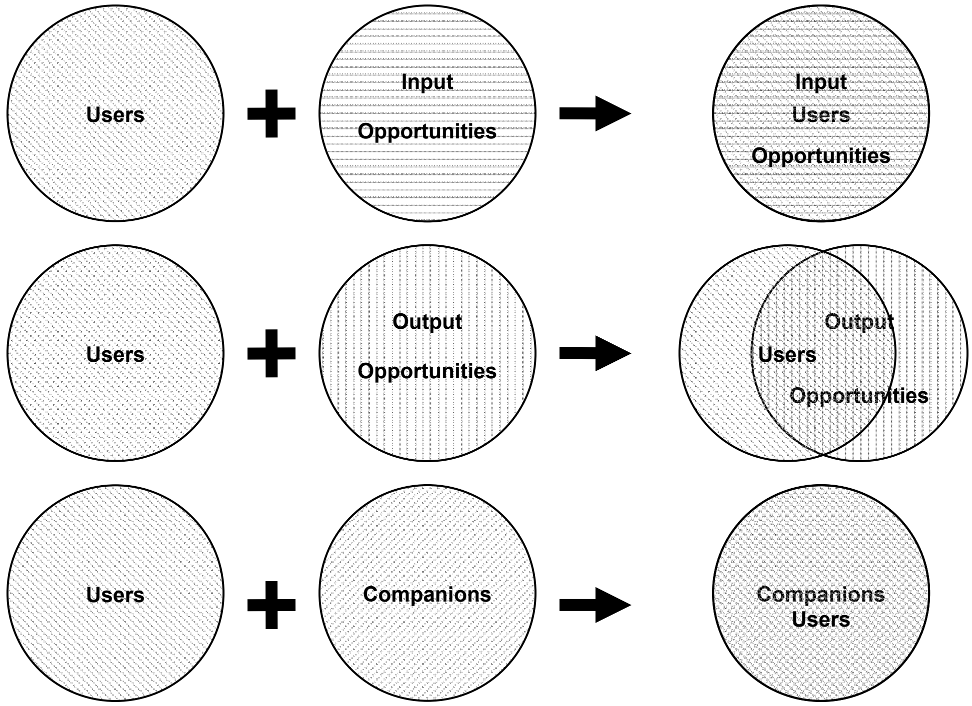


Figure 47. Venn diagram illustrating the access users have to input opportunities, output opportunities, and interaction with companions while engaging in a canonical multi-user handheld activity. Assuming that there is a 1:1 user:device ratio, a User should have complete access to Input Opportunities. Because multi-user activities are distributed across multiple devices, however, users will only have full access to output opportunities if each and every handheld device screen makes the same information available. A user's interaction with his or her Companions is usually unlimited, assuming there aren't any physical obstructions and the users are not wearing headphones.

3.3.2.4 Multi-User Handheld Device Activities: a Summary of Group Learning Opportunities

Multi-User Handheld Device Activities, like their Single-User counterparts, have the advantage of allowing visitors to have 1:1 access to input opportunities, as shown in Figure 47. They also have the potential to provide 1:1 access to output opportunities, as single-user activities do, but users may not have complete access in activities where the information to be displayed is distributed across several devices. (We saw this in the *Mystery at the Museum* activity when users had to beam information to one another's' PDAs to achieve equal access to the uncovered clues). The review of existing multi-user activities shows us that, while in theory visitors will also have complete access to their companions, in practice, different design elements can negatively impact collaborative processes.

First and foremost, narrative audio prevents visitors from conversing with each other (at least until the clip has finished playing), so designers should be wary of using it when collaboration is a goal. Secondly, although the 1:1 distribution of devices *should* encourage all members of a small group to engage equally with the activity ("symmetry of action"), some visitors may not do so unless the structure of the handheld-based activity encourages or requires it. One good way to get all group members participating is to use a game-like activity structure. Activity structures – game-like or not – can also be used to encourage or require visitors to interact with one another. One strategy to encourage interaction is jigsawing, wherein each visitor is given special information or capabilities that the other visitors do not have access to. Another is to establish a joint outcome: if the outcome for the entire group relies on the actions of each and every member, individuals who would otherwise be passive are motivated to participate (Hudson & Bruckman, 2004; Slavin, 1992).

Visitors may still need extra support even after they've been suitably motivated to work together. In several instances, visitors have been reported trying, with varying degrees of success, to establish a common ground to help them communicate, by beaming each other information, or just showing each other their PDA interfaces. The more intense the current collaborative problem, the more they seem to try to establish common referents in these manners. It is unknown whether or not the methods presented here for establishing common referents are adequate; they may not be.

Finally, much as visitors enmeshed in single-user handheld activities run the risk of getting caught in the "heads-down phenomena," so are visitors employing multi-user applications. The methods to ameliorate this risk aren't entirely proven, but some have designed activities that tie very closely to the physical exhibits in museums in an effort to get visitors to attend to them. Earlier, this was dubbed the "reification" approach. Another method may be to actually require that visitors gather information about an exhibit (by observing it or reading its label) to complete the handheld-based task. A third approach may be just to design "simpler," less-engrossing user interfaces. There was not enough research presented here to truly define a design pattern for what a "simple" interface might be, but it seems sensible to assume that the more simple the dialogue with the device, the more room there is for interactions with other objects and people. With such an interface, the outputs would probably also need to be simplistic: short segments of text or images, and little to no dynamic elements like animation. If possible, something about the output should serve as a "status indicator" of the group's overall activities or progress to help tie the visitor to the larger context. The input to "simpler" interfaces should probably also be less effortful, requiring minimal amounts of manipulation.

3.4 Multi-Machine User Interfaces in Museums

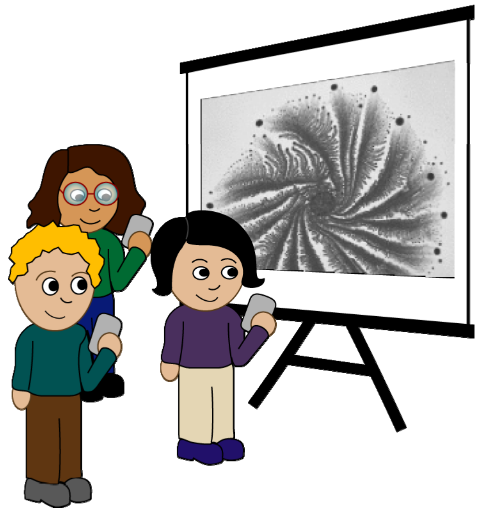


Figure 48. Illustration of one implementation style for a multi-machine user interface in a museum. Here, visitors use handheld devices that are connected wirelessly to a shared display, but alternative implementations could make use of other devices (such as touch-screen displays) to give users access to input and output.

Multi-machine user interfaces (Brad A. Myers, 2001) are systems of interconnected computational devices, engaged in the same application space, that allow groups of co-located users synchronously interact with that shared application space. The vague nature of the name, "multi-machine user interface," implies that the specific hardware used for these systems is far from being firmly determined. The term was originally coined to describe a system wherein PDAs were networked to a large shared display driven by a desktop computer, however, which colors the decision of what types of multi-device systems should be classified as MMUIs. For the purposes of this research, the MMUI label will be applied to any system that provides individual GUI interfaces on a 1:1 basis to users, which are in turn able to influence one or more shared displays (see Figure 48). Another important characteristic is that the individual interfaces to a MMUI are "private," meaning that user interactions with the devices are not necessarily a part of the "performance space" of the exhibit (Reeves et al., 2005). The shared display(s), on the other hand, are avowedly "public" in their nature, designed so that they are visible to all users, and perhaps to other visitors passing by. Moreover, the assumption is that the shared display(s) are used much as the large, shared displays of Section 3.2.2 are, to provide a gestalt view of the shared activity and serve as a tool for grounding the conversation and actions of the group members.

The moment when MMUIs began to be used for museum exhibits is not entirely clear, but researchers have directly or indirectly acknowledged the occasional need for museum visitor groups to have more input and output access opportunities than the kiosk and single-user handheld applications reviewed earlier in this chapter provide. For example, even though they were using the ultimate in terms of large shared displays – a virtual reality cave – (Tanikawa et al., 2004) found it necessary to also provide banks of networked desktop computers to allow small groups of visitors to jointly explore the VR space.

Tanikawa et al. didn’t describe their solution as a MMUI, even though it could be considered such. It may be the case that multi-machine user interfaces are employed in museums more frequently than the amount of active research on them would suggest.

Figure 49. A picture of the *Virtual Fishtank* exhibit taken from (Nearlife), which is installed at both the Museum of Science in Boston and the St. Louis Science Center. The large displays mounted on the wall together display a simulated fish tank. Visitors provide input by using the kiosks in front of the large shared display to design a fish. The fish will then be “released” into the fish tank simulation displayed on the shared display. They may also use the kiosks to tweak other parameters of the simulation: e.g., the rate at which bubbles are emitted by the treasure chest.

Owing to the affordability of touch-screen displays and the accepted wisdom that supporting groups of visitors is better than supporting just individuals, more and more exhibits like the *Virtual Fishtank* depicted in Figure 49 have been appearing (relatively unheralded) in museums since the early oughts. The usual operating principle behind MMUI setups found in museums seems to be that separate touch-screen displays (or other devices) are used to provide input to a shared multimedia application. In the case of Figure 49, this shared multimedia application is an ongoing dynamic simulation of a fish tank, whereas in the application shown in Figure 50, the shared multimedia application is an audio-visual presentation (educational movie clips, essentially). This section will discuss the access (to input, output, and companions) that these MMUI systems provide.

Figure 50. A picture of a PEACH (Personal Experience with Active Cultural Heritage) system during a trial. Museum visitors rent handheld PDAs which they use to interact with stationary flat-screen displays. Taken from (Kruppa & Aslan, 2005).

3.4.1 Input to Multi-Machine User Interfaces in Museums

Any manner of devices can be used to provide input to an MMUI, although the term "machine" implies that each component of the distributed system will have some computational abilities of its own. This point is worth stressing, because some HCI researchers of multi-user, multi-*device* systems see the blanket term "device" embracing not just computationally-capable devices like PDAs, but more standard items like keyboards and mice (Kray, Wasinger, & Kortuem, 2004). Thus we will not discuss here any systems that use multiple mice or keyboards or the like to provide input opportunities to multiple visitors (anecdotally, there has never been a multi-mouse or multi-keyboard computer-based exhibit in any of the museums that this author has visited). The two most common means of providing input to MMUIs in museums are via touchscreens[18] (see Figure 49) or via handheld devices (see Figure 50).

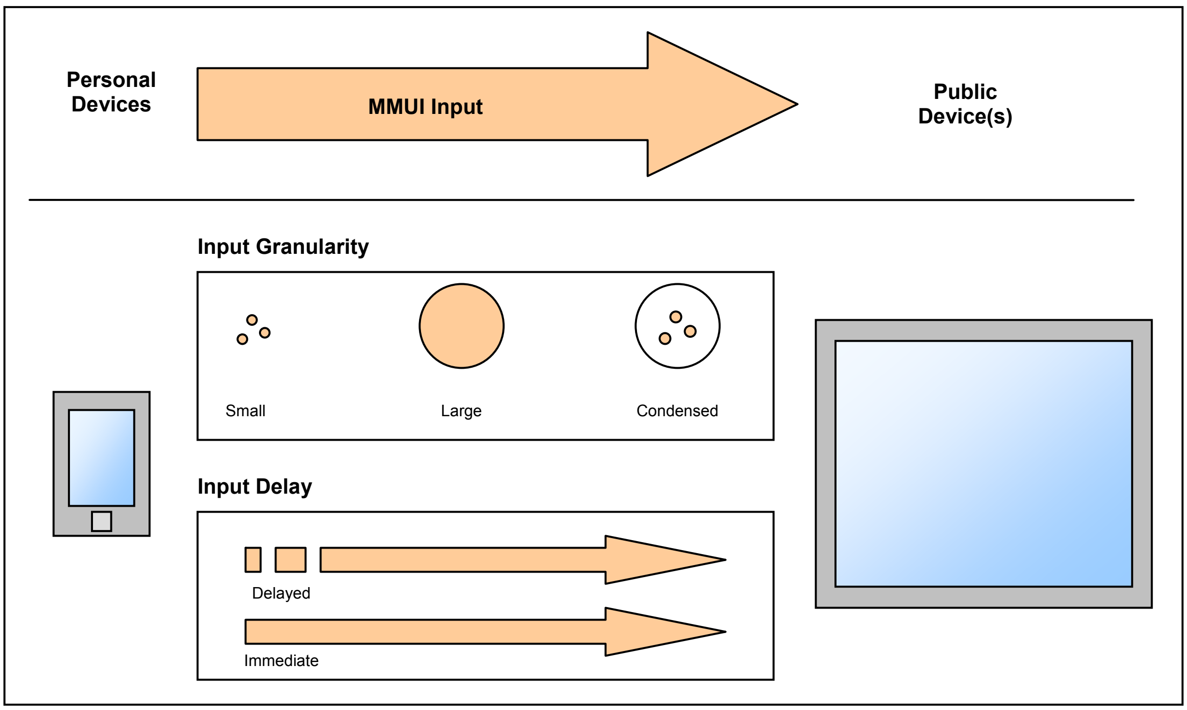


Figure 51. MMUI input is generally transmitted from the personal devices operated by individual users to a public shared device or devices. For the purposes of MMUI activity design, the input can be described along dimensions of granularity (how much information is contained in each input event vis-à-vis the number of manipulations a user performs in order to encode that information) and delay (how long it takes a user's input to reach the shared device).

One advantage of using computationally-capable machines as input devices is that they serve as a sort of "personal" interaction space, in opposition to the "public" interaction spaces that all visitors have equal access to. This adds an interesting design wrinkle, as input events entered via the "personal" spaces

[18] The touchscreen displays in these MMUIs are usually driven by their own desktop computers, and linked together via networking. While it is possible for a single computer to drive multiple displays (a la the mainframe/terminal paradigm), visitors (especially those who came of age during the personal computer era) tend to have a "1 display: 1 computer" mentality, and would thus perceive a mainframe/terminal setup as if it were a MMUI.

can affect the "public" space on varying schedules, from immediate to delayed. The input events can also have different levels of granularity as they move from the "private" to "public" spheres (see Figure 51). So, for example, in the *Virtual Fishtank* exhibit mentioned earlier, visitors spend quite a bit of time providing input solely to the "personal" device as they tweak the design of their virtual fish. This "input" is only conveyed to the "public" space after the visitor has "signed off" on his or her design efforts, which can result in a several minute delay between the onset of input activity and the point in time when that input gets relayed to the "public" space. Also, the input provided to the "private" device is of a much finer granularity (e.g., specific fin or jaw selections) than the input which eventually gets relayed to the "public" device (e.g., an entire fish). The "private" device condenses all of the user's input decisions into a single input event.

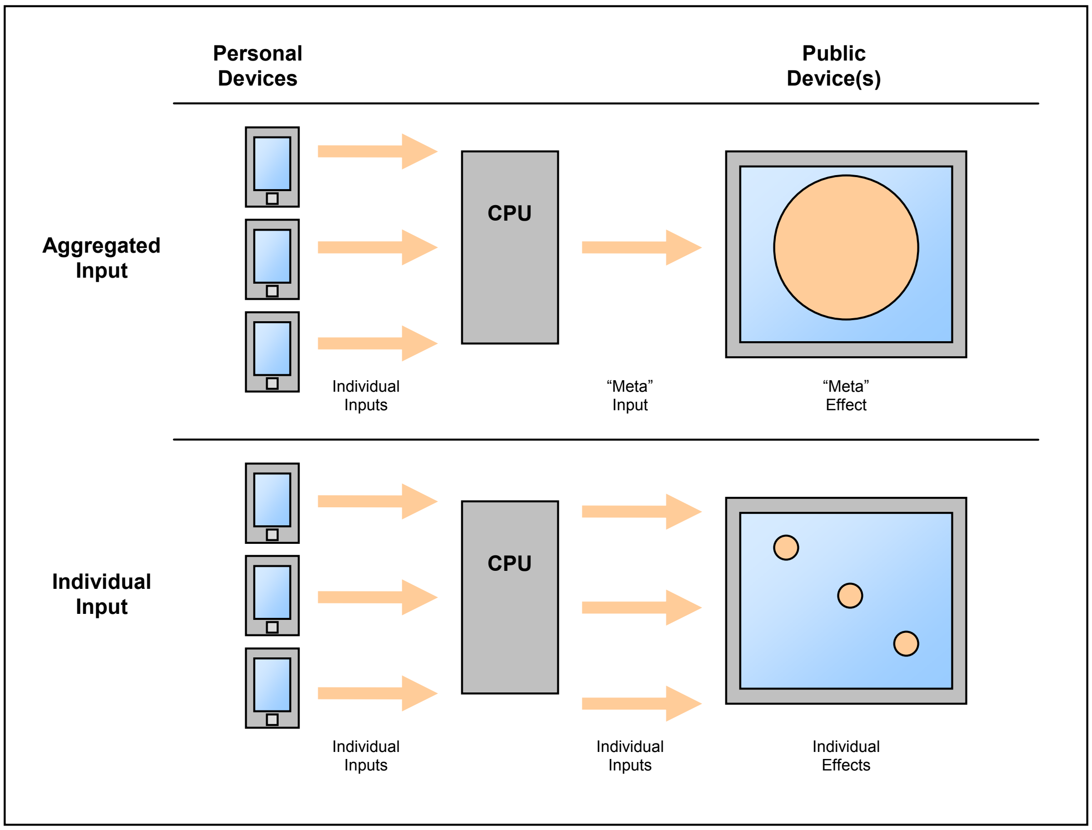


Figure 52. The input from multiple users to a MMUI can be preprocessed into an aggregate form to, effectively, act as a single "meta" input event, which produces a single "meta" effect on the overall state of the shared activity. Alternatively, the inputs from multiple users can each be processed individually, and each be responsible for a distinct change in the shared activity.

A MMUI system can process the input events originating from individual users in several different manners. As mentioned in Section 3.2.2.1 Input to Multi-User Kiosks, it can be very hard to design activities that accept inputs from multiple users in a manner that preserves their full "richness" – sometimes it's just easier to aggregate the inputs, whether or not it makes for a better shared activity (see Figure 52). This process is termed "media fusion" by some (Kray et al., 2004). Other times, though, the group activity is just better-served by aggregation, as in the MMUI where "private" PDAs are used to allow visitors to "vote" on the content that will be shown on the large "public" shared screen (Kruppa, 2004a). When input events are not aggregated, potential conflicts (e.g., contradictory input events, like "move left" versus "move right") must be ironed out via the activity design.

In some instances designers can leave conflict resolution to the users. For example, in the case of the *Virtual Fishtank*, any user can change the frequency at which virtual bubbles will be emitted from a virtual treasure chest. If a visitor turns the bubble frequency up, another visitor is free to adjust it back down. For this control sharing to work, though, the visitors need to be very aware of the current state of the shared activity – it wouldn't do for a visitor to think that he or she just turned the frequency down, only to see a constant stream of bubble emerge from the treasure chest. *Virtual Fishtank* achieves this transparency by combining two visible indicators: the bubbles themselves, visible on the public display, and a dial on their personal displays whose setting corresponds to the current bubble emission frequency.

The input to MMUIs almost always passes through a private realm (the individual device) before becoming apparent in the public realm (the shared device or devices), a "push" approach to data transfer. Some activities take the opposite "pull" approach, however. In one recent project, a mobile phone was used almost like a stylus to help users select regions on a large shared display (Hardy & Rukzio, 2008). The shared display was equipped with a hidden grid of Near Field Communication (NFC) tags, and the mobile phones (which had NFC readers at one end) were used to trace out regions on a map in order to obtain information on them. In this case, the input act was very public, as the user would literally trace the mobile phone on the shared public display. This input is then translated into informational queries, which would deliver the requested results back to the user's mobile phone display.

3.4.2 Output from Multi-Machine User Interfaces in Museums

MMUIs must distribute output amongst the different participating devices, a process sometimes termed "media fission" (Kray et al., 2004). Once again, designers would probably do well to consider which province (public or private) various outputs belong. The "private" devices are good places to display more sensitive information – for example, work-in-progress that a visitor is not ready to expose to a larger audience, or perhaps user preferences (some people may feel inhibited expressing their true interests if feeling watched). The "public" devices are good places to display information that is of common interest to all members of the group. For example, the PEACH system will display videos of common interest on the shared display screen, but stream videos only of interest to a particular user on his or her individual handheld device (Kruppa, 2004b; Kruppa & Aslan, 2005; Kruppa et al., 2005; Rocchi, Stock, Zancanaro, Kruppa, & Krüger, 2004; Stock et al., 2007). That way the idiosyncratic interests of different users can still be satisfied, while still providing a shared experience of viewing videos of mutual interest on the shared screen (see Figure 50). Although this is system presents traditional A/V content, one cam imagine the principle being generally true for most MMUIs: if the shared "public" display is present to serve as a grounding tool, it wouldn't do to place a lot of extraneous information on it. Only those items of mutual interest or importance should be directed onto the shared display. Of course, once media is "fissioned" into different displays, visitors may have a hard time deciding which display to attend to.

The PEACH system responds to the problem of directing user attention by employing an animated avatar character that "flits" back and forth between a user's handheld and the shared display, acting almost as a "shepherd" of the user's attention. They experimented with other methods (not providing any indication at all, and providing a blinking indicator on the display that the user should be attending to at that moment), looking to see which resulted in better recall of the presented content. They were concerned that simultaneously monitoring multiple displays would induce too much cognitive load in the visitors, and interfere with learning. Unfortunately, none of these three conditions was significantly superior according to objective measures, although visitors reported liking the animated character the best (Kruppa & Aslan, 2005). The problem of directing user attention from one device to another while using MMUIs is still very much an open question.

3.4.3 Access to Companions while Engaged in Multi-Machine User Interfaces in Museums

Section 3.3.2.3 Access to Companions while Engaged in Multi-User Handheld Device Activities described how, when visitors were using multi-user handheld device-based activities, they would sometimes run into problems when trying to have group conversations, owing to a lack of complete access to the distributed outputs of the activity. Without complete access, they sometimes had to go to extra trouble to establish a common ground for discussion. Some systems allowed visitors to "beam" information back and forth between devices, so that all participants would have the same information on their individual devices. The author was not clear on how much of an impediment the beaming may have been, but it is not hard to extrapolate that if an activity requires group members to converse regularly, repeated beaming might become cumbersome. In another study users would try to show each other their handheld displays to achieve grounding, to varying degrees of success. Multi-user handheld-based activities could clearly benefit from a simple and convenient method for providing conversational grounding.

One powerful aspect to MMUIs is that they include that grounding support, typically in the form of large, shared displays. Thus they leverage both the advantages of large, shared displays (conversational grounding, and support for legitimate peripheral participation) and the advantages of handheld device activities (1:1 access to input and output opportunities, and the possibility for individually tailoring interfaces). The form-factors of large, shared displays and handheld devices do not interfere with a visitor's ability to see his or her companions, so most MMUIs do not physically block collaborative conversations (it should be noted that if touch-screen LCDs are used in place of handhelds, their size and placement may affect users' lines-of-sight to each others' faces). As was true of the multi-user handheld device activities of Section 3.3.2.3, the largest impediment to a visitor's access to his or her companions may derive from how the activity is structured.

The MMUI activities reviewed in this section, despite being able to support collaborative learning, didn't seem to be designed to specifically encourage collaborative interactions. None of the activities (the

Virtual Fishtank, PEACH, and the mobile phone map query tool) employed a joint goal strategy, and visitors' choices would only impact their companions' experiences slightly. For example, in the *Virtual Fishtank*, a visitor may design a carnivorous fish which could pursue a smaller fish designed by a companion, but such interaction is very indirect, given that after a fish is released into the aquarium the creator loses all control over it. Likewise, although the PEACH system elicits a vote from each visitor to determine their content interests, any agreements or disagreements in preferences does not change *what* is displayed, just *where* it is displayed (i.e., either on the shared display or a PDA). Regardless of the number of companions present or how they vote, the sequence of video clips viewed by any visitor is identical to what he or she would view if alone. An open question for MMUIs in museums is whether strategies successfully used to encourage collaboration in multi-user device activities (e.g., jigsawing, joint goals) will still work to encourage collaboration with MMUIs.

Another major influence on the success or failure of small-group collaboration is the ability of the group to manage a joint focus of attention – those groups that are able to attain joint attention are also able to solve problems and learn better than those that do not (Barron, 2003). MMUIs, then, with their ability to provide both public and private displays, may pose a problem to the establishment of joint attention, especially if the handheld devices promote a "heads-down" effect similar to what was described in the single-user activities of Section 3.3.1.3. Can the large, shared displays used by MMUIs compete with the attention-drawing power of handheld devices? The coverage in Section 3.4.2 Output from Multi-Machine User Interfaces in Museums discussed the PEACH project, and how researchers were experimenting with different methods for shifting visitor attention back and forth between the private and public displays. The researchers were especially concerned with the cognitive overload involved in simultaneously monitoring different displays. There is not yet enough research on MMUIs in museums to know if such joint attention management issues will impact collaborative learning. It is an issue that bears further attention.

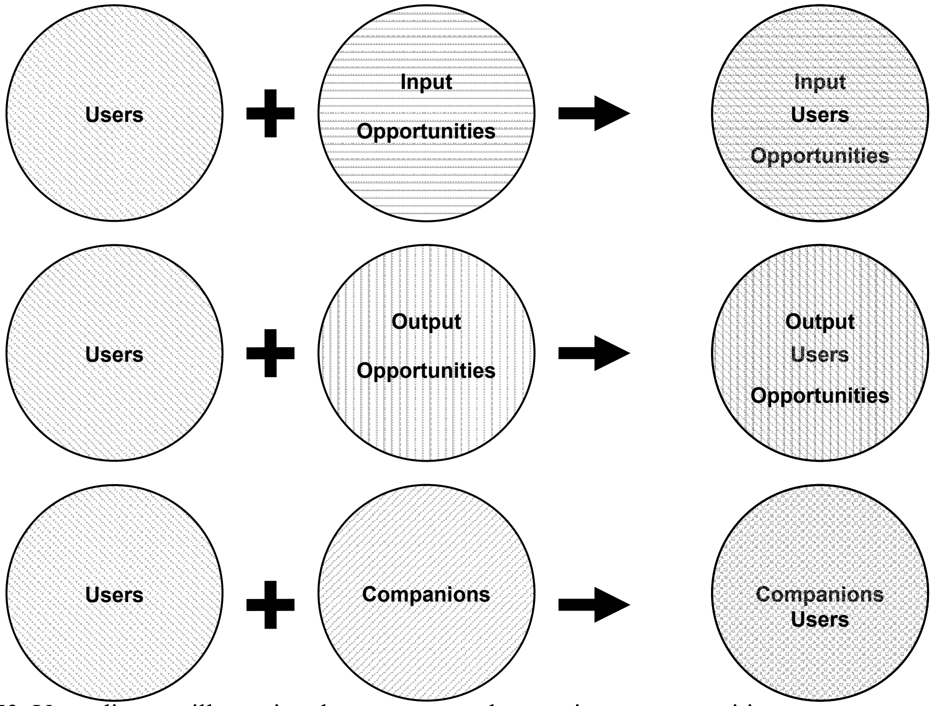


Figure 53. Venn diagram illustrating the access users have to input opportunities, output opportunities, and interaction with companions while using a canonical multi-machine user interface exhibit. Assuming that there is a 1:1 user:device ratio, much like the Multi-User Handheld Device activities described earlier, a user should have complete access to Input Opportunities. The addition of shared display(s) can help provide complete access to Output Opportunities, if systems are designed to display necessary common information on those display(s). A user's interaction with his or her Companions is usually unlimited, assuming there aren't any physical obstructions and the users are not wearing headphones.

3.4.4 Multi-Machine User Interfaces in Museums: a Summary of Group Learning Opportunities

The *potential* access MMUI users have to input, output, and to their companions is superior to all the other form-factors surveyed in this chapter. They have the potential of combining the advantages of multi-user handheld device activities (complete access to input opportunities and companions) with the advantages of large, shared displays (complete access to output, conversational grounding, and legitimate peripheral participation opportunities). If all is designed correctly, there should be no impediments. However, as with multi-user handheld device activity users, the form factor just makes collaboration *possible*, but does not necessarily encourage it.

MMUIs allow inputs and outputs to be divided ("fissioned") and distributed amongst the different interlinked devices, and the review here suggested a few potential guidelines for what to place on private devices, and what to place on public devices, to improve small-group collaboration. There may be instances wherein visitors are inhibited about providing input, as when choosing certain content preferences, or perhaps when engaged in a work-in-progress not yet ready for critical eyes. In these cases, it may be best to *assign sensitive events to "private" devices*. To take best advantage of the shared public display(s) as tools for conversational grounding, designers should try to *reserve public space(s) for shared information*. That way, visitors will be able to easily discern which aspects of the display they should attend to for the benefit of group processes – when presented with too much extraneous information, learners are far too likely to concentrate on the wrong elements (Lowe, 2003). To help with uncluttering the public space(s), designers may wish to *assign frequent, small-granularity input events to "private" devices, for later transmission as condensed input events*. Finally, it may help visitors understand their group's information-processing strategies if designers *consider aligning information flow with "location" of input initiation*. For example, if a system like the NFC-equipped one described in Section 3.4.1 is employed, visitors have the option of selection information directly from the shared display(s) (a "pull" operation). Allowing visitors to initiate the "pull" input operation on the shared display(s) implicitly communicates their area of interest to their companions. Likewise, originating "push" operations on visitors' private devices emphasizes the fact that the information about to be added to the shared display(s) was generated by individual effort.

Section 3.4.3 Access to Companions while Engaged in Multi-Machine User Interfaces in Museums illustrated that, like the multi-user, handheld-device based activities of Section 3.3.2 Multi-user Handheld Device Activities in Museums, if the activity design is not structured properly, collaborative learning will not occur. None of the existing MMUIs covered in the foregoing sections were designed explicitly to support collaboration and so, as far as is discernable from the research, none of them succeeded at promoting collaboration. Although MMUIs are perhaps the most promising form factor for encouraging small-group collaborative learning in museums, they also have the largest number of open questions associated with them. For example – although it is an easy to assume that the addition of public, shared display(s) will help ground the learning conversations of visitors like the large, shared displays of multi-user kiosks do, we do not as yet have any evidence to prove it. MMUIs have the added complication that output is distributed across both public and private displays, which in turn requires visitors to distribute their attention across all of them. We know from the review of handheld-based activities in Section 3.3 that visitor attention is all too easily monopolized by handheld displays, producing the so-called "heads-down" phenomenon. The heads-down phenomenon was observed to mainly occur when the visitors were asked to divide their attention between highly-dynamic handheld displays and relatively static, traditional exhibits or artifacts. Because MMUI users are asked to divide their attention between two dynamic displays, will the heads-down phenomenon still occur? Will MMUI users be able to manage their individual attention so as to partake of both the private and public displays, or will they need some other form of support more effective than the avatars and blinking icons used by the PEACH MMUI? Even if individual users escape the heads-down effect, the group as a whole needs to be able to manage and maintain joint attention. Will MMUI users be able to coordinate their joint attention? Finally, we know that multi-user device activity designers were able to overcome the heads-down effect by simplifying handheld user interfaces and by structuring the activities to encourage (or even enforce) active collaboration. If the form-factor of MMUIs inadvertently promotes the heads-down phenomenon, or fails to support joint attention management, might a careful design eliminate the problem? Perhaps a simplification of the private user interfaces would reduce the attentional pull of the handheld devices. Perhaps a restructuring of the activity would help assuage the problem by motivating the users to distribute their attention in a more productive fashion. All of these questions have yet to be explored.

3.5 Summary

This section summarizes the design strategies for multi-user software-based exhibits that emerged from the above review of existing software-based activities in museums. The prior work was roughly parceled out into classes on the basis of the form-factor and usage paradigm: kiosks versus handhelds versus multi-machine amalgamations, and solo activities versus multi-user activities. In the description of each class of computing technology, the means by which and the extent to which individual visitors are provided with *Access* to *Input*, *Output*, and to their *Companions* has been described and illustrated with a Venn diagram. Each of the presented classes has advantages and disadvantages with respect to one another when contrasted with the ideal HCI scenario for learning[19]: namely, a scenario where each and every visitor has unimpeded access to input, output, and his or her companions (see Figure 54).

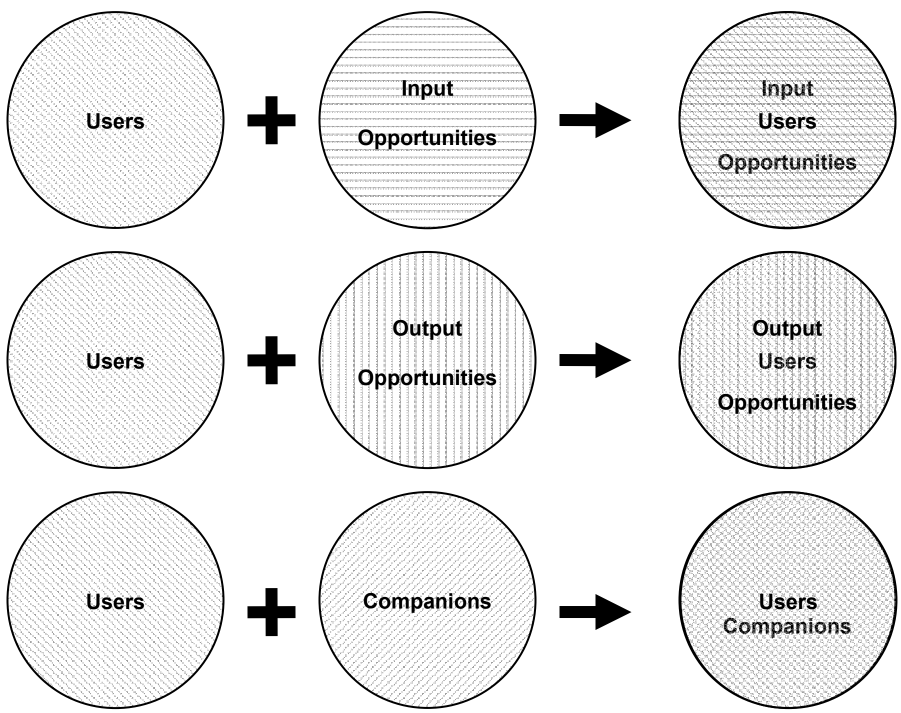


Figure 54. Venn diagram illustrating the ideal access users should have to input opportunities, output opportunities, and interaction with companions for the purpose of social-constructivist learning.

[19] It should be said that an "ideal" scenario for learning can only be defined in reference to the particular flavor of learning a system is being designed to support. In this case, a social constructivist model underlies the definition of learning used here, which is why the unimpeded ability to engage with ones companions (a nod to social learning) and the ability to engage with the exhibit via unimpeded access to I/O opportunities (a nod to constructivism) are deemed ideal.

Looking back on the examples of computing technology in museums, we can see that while single-user kiosks provide an optimal I/O experience for at least one visitor in a small group (the one who is seated/standing directly in front of the kiosk, controlling it) the other visitors in the group are not provided with the ability to engage as deeply with the software's I/O, impeding their ability to engage in constructivist learning. Multi-user kiosks improve on this by giving more equitable access to the software's output, usually via larger displays, but without a similar equitable access to input, some visitors in a group will engage more passively with the exhibit than others, a scenario that usually leads to asymmetric learning gains. Handheld-based activities have the potential to provide equitable access to both input and output, but unless they are explicitly designed to be used by small groups of visitors, they can be very socially isolating experiences. Of all of the form-factor paradigms reviewed, both Multi-User Handheld Device Activities and Multi-Machine User Interfaces have the best potential for encouraging effective small-group learning in museums. The remainder of this chapter will summarize what design strategies could be gleaned from a review of the foregoing technologies, and what open questions still remain for the design of MMUIs, since that is the form-factor adopted by this research.

3.5.1 Design Strategies for Computing Technology in Museums

The design strategies that can be derived from a reading of prior work are, for the most part, suppositions, especially given the relative lack of empirical research conducted on technology in museums. That said, the specific design strategies to emerge are:

- **Kiosk-Derived Design Strategies**
    - *Design for high interactivity*: when single-user kiosks are designed to allow for high levels of interactivity, visitors learn more
    - *Design for symmetric interactivity*: asymmetric activity leads to asymmetric learning; visitors who are afforded more interaction learn more than passive visitors, so allow all to interact
    - *Large shared displays ground conversations*: when all group members share a common point of reference, collaborative conversations are improved
    - *Visitors must be able to attribute inputs*: when visitors don't know whose actions caused changes to the shared context, it impedes group coordination and learning
    - *Large shared displays permit legitimate peripheral participation*: when visitors can see the performances of other visitors, they are encouraged to first observe and then imitate, drawing them into activities
- **Handheld Device Activity-Derived Design Strategies**
    - *Do not use narrative audio*: it interferes with visitor-visitor conversation
    - *Use activity structures to encourage or enforce visitor-visitor interaction*
        - *Joint outcomes*: Design activities so that the outcomes affect all participants
        - *Jigsawing*: gives participants different information or abilities to encourage conversation and participation

    - *Support grounding via information sharing*: because the experience is stretched across different devices, visitors may need help establishing common referents for their discussions
    - *When integrating handheld activities into an object-oriented museum, design to encourage visitor-exhibit interaction*
        - *Localization of input*: Physically tying special information or abilities that can be "collected" to individual exhibits (e.g., via RFID tags) may encourage visitors to attend to them
        - *Enforce "back-and-forthing"*: design handheld activities to require that a visitor read an exhibit's label or study an object to be able to succeed to overcome the heads-down phenomenon
        - *Simplify user interfaces*: it seems to be the case that the more richly interactive a UI is, the more likely visitors are to get engrossed in heads-down interactions
- **Multi-Machine User Interface-Derived Design Strategies**
    - *Careful design attention must be paid to the "location" of input and output events*: computationally-capable input devices allow designers to divvy up events between "private" and "public" devices
        - *Assign sensitive events to "private" devices*: visitors may not want certain preferences, work-in-progress, or information used by the system to be made explicitly public
        - *Reserve public space for important shared information*: shared spaces could quickly become cluttered with idiosyncratic artifacts of different users; try to limit shared space to displaying information that benefits the group (with the caveat that for some activities, displaying work-in-progress could very well benefit the group)
        - *Assign frequent, small-granularity input events to "private" devices, for later transmission as condensed input events*: when visitors need to engage in input processes that require many smaller input events to communicate the visitors' intent, offload these to "private" devices
        - *Consider aligning information flow with "location" of input initiation*: group processes may be aided when "pull" operations originate on public devices, as a visitor's area of interest for the "pull" is made public
    - *Careful design attention must be paid to the interpretation of inputs from multiple users*: unlike single-user applications, MMUIs can receive conflicting inputs from users
        - *Aggregate multi-user inputs into a single "meta-input" only when the activity benefits from it*: actions with large-scale effects may require consent or compromise from all users

        - *When accepting individual inputs, make effects of each input clear*: this will reduce confusion regarding input attribution, and help offload conflict resolution back onto the visitors

3.5.2 Open Questions for MMUIs in Museums

This research makes use of the MMUI form-factor, and so in this section the open questions related to the use of MMUIs for small-group learning in museums will be reviewed. They will later be referenced in Chapter 6, when setting out the research question.

- *Attention management*: MMUIs, which are distributed across multiple devices, demand visitors to distribute their attention across multiple devices. How much of an impediment to collaborative learning is this?
    - *Individual attention and private displays*: The heads-down phenomena is seen to occur when visitors are asked to divide their attention between a highly-dynamic handheld display and a (relatively static) traditional collection of museum objects.
        - Does the heads-down phenomenon still occur when visitors are asked to divide their attention between two dynamic displays (i.e., a MMUI)?
        - Do "simpler" private user interfaces to a MMUI decrease the severity of the heads-down effect?
    - *Group attention and public displays*: Large shared displays are often furnished for software-based exhibits with the intention of supporting multiple simultaneous visitors by grounding their learning conversations.
        - Do groups make use of public, shared displays while using MMUIs to help ground their learning conversations?
        - Would "simpler" private user interfaces increase the amount of attention paid to public displays, and thus increase the quantity and quality of learning conversations (by providing more grounding)?
- *Activity design*: Different activity designs have been employed by handheld-based software in museums, with the intent of encouraging participation of all members of a visiting group.
    - Does the presence of a joint outcome improve collaborative learning with MMUIs?
    - Does the use of unique roles (i.e., jigsawing) improve collaborative learning with MMUIs?
    - Does the use of interdependent roles to enforce collaboration improve collaborative learning with MMUIs?

CHAPTER 4

Related Work: Supporting Collaborative Activities with Computer Technology

The Computer-Supported Collaborative Learning (CSCL) literature is primarily grounded in formal learning environments (e.g., classrooms), and the Computer-Supported Collaborative Work (CSCW) literature is largely centered around workplaces, so it is important to stress that one should not expect the designs developed for these types of environments to automatically map to museums. CSCL and CSCW work is useful to review, though, to suss out theoretical underpinnings that can be used to describe and frame the design issues for collaborative technologies in museums. By combining a review of these areas with the more anecdotal information available on the educational use of computers in museums in Chapter 3, the reader should be able to better understand the problem space this research inhabits. The first section of this chapter, 4.1, will address the theoretical organizing principles used in CSCW and CSCL to help categorize the factors that define collaborative software systems and in turn systematically analyze how these factors impact collaboration. The subsequent two sections, 4.2 and 4.3, will address prior work in the CSCW and CSCL literature that relate to this research, identifying open questions of interest to the communities. The final section of this chapter, 4.4, will relate the open questions concerning the design of MMUIs for museums to the themes and open questions uncovered in the review of the CSCW and CSCL literatures.

4.1 Organizing Principles

Recall from the introduction to Chapter 3, Computing Technology in Museums, that the organizational theme was centered around different forms of a user's access: (1) access to input opportunities, (2) access to output opportunities, and (3) access to his or her companions. In the next two subsections, the organizational principles already in use in the CSCW and CSCL literature to categorize and structure collaborative software will be reviewed. First, Section 4.1.1 will provide a definition of "collaboration." Section 4.1.2 will address those principles that apply to both CSCW and CSCL work, whereas Section 4.1.3 will cover the underlying principles that differ between CSCW and CSCL research. The purpose in reviewing these ideas is to set the stage for later discussions concerning the framing of the research pursued in this dissertation, and in turn for situating the contributions this research makes within the bodies of established CSCW and CSCL literature.

4.1.1 Definition of Collaboration

Collaboration is a notoriously over-used term, which can signify many different things to many different people. Other terms are often used as well (for example, coordination and cooperation), and to make matters even more confusing, these terms often seem to be used interchangeably, a practice that seems to be *de rigueur* throughout much of the literature (Rummel & Spada, 2005). Although CSCL and CSCW researchers may not necessarily have different views of collaboration, they certainly seem to place different levels of importance on defining the term. In an oft-cited piece, Pierre Dillenbourg summarized many of the different perspectives on collaborative learning in the introduction to a book that emerged from a series of workshops organized around the Learning in Humans and Machines (LHM) research program (Dillenbourg, 1999). They agreed that the broadest (and thus, least useful) definition would likely be “a situation in which two or more people learn or attempt to learn something together.” As with many expansively-defined concepts, more precise definitions can be obtained by getting down to brass tacks: the measures used to try to quantify the concept. In CSCL, at least, there seem to be two broad camps: one which takes an individual perspective on measuring collaboration, and one that takes a social perspective. Figure 55 provides an illustration of the different theoretical and methodological influences on individual and social theories of learning. It is important to stress that individual and social theories of learning *are not necessarily incompatible with one another*. They each emphasize different aspects of a learning scenario.

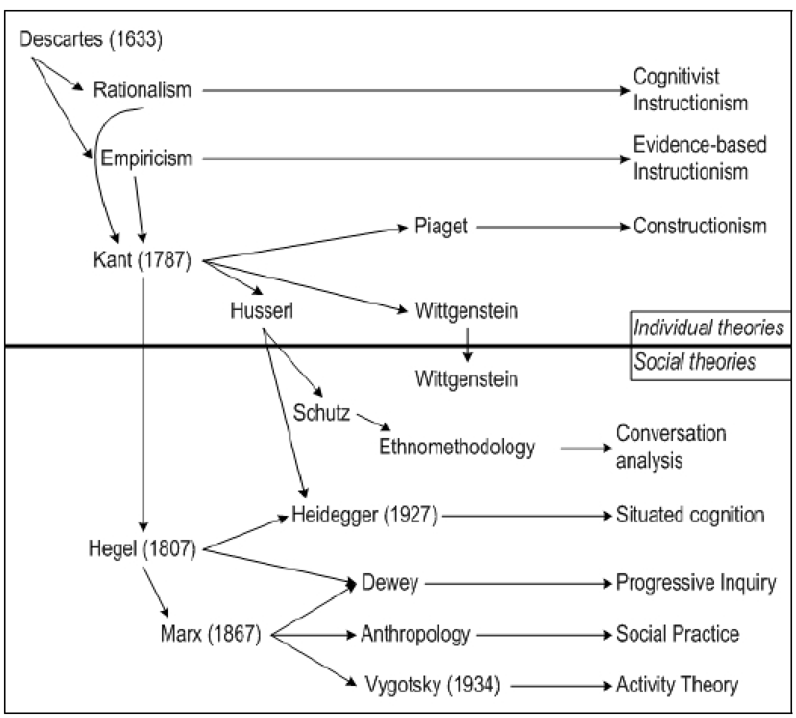


Figure 55. Diagram of the evolution of influences on individual and social theories of learning, taken from (Stahl, 2006b).

One cognitively-oriented method tries to elicit learners' mental models before and after a joint learning experience, and looks for "convergent conceptual change" (Roschelle, 1992) – in other words, did the learners come away with mental models that were more similar to each other than when they began? Thus, joint activities that cause mental models to converge would be one definition of collaborative learning. Of course, establishing a learner's mental model of a topic area is a challenge in and of itself – written long answers, concept maps, and probing interviews have all been used to this end.

A more situative perspective would detect collaborative learning not by inward mental structures, but more by the outwardly-observable relationships that form among learners. This perspective is heavily influenced by the views of Vygotsky, who subscribed to the notion that people come to understand things by first articulating them in a social context (Vygotsky, 1978). Those concerned more with a situative perspective tend to take entire activity systems as units of study (in lieu of individual minds), and often engage in more ethnographically-influenced analyses, like interaction and conversation analyses (Sawyer, 2006), that are more concerned with processes than results. Thus, through this lens collaboration occurs if people engage in behaviors like establishing and maintaining common points of reference and building on one another's remarks, which often requires that researchers engage in detailed conversational analyses.

Researchers investigating collaborative *work* don't seem to be quite as occupied with trying to define collaboration – perhaps because the work tasks they set out to support are by-and-large already pre-existing, with specific demands on the joint behaviors of participants. Also, collaborative work has the advantage of possessing its own built-in metric for the success or failure of collaboration: the quality of the resulting work product. These reasons are probably why CSCW seems to be more concerned with describing or defining the situational characteristics that shape collaboration that collaboration itself, as we will see in the next section.

4.1.2 Shared Organizing Principles

The CSCW and CSCL fields tend to share many of the same theoretical framings, probably because many CSCL researchers published in CSCW venues until they established a separate body of literature in the early 1990s (Stahl, 2006a). For example, CSCW and CSCL systems are often framed in terms of the space and time dimensions of their context of use. The choice of space-time as a framing is likely an artifact of how the CSCW field coalesced.

4.1.2.1 Space-Time Categorization: Location and Synchronicity

In the early 1980s CSCW emerged as researchers began to recognize that computers, and especially networked computers, could actively help users work with other users. At the beginning of CSCW research, the field was largely centered around what was then called "groupware," software that would allow people working at different physical locations, and often working during different periods of time, coordinate their efforts on shared tasks (Grudin, 1994). For that reason, CSCW software began to be classified along two binary parameters: (1) *location*, and (2) *synchronicity* (see Table 1). Location can assume two values: *co-located* when the workers are occupying the same physical location, and *remote* when they are not. Synchronicity can also assume one of two values, *synchronous* when the users are working at the same time, and *asynchronous* when they are working at non-overlapping intervals. This framing still applies to many CSCL applications as well, especially when HCI issues are the main research question, even though the activity in question is learning, not working.

Table 1. Adaptation of matrix used to conceptualize CSCW research originally defined by (Johansen, 1988). The cells are populated with exemplars; this is not an exhaustive listing. The dimensions not considered by this review are grayed-out.

		Synchronicity	
		Synchronous (same time)	**Asynchronous** (different time)
Location	**Co-located** (same place)	Single Display Groupware Interactive tables	Persistent displays for project management Groupware for shift work
	Remote (different place)	Chat programs / Instant Messaging Videoconferencing	Bulletin board systems (BBS) Email

It is worth noting that the majority of CSCL and CSCW research work has studied collaborative learning in *remote* or *distance* activities (as opposed to *co-located*, or face-to-face, activities), and consequently both have historically focused more on *asynchronous* activities (like message boards) than *synchronous* activities (wherein participants are working together simultaneously). To best serve the social and learning needs of small groups visiting museums, however, a computer-based exhibit should support *co-located*, *synchronous* interactions. MMUIs fall squarely into this quadrant.

4.1.2.2 Coupling

Once a collaborative system has been determined to be *synchronous* in its usage, one can then split hairs about how that synchronicity is implemented. Must all input actions from all users be initiated simultaneously, a sort of "Wonder Twins" model where no action can be taken unless all group members click mouse buttons at the same time ("Form a spreadsheet!")? Or is there a looser, more interleaved way for each user to contribute input? This method of further categorizing the synchronicity of a CSCW (or CSCL) system became known as the degree of *coupling* a collaborative system possessed (Dewan & Choudhard, 1991). *Coupling* can be thought of as how closely two users are working together on a shared artifact (Nova, Wehrle, Goslin, Bourquin, & Dillenbourg, 2007). A system with "tight" coupling requires a high degree of simultaneous focus and action, and a system with "loose" coupling allows users to have more distinct foci and for them to take individual actions. For this reason, a system's degree of *coupling* is thought to also determine the degree of collaboration present during the joint activity. See Table 2 for an overview of how different degrees of coupling manifest in different distributions of access to input and output.

Table 2. Matrix used to illustrate the degree of coupling possible with a *synchronous* multi-user, multi-device shared GUI workspace. The degree of coupling refers to how closely users are working together on a shared artifact.

Coupling Degree	I/O Access Distribution		Collaboration Continuum
	Input	Output	
Tight	Joint Execution (e.g., simultaneous button press)	Strict WYSIWIS (all views are identical)	Enforce
↕	Token or Turn-based input	Relaxed WYSIWIS (Individual portals on shared view)	Encourage
	Concurrent input		
Loose	"Commit" style input	Custom views	Enable

In operational terms, coupling is a measure of how large of a role other users have in affecting the inputs produced by and outputs viewed by a given user. A strict definition of tight input coupling means that a user must have direct coordination with users to provide input to the shared workspace, as with *joint execution* systems wherein users must enact input events simultaneously, like when two users need to press a keyboard keys at the same time (Light, Foot, & Colboum, 1997). Moving down along the coupling spectrum, *token or turn-based input* requires only that a user have permission from his or her companions when attempting to provide input (i.e., only one user will be providing input at a time). This can be enforced physically, as when users share a single means of input like a mouse, or virtually, through some sort of semaphore or locking mechanism (Inkpen, Booth, Gribble, & Klawe, 1995; Olson, Olson, Mack, & Wellner, 1990).

Table 3. Matrix used to illustrate the different levels of involvement a partner must have with an input-providing user in collaborative computing systems with different degrees of input coupling.

Coupling Degree	Input Coupling Mechanism	Involvement Of Partner
Tight	Joint Execution (e.g., simultaneous button press)	Action
	Token or Turn-based input	
	Concurrent input	Attention
Loose	“Commit” style input	

As the input coupling becomes even more relaxed, we move from a realm where some sort of *action* is required on the part of a user’s companions for the user to provide input, to a realm where only the *attention* of a user’s companions comes to bear on his or her input attempts (see Table 3). For example, *concurrent* input allows users to provide inputs to the system at the same time[20], although they are usually able to see the impacts of their partners’ input actions as they occur. The loosest input coupling interpretation possible for *synchronous* systems is “*commit*”-style input, where users are free to make a larger series of inputs (perhaps in a private workspace) before submitting those changes, all at once, to the shared system. So in a very loosely-coupled input system, a user’s companions may not even be privy to his or her “in-progress” actions, but only to the eventual results once the commit takes place.

Output can also be coupled to various degrees. In some of the earliest *co-located*, *synchronous* applications, a WYSIWIS (“What You See Is What I See”) approach was experimented with (Stefik, Bobrow, Foster, Lanning, & Tatar, 1987; Stefik, Foster et al., 1987). In a strict WYSIWIS system, all users see exactly the same representation as all other users – so, for example, if one user scrolls a text document, *all* users will see the “scrolled-to” portion of that document[21]. A more relaxed WYSIWIS system still provides a common visual work space, but each user is in control of his or her own “viewport” onto that workspace. Put in MVC (Model-View-Controller) terms (Krasner & Popo, 1988), each user has his or her own view of a shared model. So, returning to the text document example, each user may scroll (Control) to a different region (View) of the document (Model). The assumption, though, is that the users’ viewports all are similar in terms of how they visualize the model’s data – the users are just free to “look” in different places in the shared model. In the loosest of output couplings, each user can have his or her own custom

[20] To be technical, the inputs from different users are typically queued in a FIFO (First-In, First-Out) manner, although this ordering is seldom noticed by users of the system unless they provide contradictory inputs.
[21] Note that this assumes parity in terms of display devices.

view of the shared model data – the manner in which this data is visualized may or may not bear any relationship to how another user's view visualizes the same data.

The degree of input and output coupling obviously has a direct impact on the collaboration that will occur amongst users of the shared system. (Benford et al., 2000) conceptualized collaborative software as existing along a continuum, from systems that merely *enable* collaboration to systems that *enforce* collaboration, with systems that merely *encourage* collaboration lying somewhere in-between. A system that is tightly-coupled is one that, perforce, *enforces* collaborative behaviors, since no user can act without the participation (and awareness) of his or her co-users. Systems that merely *enable* collaboration, however, have no structural components that require users to behave in manners that promote effective collaboration – the onus of collaboration rests on users making good use of the system's capabilities. Loosely-coupled systems, then, only *enable* collaboration, since users are free to see and do what they wish, independently of other users' perspectives and actions. It is also worth mentioning that the input and output of a collaborative system need not be coupled to the same degree, although they usually are.

4.1.2.3 Application to Technology in Museums

Returning to the research area at hand, we can see that the educational software systems designed for museum use described in Chapter 3 were nearly all *synchronous* and *co-located*. They differed quite a bit, however, on the degree of *coupling* their designs embodied. The notion of *coupling*, as described in CSCW and CSCL work, is typically only applied to use contexts wherein each would-be collaborator has a computational device of his or her own. Several of the categories of software covered in Chapter 3 (like the single-user kiosks described in Section 3.2.1 and to a lesser extent the multi-user kiosks of Section 3.2.2) do not provide users individual access to devices, let alone input and output opportunities, and so viewing such systems through the coupling lens is not very productive. When considering some of the other categories (especially the multi-user handheld device activities of Section 3.3.2 and the multi-machine user interfaces of Section 3.4), however, coupling degree becomes a very useful categorization tool.

For example, the hidden-picture quiz games of (Yatani et al., 2004) represent a system that is tightly coupled with respect to output, as each user has a WYSIWIS display of the same hidden picture. The input was also tightly coupled, requiring that both users correctly answer the question for each grid for it to be revealed. By way of contrast, the treasure-hunt style game of (E. Klopfer, Perry et al., 2005b) was very loosely-coupled in both output and input. Users had role-specific displays, "custom views" as Table 2 would describe them. Input took place within the confines of these private custom views, and a user's input would only affect his or her partners when they collectively chose (by aligning the IR ports of their devices) to share information. Looking at these two systems, it is clear how the tightly-coupled hidden picture game enforces collaboration, whereas the loosely-coupled treasure hunt merely enables it, leaving the details of collaboration to the users. Which approach works better for encouraging collaborative learning? It is hard to say by comparing these two examples, because they were structured for different types of museums (science versus art) and for wholly different cultures (Japanese versus American). It is interesting to note, however, that the users of the hidden picture system that enforced collaboration were just as likely to go off individually as they were to tour the museum in pairs, whereas the users of treasure-hunt system that merely enabled collaboration tended to travel in groups, even though their individual roles allowed them greater freedom. This issue of the potential impact of *coupling* on collaborative behaviors will be returned to in the discussion in Chapter 8.

4.1.3 Disjoint Organizing Principles

The prior section covered those organizational schemas that are used in both CSCW and CSCL to categorize collaborative software, but not all of the theories underlying CSCW and CSCL are in harmony. To an extent, CSCW can be thought of as being dominated by User-Centered Design (UCD), which stresses (among other things) the need to construct highly-efficient user interfaces to help users complete their tasks quickly and accurately (Norman & Draper, 1986). To an extent, CSCL can be thought of as being dominated by Learner-Centered Design (LCD), which stresses (among other things) the need to occasionally create an inefficient interface in order to spur learning (Stahl, 2006a).

4.1.3.1 UCD and LCD

When an interface slows learners down and makes them take more time to complete a task than would be absolutely necessary, it can encourage learners to reflect on what they are doing in a way that leads to deeper learning. When an interface makes it easy for a learner to make mistakes unless he or she is absolutely certain about the material, it can help users learn by forcing them to confront their (sometimes incorrect) mental models. So, an exaggerated view of UCD would have interface designers attempting to reduce tasks to a minimal number of steps, to shorten the overall task completion time as much as possible, and to reduce the possibility that a user would make a mistake. An exaggerated view of LCD would have interface designers inserting extra steps in a task to give users more exposure to the ideas to be learned, building in time for reflection at junctures when they deem it necessary for users to think a bit more, and constructing “tricky” input opportunities to try to elicit evidence of incorrect mental models. In comparing these two extremes it is obvious that the two different design philosophies would quite naturally lead to very differently-structured interfaces.

4.1.3.1 Application to Technology in Museums

The HCI style required for museums, in many ways, straddles the UCD and LCD design philosophies. The primary aim of museum software is to promote learning, for which LCD would seem to be well-suited. At the same time, though, a user interface in a museum must allow users to very quickly figure out what needs to be done, how to do it, and allow them to do it. We know from literally decades of visitor studies research that visitors typically spend only two or so minutes at an individual exhibit (Borun et al., 1996; Diamond, 1986; McLean, 1999; Sandifer, 2003), and that visitors will not learn from interactive exhibits, and may even walk away, if they cannot figure out how to manipulate them in short order (Allen, 2002; Diamond, 1986; Fehrer, 1990). This sort of high-stakes usability challenge is one for which a UCD perspective would be helpful. It behooves those interested in devising successful user interfaces for museum exhibits, then, not to follow either an LCD or an UCD design philosophy, but rather, to find design principles that maximize both opportunities for learning as well as the usability of the software. Now we will move on to a brief overview of related work in CSCW.

4.2 Computer-Supported Collaborative Work (CSCW)

The field of Computer-Supported Collaborative Work (CSCW) emerged in the mid-1980s as researchers began to recognize that computers, and especially networked computers, could actively help users work with other users. Using the matrix in Table 1, we can see where many popular CSCW applications fall: email is *asynchronous* and *remote*, whereas videoconferencing is *synchronous* and *remote*. The *remote* collaboration tools are the ones that seem to have most infiltrated the public consciousness, shaping people's ideas about what forms collaborative software may take. This is probably because, in terms of computing equipment, *remote* collaboration tools tend to require only standard desktop computers (as opposed to more bespoke hardware). *Co-located* collaborative software, on the other hand, tends to require specialized hardware, like the *synchronous*, *co-located* touch-sensitive interactive tables, and the *asynchronously* used, *co-located* large displays used for project management. For the purposes of this research, the *synchronous*, *co-located* quadrant is most relevant. Owing to the wide variety of different form factors and proofs-of-concept developed in this quadrant (and a concomitant lack of consistent research goals and methods), the following review will focus the three form factors most relevant to the design of multi-device museum exhibits: single-display groupware, interactive tabletops, and multi-machine user interfaces. The purpose is to introduce the reader to the common themes underlying the research questions often asked by the CSCW researchers who have studied these form factors.

4.2.1 Single-Display Groupware (SDG)

Shared visualizations are a powerful means to ground collaborative conversations (Clark & Wilkes-Gibbs, 1986), and have been shown to improve the speed and accuracy with which partners complete joint tasks (Fussell, Kraut, & Siegel, 2000; Kraut, Fussell, & Siegel, 2003). Thus, a large and ever-growing body of research into co-located, synchronous systems involves the use of a shared display. The means for providing output is often dictated by what is technologically feasible at the time – and in general, the preference seems to be for the largest, highest-resolution display that can be obtained by the research group. Many of the research questions posed, then, concern the means of providing input.

4.2.2.1 Input to Single-Display Groupware

Groups that share a single input channel, for example, a single mouse, often experience conflict as members compete for access (Inkpen et al., 1995; Stewart et al., 1999; Stewart, Raybourn, Bederson, & Druin, 1998). The most direct approach to providing input access to all users is to "simply" augment existing displays with extra mice (Bier & Freeman, 1991; Birnholtz et al., 2007; Inkpen et al., 1995; Pawar, Pal, Gupta, & Toyama, 2007), although the effect of this approach is that the conflict just migrates from the physical realm to the virtual. Depending on the task structure, there are many different ways of resolving these conflicts.

Input devices can be "registered" to particular users, so the software can keep tie user input to modal data structures (Bier & Freeman, 1991), although such software-based conflict resolution is likely to be extremely task-dependent. One extreme example is a graphical layout program that allows a user to delete only those items that he or she has created, not those created by other users (Morris, Huang, Paepcke, & Winograd, 2006). In this example, impact on the shared space is limited by "ownership," which turns the SDG into a space that supports parallel work, but not necessarily collaborative work. This is perhaps the loosest coupling of input available for the joint use of shared displays.

Designers can structure software so that it allows all users to have equivalent power over all aspects of the shared space, but in these scenarios designers must decide at which level of input "granularity" visitors can interfere with one another's actions. Oftentimes, performing a task requires a series of atomic-level actions before anything meaningful can be made out of a user's input. To prevent users from inadvertently (or intentionally) interfering with such tasks, some systems block inputs from other users during certain types of serial inputs (Tse, Shen, Greenberg, & Forlines, 2006). This type of structuring locates the conflict at a level of granularity below the all-encompassing "mode," but above the level of atomic actions – perhaps the best way to characterize this is to describe it as providing "task" level ownership. Another approach that situates the conflict in the nebulous region between action and mode is to create on-screen "tools" that users can acquire to obtain unfettered access to specific roles within the shared application. For example, a shared drawing program uses a "local tools" motif that allows users to assume roles (pen, eraser, etc.) that are mutually excusive (Bederson et al., 1996; Hourcade, Bederson, Druin, & Taxén, 2002). The advantage of using on-screen "tools" is that they reify and make visible to other users exactly which roles different users have assumed.

All of the preceding schemas for providing input have taken a loosely-coupled approach, but some researchers have looked into tightly-coupled input mechanisms. For some types of actions, especially those that have global effects, enforcing a tight coupling may be the best approach. For example, in a joint task that involves laying out graphical elements, there are times when the group may wish to clear the screen – but because this action impacts all users, the input schema is designed to require all users to simultaneously provide the "clear" command (Morris et al., 2006).

4.2.2.2 Output from Single-Display Groupware

The tendency to use large displays for SDG isn't just gadget lust - larger displays have been shown to facilitate improved communication between group members (K. O'Hara et al., 2003). Many SDG displays are just large monitors or projected images, but for displaying information with a third visual dimension, immersive VR caves can be used (Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992). Other research has employed tabletop displays (Dietz & Leigh, 2001), which works best for two-dimensional data that does not need to be oriented in any particular direction to be understood (i.e., text is problematic, although imagery can still be used in most cases). The exact nature of the display device used seems to be dependent more on the type of data to be displayed than the collaborative task itself.

The very nature of SDGs ensures that users will have tightly-coupled output – viewing the exact same public display – but one theme that does come up is the extent of on-screen privacy available to users. There is an entire spectrum, where individual manipulations (and their effects) can either be hidden or made plainly obvious on a shared display (see Figure 56). How individual users (and/or their actions) will be represented on that shared space, and to what extent this individual information will be kept private, can have a big impact on the nature of collaborative work. Many tasks, especially those associated with creating or editing, benefit from having on-screen representations of the users, like color-coded cursors (Bier & Freeman, 1991) . For some tasks, especially those with global effects (like navigation) it can be very useful to make the actions – and their effects – very apparent on the shared display (Stanton, Neale, & Bayon, 2002). For other tasks, having a private work area is important (Stacey D. Scott, Grant, & Mandryk, 2003). This has led some researchers to experiment with providing alternate, private channels, via methods like shuttered glasses that allow users to view a private region in conjunction with the shared space (Shoemaker & Inkpen, 2001), or via individual audio (Morris et al., 2004).

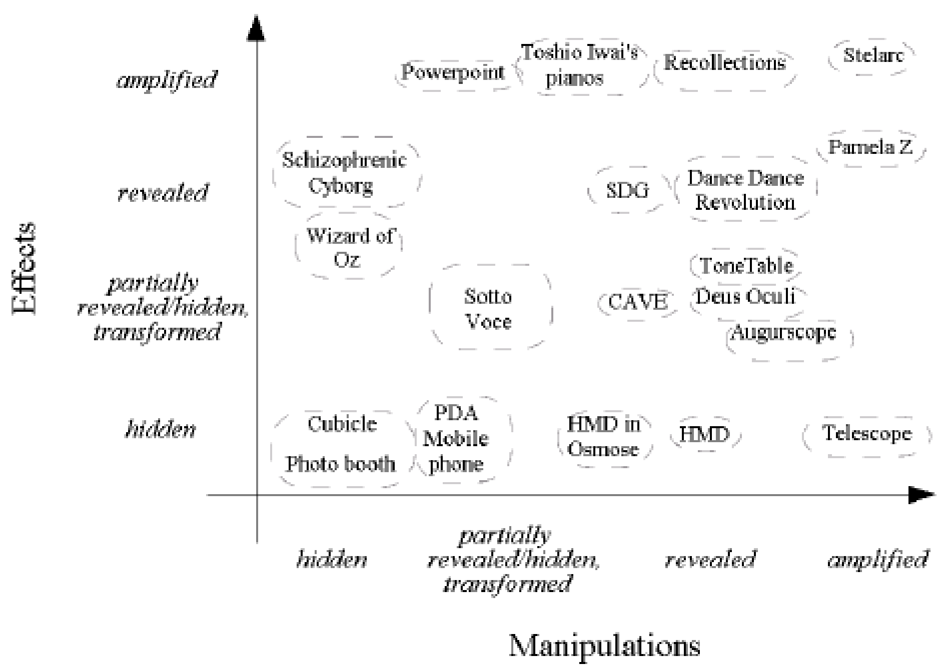


Figure 56. A classification of collaborative user interfaces by the awareness spectators have of both the "manipulations" (inputs) provided by users and the "effects" (outputs) of those manipulations. Notice that SDG systems are classified as "revealed" for both inputs and outputs. Taken from (Reeves et al., 2005)

4.2.2.3 Access to Companions while Engaged with Single-Display Groupware

The purpose of SDG was to provide groups of users with unfettered access to both the display and each other. The presumption was that input conflicts of the sort discussed in Section 3.2.2.1 Input to Multi-User Kiosks would be resolved via verbal communication. There is some evidence that the way in which input from multiple users is coupled may have an impact on user-user interactions. (Birnholtz et al., 2007), for example, found that loosely-coupled inputs led to more selfish behavior in a joint task. The task in question also offered individual-level rewards, though, so most likely there is an interaction between the reward structure of a task and the input coupling. Not nearly enough research has been done on the impact of input coupling and task structure on collaboration to make any specific predictions.

4.2.2.4 Single-Display Groupware: a Summary of Open Questions and Themes

The types of work activities SDG systems have been proposed to support (collaborative editing, drawing, layout, design reviewing, information exploration, etc.) all have very different activity structures, so there is also a fair amount of variety in what questions CSCW researchers pursue. A review shows that while SDG implementations by definition have tightly-coupled output, they vary in the degree of input coupling. Loose input coupling permits simultaneous work, but there are hints that when it is combined with an activity structure that contains individual-level rewards, the quality of the collaborative effort decreases. It is an open question if poorer collaborative performance is an unavoidable consequence of loose input coupling, or if it can be avoided by making use of an activity structure with group-level rewards, but no individual rewards.

Users of SDG systems are, by definition, exposed to the same display. Actions in this space are more public than actions on a single-user desktop would be, adding a performative nature to input actions that user interface developers do not usually have to grapple with. To this end, how users and their actions are represented on-screen has been flagged as an important issue, but not systematically explored. The general consensus seems to be that being able to tell the on-screen representations of users apart is a good thing, and there is a suggestion that on-screen representations of actions should be emphasized proportional to their impact on the shared context. Acknowledging that not all actions should necessarily be made public, some SDG implementations are concerned with providing some sort of private workspace to users. This notion of public and private displays, and public and private work will be addressed further in the next section. Multi-machine user interfaces are often structured to provide just this sort of public/private divide.

4.2.2 Multi-Machine User Interfaces (MMUIs)

The term “multi-machine user interface” (MMUI) was coined to describe SDG systems where input is provided via handheld computers (Brad A. Myers, 2001), but the concept was explored many years earlier with desktop computers serving as the individual interfaces(Stefik, Bobrow et al., 1987; Stefik, Foster et al., 1987; Tani, Horita, Yamaashi, Tanikoshi, & Futakawa, 1994). Xerox PARC’s Colab is perhaps the earliest example of such a system, where individual terminals were provided to give user a relaxed WYSIWIS (“What You See Is What I See”) perspective on the SDG content (see Figure 57). The term “multi-computer user interface” has also been used(Rekimoto, 1998), but MMUI seems more broadly applicable in this era of ubiquitous computing.

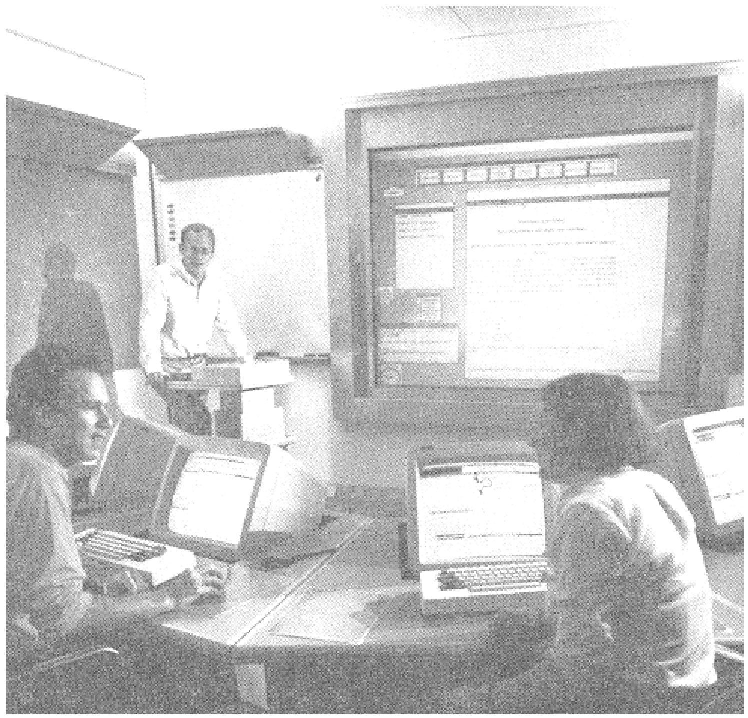


Figure 57. View of Colab, Xerox PARC’s early experiment with creating a synchronous, co-located collaborative meeting space. Each of the computer terminals was networked to the others (and the large touch-screen display) so that they could share views in a “relaxed” WYSIWIS manner (Stefik, Bobrow et al., 1987).

The motivation behind the creation of most of the MMUI systems is to reduce the *physical asymmetry* of the human-computer interaction experience, a term taken from (Rodden et al., 2003). This is the same as the concept of equal access to input and output possibilities that was flogged throughout Chapter 3. Many of the same themes present in SDG research (coupling, public/private actions) appear in MMUI research as well, but the presence of extra displays allows for a greater variety in the ways in which these themes have been explored. Additionally, because multiple displays are available, decisions must be made concerning not just what user will be shown, but where that content will be shown.

4.2.2.1 Input to Multi-Machine User Interfaces

A good proportion of MMUIs make use of handheld devices (either PDAs or cellular phones) as the means by which user provide input to the system (Ganoe, 2002; Greenberg, Boyle, & Laberge, 1999; Brad A. Myers, 2000; Brad A. Myers, 2001; Brad A. Myers, Stiel, & Gargiulo, 1998; Paek et al., 2004; Rekimoto, 1998; Sugimoto, Hosoi, & Hashizume, 2004). Most of these systems serve as proofs-of-concept, and many were created with the purpose of supporting design or planning meetings. What is interesting about these meeting-oriented systems is that all of them locate detailed input on the private devices. This is not an examined concept – it was not tested against other methods – but the decision parallels the design recommendation that emerged from the analysis of museum-based MMUIs: *assign frequent, small-granularity input events to "private" devices, for later transmission as condensed input events*. The purpose here seems to be to keep the shared display uncluttered, and to take advantage of the perceptual immediacy of a handheld device to help users focus on details. Another system, which uses tablet PCs as private devices, located detailed annotation work on the private devices not to keep the public space uncluttered, but rather to prevent users from being inhibited about engaging in annotation (Forlines, Esenther, Shen, Wigdor, & Ryall, 2006). None of the systems reviewed had tight coupling of input – most was loosely-coupled synchronous input, with a few systems even allowing for loosely-coupled *a*synchronous input, as when users would "snarf" content onto their private devices for manipulation and later return it to the shared display (Brad A. Myers, 2001).

4.2.2.2 Output from Multi-Machine User Interfaces

The presence of multiple displays provides designers with a variety of schemas for dividing ("fissioning") output (Kray et al., 2004). Sometimes the output is "*stretched*" across multiple devices, as in this scenario where Google Earth output is distributed across multiple machines (Forlines et al., 2006). In persistent, room-wide MMUIs like "Roomware," it may be the case that certain displays are reserved for displaying certain types of information in the service of certain types of work, a *task-based partitioning* (N. Streitz, Prante, Müller-Tomfelde, Tandler, & Magerkurth, 2002; N. A. Streitz et al., 1999). With many MMUIs, though, the private devices are used to provide customized "*portholes*" onto a shared context (R. B. Smith, Hixon, & Horan, 1998; Stefik, Bobrow et al., 1987). A fourth style of "fissioning" is to deliver *complementary* but distinct information to the private devices, providing specialized views with different details than what the shared display provides (Sugimoto et al., 2004). While many systems ensure that despite being "fissioned," the views across devices are synchronized; some allow for asynchronous work to be done on the individual displays (Greenberg et al., 1999; Brad A. Myers, 2001).

Many of these systems were proofs of concept, and so by and large they did not tend to examine attention division behaviors. One study revealed that although the users reported enjoying being able to shift their gaze between the SDG and the private handheld display, they didn't often do so. The researchers noticed that the users tended to get lost in the private view, ignoring the shared display for long periods of time, even to the detriment of the joint task execution (Sugimoto et al., 2004). Although not labeled as such, this is a clear example of the heads-down phenomenon noted by museum employees (see Section

3.3.1.3 Access to Companions while engaged in Single-User Handheld Device Activities). The researchers attempted to draw user attention back to the shared display by displaying highlighted indicators on the private display when changes were made to the shared display, but users either failed to understand the indicators or found them to be annoying.

4.2.2.3 Access to Companions while Engaged in Multi-Machine User Interfaces

Many of the MMUI systems, especially those devised for group meeting scenarios, were designed to allow multiple users to insert new content into the shared display without unduly affecting the actions of other users. The usage tended towards parallel execution, rather than true collaboration, perhaps because the activities supported by most systems are rather open-ended – nothing about the structure of most of the MMUIs presented here enforced collaboration. Very few of the studies examined whether or not visitors were interacting, or how the systems affected interactions. In one study, smaller sub-groups of users were observed to utilize the private displays to support side discussions (Sugimoto et al., 2004), but this was the same study that noted that it was common for individual users to become lost in the heads-down effect and miss out on whole-group discussions. The extent to which users engaged in one or the other behaviors is unclear. The designers added a tightly-coupled input element (an “agree/disagree” button) on the shared screen to try to overcome the heads-down effect, but the results were mixed – some groups found it helpful, and some found it to be unnecessary.

4.2.2.4 Multi-Machine User Interfaces: a Summary of Open Questions and Themes

The systems reviewed in Section 4.2.1 Single-Display Groupware (SDG), although also concerned with public/private divisions of work, did not have to contend with public/private divisions of visual attention. The research on SDG systems never hinted that users might have difficulty attending to the shared display. The MMUI systems here demonstrate a wide variety of methods for “fissioning” the display output onto the different public and private displays, but there is a suggestion that when the private interfaces display detailed visual information, the users get lost in the heads-down phenomenon just as museum-going handheld users did in Section 3.3. The two design strategies experimented with to alleviate the phenomenon, (1) forcing attention to the shared display via tightly-coupled inputs on the shared display, and (2) reminding users to make use of the shared display via passive indicators on the private display, both received mixed results. Although MMUIs allow users to shift between private and public work, some users need guidance in doing so, and just how to do so effectively is very much an open question.

4.3 Computer-Supported Collaborative Learning (CSCL)

Computer-Supported Collaborative Learning grew out of the CSCW field as researchers who studied collaborative learning contexts (especially contexts with children) realized that their interests, while similar, often had different emphases. Akin to how the difference between UCD and LCD came about, researchers working in learning contexts realized that while the technology form factors might be similar, the processes that the technology supported were fairly different. Oftentimes, CSCW research looks to structure new technology in the service of processes already taking place in a work context, whereas CSCL research looks to shape learning processes via the structure of the technology. This section, then, will part from the usual Input/Output/Access to Companions structure used thus far, and instead forefront the learning activities and processes that CSCL researchers are attempting to support with different form factors. To understand software designed from a Learner-Centered Design perspective, one must understand what types of learning activities were being supported, and how. These will be pulled out as themes.

4.3.1 SDG in CSCL

4.3.1.1 Classrooms using SDG

The primary purpose of Single-Display Groupware in classrooms seems to be to provide a shared visual focus around which learning conversations can be grounded (Clark & Wilkes-Gibbs, 1986). Many modern classrooms make use of projectors to provide a large, shared display for lecture purposes, but the output – the projected image – can be structured in ways that promote more or less conversation. For example, in lieu of displaying didactic text-based slides, teachers can display data in different graphical formats and invite students to attempt to interpret it (Chang, 2004).Supporting lectures with a shared visualization helps students maintain a joint focus on material, but by encouraging a more active approach to viewing the visualization, students are able to engage in a more active process of integrating the presented content with their already-extant mental models of the material. Eliciting participation in active learning processes is a strong focus of many CSCL designs. Rather than just inviting students to think about things, many CSCL applications actively enable (or enforce) participation.

One way to do increase participation is to reduce the number of students using the SDG to small groups, and approach that will be discussed in the next section. Another is to locate the "shared display" in a virtual space that all students can access via their own desktop computers. Although synchronous, and "co-located" in a virtual sense (seeing as avatars can "look" at one another) these environments are usually designed to encourage individual exploration, often of some sort of historically-based city (Di Blas, Poggi, & Reeve, 2006; Ketelhut, 2007). The tasks allow for open-ended exploration, which means that individual users aren't experiencing the same shared display at all, but operating on their own. In an effort to encourage more on-task interaction between students in a shared VR environment, some researchers experimented with activity structures. They found that activity structures that created positive interdependence between users resulted in an increase in on-task communication levels (Steiner & Moher, 2002). This is mentioned because positive interdependence will become a strong theme in Section 4.3.2 Handheld Devices in CSCL.

4.3.1.2 Small Groups using SDG

Sometimes regular desktop computers are used as SDGs, where small groups of two to four students will gather around a single computer, and use it as a focus for their attention and a locus for their participation. The idea is to shift students from a passive to an active stance by giving them more of a chance to interact with the SDG than a whole-classroom scenario of the sort presented in Section 4.3.1.1 Classrooms using SDG, would allow. The trouble with single desktops, though, is the same as that for single-user kiosks: their form factor inherently limits access to input and output opportunities. Moreover, the student who has access to the input device is often reluctant to surrender it to another user, even if he or she is not using it to provide input (S.D. Scott, Mandryk et al., 2003). In other instances involving dyads, although one user might be physically operating the input device, the other student may be ordering them around (Edelson & Reiser, 2006). The cognitive load inherent in both interacting with a user interface and listening to a partner's directions can prevent the operational user from engaging in higher-level thought processes on his or her own, so although the user is physically interactive with the SDG, he or she is not necessarily as *mentally* interactive with the problem scenario as the directing partner. The trouble with that is that the less interactive a participant is in a problem-solving scenario, the worse his or her individual learning outcomes (Mevarech, 1994). As with museum learning, it is not enough to be hands-on; the learning must also be minds-on.

Providing input devices (e.g., mice) to each and every student in the group is one way to try to equitably engage all learners in the activity. Much of the work with SDG and multiple mice has been tested in the context of open-ended, loosely-coupled shared drawing programs (Benford et al., 2000; Hourcade et al., 2002; Stanton et al., 2002). "Learning outcomes" for this sort of context are very difficult to define, let alone assess, although students do participate when they have access to input opportunities. In at least once case, an activity with more easily-assessed learning outcomes (mathematics skills) was tested with multiple mice, but problems with recording performance data prevented an analysis (Stacey D. Scott, Mandryk, & Inkpen, 2002). This latter activity was adapted from a single-mouse activity, and as such, required tight input coupling between participants. Thus, loosely-coupled inputs have been used for open-ended activities, and tightly-coupled inputs have been used for activities with clearer performance metrics.

Some desktop-centric SDG does not make use of multiple mice, however, and rely instead on activity structure to try to engage all participants. A good portion of work originating from Northwestern's Learning Sciences program and Berkeley's School of Education in the last 15 years has been concerned with structuring desktop-based activities to elicit "minds-on" participation from students. Berkeley's Knowledge Integration Environment (KIE) / Web-based Integrated Science Environment (WISE) system was not originally designed as CSCL software, but shortages of computers (microcomputers) in the early years often necessitated that desktops be used by small groups of users. In general, the various flavors of KIE/WISE attempted to engage learners in a self-directed inquiry learning process not terribly different from what museum visitors are encouraged to undergo (E.A. Davis, 2003; Elizabeth A. Davis, 2004; Linn et al., 2004). They found that if students were asked to specialize, to take on a role related to mastering one particular concept, their ability to integrate knowledge improves (Cuthbert, 1999).

The software programs arising from Northwestern, many under the aegis of the Learning Technologies in Urban Schools (LeTUS) center, were similarly devoted to inquiry learning, although with a slightly stronger emphasis on engaging students in authentic practices related to the scientific domain under study (Edelson et al., 1999; Edelson & Reiser, 2006). Like KIE/WISE, many of the LeTUS applications were used by multiple students at a single computer more by default than design. In response, the Progress Portfolio tool was created to help these groups (usually dyads) collaborate better (Kyza, 2002). The operating principle was to make the group members' efforts visible to one another.

4.3.1.3 SDG in CSCL: a Summary of Open Questions and Themes

The major theme driving the use of SDG in CSCL is the same as for CSCW: that a shared visual focus can help with conversational grounding. Students, though, must need more encouragement than workers to make use of the SDG for its intended purpose, because a great deal of CSCL SDG research focuses on ways to encourage students to become engaged actively with the shared context. Perhaps there is a presumption that the natural motivational structures present in the workplace (e.g., earning raises, keeping ones job) is enough to encourage worker participation in SDG-based activities. Three techniques in particular emerge as strategies to increase participation: using task structures with positive interdependence, assigning specialized roles to participants, and making the efforts of each group member visible. These themes will appear again in the next section on handheld devices in CSCL.

The coupling of output in the SDG paradigm is tight by definition, but as in CSCW research on SDG, CSCL SDG applications have some variety in the degree of coupling of input. Recall from Section 4.2.2.4 Single-Display Groupware: a Summary of Open Questions and Themes, that CSCW users of SDG systems behaved in a more self-interested manner when their inputs were loosely coupled. It seems to be the case for CSCL SDG applications that loosely-coupled input seems to get paired with open-ended learning outcomes, whereas tightly-coupled input is paired with specific learning goals. The implication one might make is that for learning scenarios where the *process* is more important than attaining a specific *outcome*, loose coupling of inputs is appropriate. There is not enough evidence to make any conclusions, though, and so the impact of input coupling on collaborative behaviors while using an SDG is still an open question.

4.3.2 Handheld Devices in CSCL

The relative portability, networkability, and affordability of handheld-based devices (e.g., handhelds, calculators, cellular phones) *vis-à-vis* desktop computers makes them an attractive form factor for CSCL researchers (Roschelle & Pea, 2002). By providing a handheld device to each and every learner, IO access problems are certainly avoided. In this context, researchers are free to think about how to structure activity patterns for collaboration that are flexible along many different dimensions, like time, location, and the person-to-person connection "topology." With such freedom, careful though must be given to attention management, as well as the degree of coupling between different learners (Roschelle & Pea, 2002). In particular, learners need to be able to shift between "private" interactions with their handheld device and "public" interactions with their co-learners smoothly (Vahey, Tatar, & Roschelle, 2007). This section will describe some of these collaborative "topologies," trying to highlight how learner attention is (or is not) managed.

Many handheld applications designed for use in classrooms have only a small emphasis on collaboration, limiting student-student interactions to the sharing of data. Sometimes data (especially student-created artifacts like concept maps) are shared synchronously and directly between devices (Curtis, Luchini, Bobrowsky, Quintana, & Soloway, 2002; Luchini, Quintana, & Soloway, 2004), but more often the sharing takes place asynchronously by placing it on a central server for later retrieval by others (Lundby, Smørdal, A., & Fjuk, 2002; Milrad, Perez, & Hoppe, 2002). When the primary IO is conducted solely through the handheld device and collaboration episodes are limited to sporadic data sharing, attention management is not much of a concern – although technically qualifying as CSCL applications, the collaboration is effectively conducted asynchronously. For this type of work, the inputs and outputs of the students are naturally very loosely coupled.

Handhelds, unlike desktop-based UIs, allow input/output coupling not just to take place between would-be collaborators, but also between the handheld and the physical environment it is places in. Probeware, the use of handheld computers to support in-the-field data collection (Farooq, Schafer, Rosson, & Carroll, 2002; H. Smith, Luckin, Fitzpatrick, Avramides, & Underwood, 2005), strongly couples handheld IO to the physical context from which data is being collected. In a sense, then, probeware, and the related "augmented reality" handheld applications (Benford et al., 2005; Cole & Stanton, 2003; Eric Klopfer, Perry, Squire, & Jan, 2005a; Squire & Jan, 2007; Squire & Klopfer, 2007), are providing "portholes" onto a shared context, as some of the CSCW MMUI applications do in Section 4.2.2.2 Output from Single-Display Groupware. While probeware provides a "porthole" onto data not normally visible to the human eye (e.g., water temperature in streams), augmented reality applications provides a "porthole" onto data that may be wholly imaginary, like the tracks of a creature called a "snark," the path of leaking toxins in a simulated chemical spill, or a simulated African savannah.

By and large neither probeware nor augmented reality researchers looked at issues surrounding user attention management or encouraging user participation, but one research team created a pair of augmented reality applications, one which displayed dynamic event-driven information, and one which displayed static, user-initiated information. They noted that in the case of the dynamic application, users would become "engross[ed] in the technology," often ignoring their collaborative partner, but by designing the second applications interface to show only static information, they were able to alleviate the problem (Cole & Stanton, 2003). This is in strong accordance with the vague suggestions emerging from a review of handheld UI design for museum use, suggesting that "simple" UIs may alleviate the heads-down effect (see Section 3.3). There is a suggestion that the other augmented reality researchers may have overcome the heads-down effect by structuring their activities to require positive interdependence in task execution, combined with individual accountability so that all group members were aware of one another's contributions (Eric Klopfer, Perry et al., 2005a; Squire & Jan, 2007; Squire & Klopfer, 2007). In this case, the inputs and outputs were not strongly coupled, but the *tasks* permitted to the users were strongly coupled – in order to accomplish tasks, each member had to participate, creating positive interdependence.

Many other handheld-based activities use task structures with positive interdependence to encourage participation. Some are fairly simple in establishing interdependence: this small-group discussion tool flatly requires participation, preventing users from moving on until all participants have weighed in on the matter (Cortez et al., 2004). The problem with this sort of design is that it can be subverted – students can engage physically (clicking on the agree button) without engaging mentally (actually deciding to agree with the rest of the group). Others establish interdependence via hidden information, like some of the mathematics learning exercises used in NetCalc, which require users to play a "Mastermind" like guessing game regarding their partner's "secret" line equation (Vahey, Tatar, & Roschelle, 2004). Yet another approach to establishing positive interdependence is to give each and every participant a different but unique role, an approach known as "jigsawing" (Aronson et al., 1978). Sometimes jigsawing is accomplished by giving each participant a different, but equivalent "piece' that needs to get arranged *vis-à-vis* the pieces assigned to others. For example, this handheld-supported concept-mapping exercise requires each student to generate a map covering a portion of the content, which will then be collected and quite literally assembled in a jigsaw fashion (C.-Y. Lai, Wu, Kao, & Chen, 2005; C. Y. Lai & Wu, 2006). Similarly, in language activities each student is given a distinct phoneme, which must be used with those provided to partners to assemble words (Zurita & Nussbaum, 2004). Other times, jigsawing is accomplished by giving participants a distinct role to play in the context of the activity, as in these two code-breaking activities where each user plays a role (e.g., "Presenter," "Publisher," etc.) that required a wholly different relationship to the data (Goldman, Pea, & Maldonado, 2004; White, 2006). Sometimes these roles can be "pipelined" such that one user's ability to act directly relies upon another player completing his or her role (DiGiano et al., 2002).

Participatory simulations are perhaps the strongest example of establishing positive interdependence between students: the activity is whole comprised of and defined by student-student interactions. The idea behind participatory simulations is a simple one: users assume the role of an entity in a simulation, and must act out that role with the help of a portable device that helps manage the details of the simulation (Colella, 2000; Danesh, Inkpen, Lau, Shu, & Booth, 2001; E. Klopfer & Yoon, 2005; Mandryk, Inkpen, Bilezikjian, Klemmer, & Landay, 2001). So, for example, if one is playing the role of a fish tasked with passing on one's genes, one must "mate" with students playing the role of other fish by transmitting data from one device to another. The task – mating – requires at least two participants to work; the task would not exist without the participation of the players, therefore has a very high interdependence.

4.3.2.1 Handheld Devices in CSCL: a Summary of Themes and Open Questions

A slew of themes emerged from this review, many of them relating to engineering the activity structure to encourage more participation in the joint activity. As in Section 4.3.1.3 SDG in CSCL: a Summary of Open Questions and Themes on SDG in CSCL, positive interdependence and the assignment of specialized roles appeared as suggested methods for encouraging participation. A few additional means for encouraging participation by establishing positive interdependence also emerged here: by fiat (which may or may not work well), by hidden information, by assigning unique (but not hidden) pieces to participants, by pipelining stages of the activity, or by structuring the activity itself out of player-player interactions (as in participatory simulations).

Compared to the amount of attention given over to activity structure designs, the user interface designs of the handheld applications were virtually ignored. There was a suggestion that dynamic handheld interfaces can lead to the heads-down effect, whereas simpler static interfaces promote more interaction with collaborators and context. It seemed that the outputs for the handheld activities were loosely coupled at most (as when the handhelds served as "portholes" onto a shared view of data in probeware or onto a virtual world in augmented reality applications). In some cases the loose coupling was viewed as a problem – in one participatory simulation activity, the researchers experimented with "merging" the handheld outputs into a larger tightly-coupled display to help participants better interpret the activity (Mandryk et al., 2001). In other cases the loose coupling was viewed as an asset, especially when it allowed users to have different views of the shared context that would support their different roles in the activity structure (Goldman et al., 2004; White, 2006).

The coupling of inputs also tended to be loose, although some designs capitalized on device-device communications (e.g., IrDA beams) to force a tight coupling for certain types of input actions. For example, in participatory simulations, input was tightly coupled – no user could provide "input" to the simulation without the participation of another user. Unlike SDGs in CSCL, where there seemed to be a matching of loosely-coupled inputs to open-ended learning goals (see Section 4.3.1.3 SDG in CSCL: a Summary of Open Questions and Themes), the handheld activities here were not consistent about paring input couplings with goal types. Participatory simulations had the tightest of input couplings but relatively open learning goals, whereas some of the activities with well-defined goals like code-breaking or determining water quality had some of the loosest input couplings. With the handheld activities, it seems, input coupling was not used to encourage or enforce collaboration; rather, the design of the activity structure was used for that purpose.

4.3.3 MMUIs in CSCL

One of the handheld-based activities in Section 4.3.2 Handheld Devices in CSCL was modified to allow the handhelds to be linked together to form a larger, shared display space to better support group discussions (Mandryk et al., 2001). At least one other research group took a similar tack, allowing users of tablet PCs to link them together to form a single larger display. This was done after researchers noticed that students would gather around one of the tablets when they needed to discuss ideas, even though each student had a tablet of his or her own (Deng, Do, Chang, & Chan, 2004). What's more, most of the augmented reality applications described in Section 4.3.2 Handheld Devices in CSCL had a "debriefing" phase after students got back to the classroom, using a projector to display data from the augmented reality episode to help the class engage in discussions. It seems from these examples that some types of collaborative activity, especially *discussions*, are best supported by a single, shared display. And yet, individual devices allow for much more *participation*. As discussed in Section 3.4.4 Multi-Machine User Interfaces in Museums: a Summary of Group Learning Opportunities, the promise of MMUIs is a sort of “best of both worlds” combination: the grounding benefits of a single shared display combined with the accessibility and participatory benefits of handhelds. The different MMUIs experimented with in CSCL show a wide range of activity structures, with consequently very different emphases on public and private interfaces.

Some MMUIs use handheld devices to act more as remote controls than anything else, with little to no information being displayed on the handheld screens. Examples include the NetLogo-based activities that made use of graphing calculators to allow whole classes participate in classroom-wide exercises (Resnick & Wilensky, 1998; Wilensky & Stroup, 1999). Some of these exercises involved one-time inputs, as when each student would guess about the one of a point’s coordinates when supplied the other coordinate and the line equation. The collective results would be displayed on-screen, so the teacher and the class as a whole could see the degree of agreement between participants. This is the same idea behind classroom response systems, which use devices as varied as laptops and special-purpose clickers to allow classes to respond to quizzes in the midst of teacher lectures (Vahey et al., 2007). Although the SDG applications described in Section 4.3.1 SDG in CSCL *relied* on making the attribution of individual efforts public in order to motivate participation, for these MMUIs, the opposite strategy seems to hold. When students were interviewed they reported that they found the anonymity of performance on the group display to be important to their desire to participate (Curtis et al., 2002). Curiously, though, the same study reported that students expressed a desire to know "where they were" on the group display. The students seem to want to know how their actions are impacting the shared activity.

A slightly more involved activity structure is that used by a control-of-variables exercise (Moher et al., 2003), a traffic control exercise (Wilensky & Stroup, 2000), and a cardiovascular system control exercise (Alexander Repenning & Ioannidou, 2004; A. Repenning & Ioannidou, 2005). In each of these, a jigsaw activity structure is used, where each student is given ownership over a particular element of an ongoing activity displayed on the shared display (usually a projected image at the front of the classroom). The students must coordinate their actions to attain a desired joint outcome – whether it be flipping a field of tiles to display a single color, changing the color of traffic lights to regulate the flow of cars in a city, or controlling the pulse and respiration of a human body, respectively. Like the response-style systems described prior, the handheld devices don’t display much on their screens – if used at all, the displays are simplistic. Unlike the prior systems, though, identifying which student is responsible for which input is not just important to each of these activities, it is a critical component. So it may be the case that anonymity is only needed for “one-off” MMUI activities like those described above. When an activity is continuing, and especially when it involves a coordination of efforts, the MMUI should make it possible for students to identify their individual contributions to the shared activity.

The MMUI activities discussed up until this point have involved the use of handheld devices in a very unidirectional fashion: all input is provided via the handhelds and immediately conveyed to the activity on the shared display. Owing to their ability to display images as well, though, there is no reason why information cannot be taken from the shared display and placed on the individual devices, like the "snarfing" discussed in Section 4.2.2 Multi-Machine User Interfaces (MMUIs). For example, the Group Scribbles application (Roschelle et al., 2007), allows students to both post and take Post-It™-like notes from the Shared Display for further manipulation on their individual devices (tablet PCs in this case). Several Group Scribbles activities have been devised around this ability to both push and pull data. Similarly, the collaborative small-group applications designed for the MUSHI (Multi-User Simulation with Handheld Integration) framework supported similar push and pull activities (Lee et al., 2005; Lyons, Lee, Quintana, & Soloway, 2006a, 2006b; Vath et al., 2005). For example, the MUSHI-Life simulation was designed to allow users to capture information (and sometimes entire organisms) from one ecosystem displayed on an SDG, and (in the case of the organisms) deposit them in a different ecosystem displayed on another SDG. With both of these applications, the displays on the personal devices are employed to allow users to manipulate or study snippets of data originally made available on the SDG, and possibly reintroduce them to the shared display afterwards. Because there exists no name for this collaborative use pattern, it will be dubbed the "pull-and-push" activity structure.

"Pull-and-push" structures do not tightly couple students' private inputs or outputs; although the SDG display is shared, the private devices are very loosely coupled to it and to each other. As a consequence, users may need other activity structures to motivate participation and to help them to divide their attention between private and public displays. In the case of Group Scribbles, the teacher helps marshal the students' attention, declaring when students should attend to their personal devices, and when they should attend to the SDG. The MUSHI-Life application employed a "porthole" style division of output between the handheld devices and the SDG, where each handheld displayed a highly dynamic "zoomed-in" region of the ecosystem. Although the augmented reality users in Section 4.3.2 Handheld Devices in CSCL got lost in the heads-down phenomenon when using dynamic personal displays, the participants in the MUSHI studies were able to successfully divide their attention between the private handheld devices and the public SDG, and moreover, the levels of participation were relatively even within each dyad tested. That said, the tasks that the users were engaged in during the MUSHI studies were of short duration, and were marshaled by the researchers, who moved the participants from one phase of the activity to the next.

This section has been exploring MMUIs in CSCL along two continuums: one of information flow directionality, from push to pull-push to just pull (as we will see shortly), and one of duration, from very short collaborative activities to activities that require more time. The longest of all are probably "embedded phenomena," wherein multiple displays are mounted throughout a classroom to provide "portholes" onto a simulated world (Moher, 2006; Thompson & Moher, 2006). Some of the simulations involve insect swarms, earthquakes, and planetary orbits. In these MMUI activities, an event (either a teacher's call-to-action, or a simulated "earthquake," or the like) spurs students to claim an MMUI as an observational post, from which they "pull" data to be used in larger, non-technologically enhanced activities, like plotting a point on a paper chart. Embedded phenomena activities run in the background during normal classroom activities for several weeks at a time. There coupling of input and of output across devices is very loose and is largely immaterial – the MMUIs are used, like those in the MUSHI-Life studies, as a way to capture observational data from the shared simulation. In lieu of using coupling to get students to participate, jigsaw-style activity structures are employed to engage students in the activity.

Returning briefly to the information flow – it seems that when students are expected to have briefer interactions with an MMUI, a "push" model of information flow, moving information from private to public devices, is often used. When students are expected to have lengthier interactions with the MMUI, the information flow becomes bidirectional or even reverses into a "pull" model. The information flow patterns indicate how work is fissioned between the private and public devices – increasing the degree of "pull" also increases the amount of time spent in "private" work tasks. The onus for keeping students engaged and participating falls increasingly on external structures (curriculum, directions from teachers or other activity leaders) as the emphasis shifts from push-to-public data flow to a pull-to-private model of data flow.

4.3.3.1 MMUIs in CSCL: a Summary of Themes and Open Questions

Far fewer MMUI systems have been implemented in classroom environments than either SDG or handheld-only systems, so the themes are not as clear-cut. Like handheld-based applications, activity structures were employed to keep students engaged in the activities, but unlike handheld-based applications, MMUIs often relied on external structures (like guidance from an observing teacher) to marshal the actions of the students, rather than just incorporating the activity structures into the design of the software itself. With the addition of both private and public display devices, designers must consider how to "fission" the output onto the different displays. All methods suggested in the CSW-oriented Section 4.2.2.2 Output from Multi-Machine User Interfaces are used here: outputs are *stretched* across displays to make impromptu shared workspaces, sometimes the different displays are used to show *complementary* information, and sometimes the private devices are used as *portholes* onto a larger shared context. As with the CSCW research, no real attention was paid to the effect that these different types of fissioning have on student attention management.

4.4 Summary

The review of CSCW and CSCL research provided several framings that are of use when trying to conceptualize multi-user software-based museum exhibits. First, multi-user museum exhibits can be defined by the spatial and temporal properties of the collaboration intended for them; in this case, the collaboration is co-located (the same place) and synchronous (the same time). By narrowing the field of applicable prior work to co-located, synchronous applications, it is easier to pull out themes and open questions relevant to the research at hand. Secondly, a review of the theoretical underpinnings of CSCW and CSCL shows that to the extent that they are motivated by different design philosophies (User-Centered Design and Learner-Centered Design, respectively), one should be careful in how those themes and open questions are interpreted.

The review of co-located, synchronous CSCW applications revealed that the coupling of individual users' input and output is often used as a strategy to encourage collaboration. When users are provided with loosely-coupled output, especially when that output is detailed and is delivered via a private interface, they are prone to getting lost in the heads-down phenomenon. When users are provided with loosely-coupled input, they are more prone to act in the service of individual self-interest, rather than for the group's benefit. One research group found that limiting input to a single user (rather than providing mice to all users) improved collaborative discourse. If only CSCW research was reviewed, one might conclude that very tightly-coupled, WYSIWIS (What You See Is What I See) systems are the best for supporting collaborative work.

The review of CSCL applications revealed a different goal structure, though – although collaboration was valued, so was the participation of each and every student. The desire to engage every student in the learning activities led researchers to try to provide each student with access to both input and output, and also to strongly rely on activity structure design to encourage participation. The nearly-universally acknowledged strategy to encourage collaboration was to create an activity structure with positive interdependence between students. Several methods of establishing positive interdependence were used, but the jigsaw method proved to be the most popular (and presumably, the most successful).

The CSCL applications showed wide variations in input and output coupling – unlike many CSCW applications, tighter coupling was not necessarily acknowledged to result in better collaboration. The implication was that looser coupling of input and output would encourage more engagement on the part of the students, by allowing them more autonomy, and that positively interdependent roles would encourage the students to collaborate even if using loosely-coupled devices. Although loose coupling of output was not definitively shown to result in the heads-down phenomenon, one study showed that when the loosely-coupled output provided to handheld devices was more dynamic and detailed, the heads-down effect occurred, but when the output was "simpler" and static, it did not. CSCL applications also showed that privacy is sometimes a concern – sometimes learners may be embarrassed about their performance, and thus wish to preserve anonymity on any shared display. For activities where coordination is required, though, user contributions must be identifiable.

From the review in Chapter 3 of computing technology in museums, a fair number of design recommendations and several open questions were uncovered. Very few, if any of these, have been directly investigated by CSCW or CSCL research, but many of the same themes are present in all venues. The heads-down phenomenon, although not labeled as such, seemed to appear in CSCW and CSCL applications when outputs were loosely coupled, dynamic and/or detailed, and were placed on private, handheld devices. This parallels the observations of museum professionals that "simpler" handheld UIs were better for staving off the heads-down effect. Much of CSCL research used loosely-coupled inputs and outputs to encourage engagement in the learning activity, however. It us unclear if, in a museum context where the paradigm is self-directed learning (not compulsory learning), these same strategies apply. Does looser output coupling still result in more engagement, when the context is a museum and not a classroom? If visitors are engaged more by loosely-coupled handheld devices, are they also more prone to the heads-down effect? Museums lack an authority figure, like a teacher, to marshal the attention of the learners. Is a positively-interdependent activity structure enough to shake them from any heads-down behaviors? All of these are open questions.

CHAPTER 5

Design Rationale and Implementation

The design of the computer-based exhibit researched in this work had two main drivers: (1) to construct a computer-based exhibit that would support multiple users and their collaborative learning processes, and (2) to use this exhibit to examine some open questions related to placing computers in the context of a science museum floor. From Chapter 2, the reader should now have an understanding of how the cultural-historical, social, and physical contexts found in science museums can shape visitors' expectations and needs regarding computer-based exhibits. The recommendations that emerged from that review can be found in Section 2.4. From the review of existing computer technology in museums in Chapter 3, the reader should have a sense of how computers have been and are currently in use in museums. A series of design strategies that emerged from the review, and a listing of open questions, appears in Section 3.5. Chapter 4 took a more theoretically-grounded review of the design of multi-user, co-located, synchronous collaborative software, situating different design paradigms within the schools of thought that spawned them.

This chapter will cover the design of the software used in the experiments described in Chapter 6. The aim of the first several sections of this chapter is to convey a detailed understanding of the software. The different software components, and how they were designed and built, will be described in Section 5.1 Description of Software Design. Because the software is distributed across multiple devices (a server that drives a large display, and multiple handheld devices), the server-based portion of the software will be described in Section 5.1.1, and the handheld-based portions of the software are described in Section 5.1.2. A use case, which illustrates in narrative form how the software might be operated in an ideal case, is presented in Section 5.2. A discussion of this use case will reference the relevant themes and prior work from Chapters 2, 3, and 4 as needed to highlight the design decisions.

5.1 Description of Software Design

This section will describe the software used in the study presented in Chapter 6. The description here will be given in a somewhat workman-like manner – no discussion of the design strategies or underlying theories, just a bare-bones description of what the software is, a few details on how it was constructed, and how users operate it. A larger discussion of the design decisions will be saved for Section

5.2. The single exception is Section 5.1.1.1, which provides some necessary background information on the underlying type of simulation used by the software: cellular automata.

5.1.1 *MUSHI-Lignancy* Simulation

MUSHI-Lignancy is a multi-user role-playing game centered on an open-ended simulation of a complex system: namely, a simulation of cancer growth in human tissue. It was designed and built by the author for the second-generation MUSHI (Multi-User Simulation with Handheld Integration) platform that was co-created by the author and Joseph Chigwan Lee. The *MUSHI-Lignancy* simulation was designed to be an abstract representation of a patient suffering from cancer, and to exhibit emergent phenomena that occur within real cancer patients: tumor growth, angiogenesis, metastasis, and secondary cancers owing to treatment-related DNA damage[22].

5.1.1.1 Cellular Automata

The *MUSHI-Lignancy* simulation is implemented as a collection of cellular automata. It is important for the reader to know a little about cellular automata and their place in education, so this section will offer a brief review.

Mathematician John Conway first introduced cellular automata with his seminal *Game of Life* in the late 1960s, wherein dots in a grid would change color from black to white and back, using just a simple set of rules and the state of the surrounding dots to govern the color selection for a given dot (Gardner, 1970). Despite the simplicity of the rules, rather than displaying predictable outcomes or just chaos, the *Game of Life* exhibited a host of unexpected “emergent phenomena.” Emergent phenomena can be thought of as larger patterns that can be seen, or “emerge,” as a result of the combined interactions of very many smaller elements (Holland, 1998). In the case of the *Game of Life*, simple rules about whether to change color or not resulted in emergent patterns given names like “blinker,” “glider,” or “beehive,” that would persist, replicate and interact with one another in interesting ways (see Figure 58).

[22] The different medical phenomena mentioned here will be described in depth later on, when describing the specifics of the simulation.

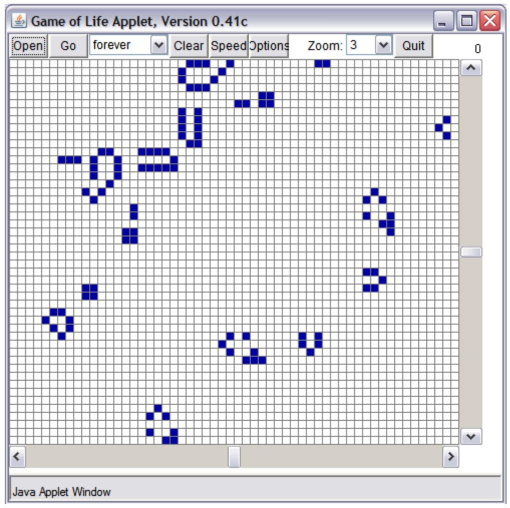


Figure 58. A screenshot of a *Game of Life* applet, taken from (Callahan, n.d.). The squares in the grid will change color, from white to blue or vice versa, on the basis of the colors of the adjacent squares. The set of simple rules controlling color change, combined with a random initial assignment of colors to the squares, can result in interesting phenomena, many of which are self-replicating.

The canonical example of a complex system is that of an anthill: individual ants are not intelligent, but by applying simple rules (smell pheromone A, go left; smell pheromone B, go right) they can collectively construct complicated anthills with separate chambers for food, waste, and pupae. Most scientific phenomena can be viewed as complex systems, from natural selection to protein folding bacteria colonies to global warming (see). Viewing science topics through a systemic lens (as opposed to the more traditional Western reductionist lens) has gained popularity with both professional scientists and science educators alike (American Association for the Advancement of Science, 1993). A great many educational researchers have begun to explore how to teach people about complex, emergent systems (Goldstone, 2006; Cindy E. Hmelo-Silver & Azevedo, 2006; C. E. Hmelo-Silver & Pfeffer, 2004; M. J. Jacobson, 2000, 2001; Michael J. Jacobson & Wilensky, 2006; Lesh, 2006; Liu, Hmelo-Silver, & Marathe, 2005; Liu, Marathe, & Hmelo-Silver, 2005), and not unsurprisingly, many have turned to computer simulations of complex systems to help with the task (Goldstone & Sakamoto, 2003; E. Klopfer, 2003; E. Klopfer & Yoon, 2005; E. Klopfer, Yoon, & T., 2005; Penner, 2000; Wilensky & Stroup, 2000).

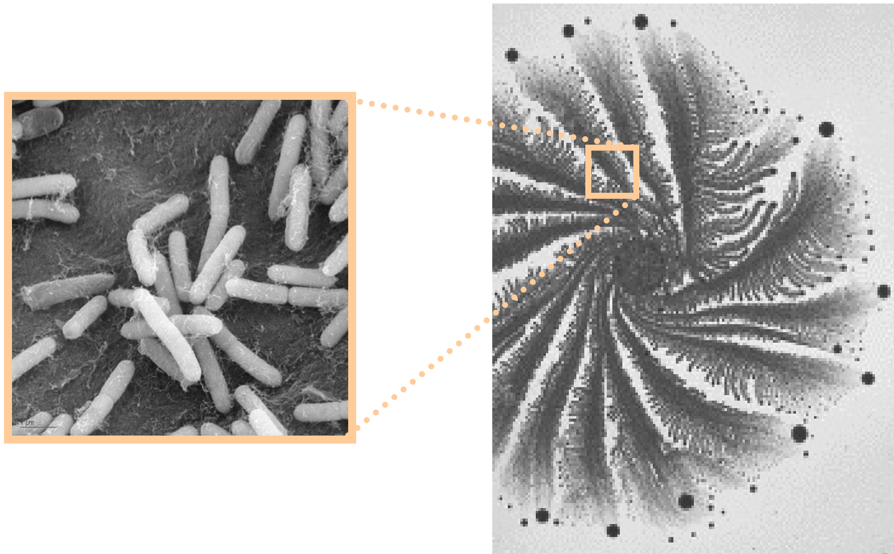


Figure 59. The photo at the left is an image of bacteria. The photo at the right is an image of an entire colony of bacteria, illustrating the complex, emergent pattern that forms when the bacteria is allowed to grow in an agar medium.

5.1.1.2 MUSHI-Lignancy as a Cellular Automaton Simulation

In *MUSHI-Lignancy*, the three types of automata are cancer cells, healthy cells, and blood vessel segments (see Figure 60). Each simulated automaton maintains information about its current state in the form of variables. For example, a healthy cell will maintain a current "health" variable, a variable that tracks cell age, and a "cumulative radiation exposure" variable. Each automaton shares a rule base with other automata of its type (i.e., all healthy cells obey the same set of rules, all blood vessel segments obey their own shared set of rules, etc.). There is no "controlling hand" to the simulation – the next state of the simulation is an outgrowth of each automaton performing its own state update.

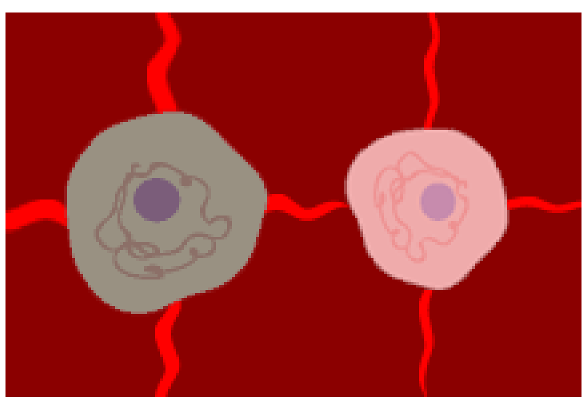


Figure 60. Illustration of the three key elements of the MUSHI-Lignancy simulation: cancer cells, healthy cells, and blood vessel segments. On the left is a cancerous cell, identifiable by its grayish-green pallor. On the right is a normal cell. The red "squiggles" are blood vessels.

MUSHI-Lignancy uses a grid of interconnected cells to represent the simulated "patient" suffering from cancer. The cells are placed so that they are centered on the intersections of a grid, with the blood vessel segments connecting at the grid's intersections (see Figure 61). All cells are dependent on blood flow for survival, but they share blood vessels and thus implicitly compete for blood supply.

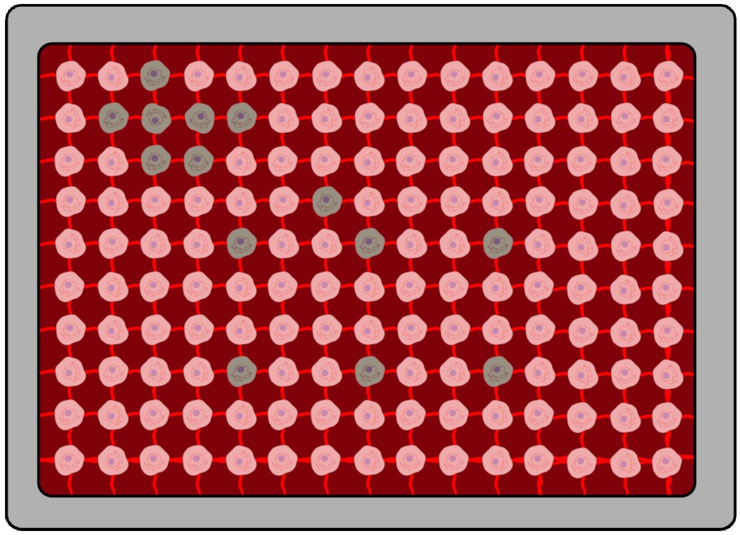


Figure 61. Illustration of the MUSHI-Lignancy simulation, as sown on the large shared screen. Cells are arranged in a grid, connected via blood vessels. Clusters of the darker cancer cells act as "tumors."

When it comes time to update the state of an automaton, the automaton uses its current state and other data (e.g., the state of neighboring automatons) as fodder for its rules, which will be applied to calculate the particulars of the automaton's next state. By carefully selecting a stable of state variables and tuning a collection of rules, it was possible to create a simulation that is relatively simple in its definition, but nonetheless exhibits several complex emergent phenomena that are hallmarks of cancer in real life. The *MUSHI-Lignancy* automata have a total of 11 different variables (see Table 4), and yet the simulation exhibits analogues to real-life phenomena like tumor growth and its associated angiogenesis, metastasis, and radiation-induced secondary cancers.

Table 4. Variables used in the MUSHI-Life simulation.

Cell Variables		**Blood Vessel Variables**	
Name	**Type: Range**	**Name**	**Type: Range**
`Cell_Type`	enum: normal or cancer	`Vessel_Flow`	double: 0 – 1
`Cell_Health`	double: 0 – 1	`Vessel_Growth`	double: 0 – 1
`Cell_Growth`	double: 0 – 1	`PersistentDamage`	double: 0 – 1
`Cell_GrowthDir`	int: 1 or -1	`CANCER_GROWTH`	double: >1
`PersistentDamage`	double: 0 – 1		
`RadiationExposure`	double: 0 – 1		
`GROWTH_INCREMENT`	double: 0 – 1		

An example of how an automaton state-change works will be provided by first describing a phenomenon, angiogenesis, then describing how the phenomenon is represented in the *MUSHI-Lignancy* simulation, and finally describing the mechanics behind the simulated phenomenon. Within real bodies, cancer cells will emit chemical signals that trigger nearby blood vessels to grow at a rate than normal. (Healthy cells also emit these chemical signals, but not at nearly the same volume as cancer cells). The growth of new blood vessels is known as *angiogenesis*. Angiogenesis in turn allows clusters of cancer cells (known as “tumors”) to reproduce more rapidly by supplying the blood needed for cell growth. (This is the reason why many cancer treatments are “anti-angiogenetic” – the idea is that by interrupting the growth of new blood vessels, cancer tumors will be starved of the blood they need to grow. Healthy cells, which rely less on new blood vessels than tumors do, will not be as affected).

$$\texttt{Vessel\_Growth} = \frac{\sum_{n} \texttt{n.Health} + \sum_{c} \texttt{CANCER\_GROWTH} \cdot \texttt{c.Health}}{\sum_{n} 1 + \sum_{c} 1}$$

Where `n` ∍ {normal neighboring cells}
and `c` ∍ {cancerous neighboring cells}
and `CANCER_GROWTH` *is a constant*

```
next_state_vessel_flow  = Vessel_Flow
                        + normalized_Vessel_Growth
                        - PersistentDamage
```

Figure 62. Equation used to calculate a blood vessel segment’s flow. The health of the surrounding cells is used as a proxy for how much pro-angiogenetic chemical signaling a blood vessel segment would receive. The health of the cancer cells is multiplied by a constant, CANCER_GROWTH, to emulate the increased signaling cancer cells perform. The next blood flow state for a segment is computed by adding (a normalized version of) the growth to the current Vessel_Flow, and subtracting any persistent damage.

MUSHI-Lignancy visually represents angiogenesis, the growth of new blood vessels, by changing the apparent "thickness" of the blood vessel segments, which directly corresponds to the segment's `Vessel_Flow` variable. Blood vessel "thickness" is an abstracted way to indicate the overall amount of blood flow passing through the region represented by the blood vessel segment. So, if a blood vessel segment automaton is next to a thriving cancer cell automaton, it is likely to grow in "diameter" in order to meet the cancer cell's need for blood flow (notice that the blood vessel segments adjacent to the cancer cell in Figure 60 are larger than the segments abutting the normal cell). The equation that explains how this might occur is detailed in Figure 62. When the segment is surrounded by greater numbers of cancerous cells, it will have a higher `Vessel_Growth` value, which is an analogue for the amount of pro-angiogenetic signals the segment is receiving. There is a fixed amount of blood available for the "body", which gets distributed via the interconnected blood vessel segments. After all of the blood vessel segments have computed their preferred next-state level of blood flow, the fixed amount of blood in the patient is apportioned out proportionally to the requests, giving each segment a `normalized_Vessel_Growth`. The blood vessel's next state for blood flow is based on the current `Vessel_Flow`, decreased by whatever `PersistentDamage` might have been done to the segment, and increased by `normalized_Vessel_Growth`.

```
Cell_Growth =  Cell_Growth
            +  (Cell_GrowthDir * GROWTH_INCREMENT);

cell_fitness  = Cell_Growth - PersistentDamage;

next_state_cell_health  =  HEALTH_FACTOR * Cell_Health
                        +  BLOOD_FLOW_FACTOR * cur_blood_flow
                        +  FITNESS_FACTOR * cell_fitness;
```

where `GROWTH_INCREMENT` is a variable set when the cell is created, and `HEALTH_FACTOR`, `BLOOD_FLOW_FACTOR`, and `FITNESS_FACTOR` are constants that sum to 1.0

Figure 63. Equations used in MUSHI-Lignancy to compute the next-state attributes for a cell. Cell_Growth represents the amount of growth a cell is attempting to accomplish. A cell has a natural tendency to either wax in growth (when Cell_GrowthDir is positive) or wane in growth (when Cell_GrowthDir is negative). This way, we can simulate a cell's "natural" lifespan. A cell will wax or wane by a given GROWTH_INCREMENT each update, which is a bounded randomized number assigned when the cell is created. Even if the cell's growth would otherwise be on the upswing, PersistentDamage can impact the. growth as shown. The next cell health is computed by combining a ratio of the current cell health, the current blood flow, and the current fitness (a composite of the growth and the persistent damage).

The relative "health" or vitality of the cells is also indicated by diameter: as cells near death, they shrink in size, and as they gain vitality, they wax large, in direct correspondence to the `Cell_Health` variable. A cell's `next_state_cell_health` is computed as shown in Figure 63, using a weighted sum of its current Cell_Health, the current amount of blood flow (`cur_blood_flow`, computed by averaging the `Blood_Flow` of neighboring blood vessel segments) , and a measure of the cell's "fitness." The notion of cell "fitness" was incorporated to give cells a life cycle: in the absence of any interference, they are "born," grow for a while until they reach a peak of fitness, and then they "waste away" and die. In real bodies, cancer cells have a much longer life cycle – part of the problem with cancer cells is that they often will not undergo normal apoptosis, or cell death. So in *MUSHI-Lignancy*, cancer cells have a longer life cycle than normal cells, which is enacted by giving them a smaller `GROWTH_INCREMENT` constant. (What is not shown by Figure 63 is that once `Cell_Growth` reaches or exceeds a value of 1, `Cell_GrowthDir` becomes negative).

Cancer cells spread in *MUSHI-Lignancy* in one of three ways. The first emulates tumor growth, wherein cancer cells reproduce in a fast-growing cluster. If a cell is next to several cancer cells, when it updates its state, there is a probabilistic chance (that is based on its own health, and the health of the neighboring cancer cells) that it will be "supplanted" by a cancer cell. Because *MUSHI-Lignancy* is grid-based, the uncontrolled growth normally seen with tumors cannot be replicated, but by having "weaker" normal cells surrounded by "stronger" cancer cells get "replaced" by cancer cells, tumor-like clusters of cancer cells will form on the grid of simulated cells.

The second means of cancer propagation emulates metastasis, wherein cancer cells can travel through the body (usually via the bloodstream) and appear elsewhere, creating new tumors. When a cell reaches a "natural death" in the simulation, by virtue of its `Cell_Growth` variable being decremented to a value of 0 or below, the likelihood for it being replaced by a cancer cell as a result of metastasis is calculated. Metastasis was initially implemented via a rather complex procedure that involved generating a measure by computing a composite distance to cancer cells, weighted by the blood flow of the narrowest blood vessel segment along the shortest path that would connect the dying cell to the cancerous cell. This computationally-intense procedure was replaced by basing the probability on the number of cancer cells present within the simulated body, to no observable difference.

The third means of cancer propagation only occurs if the simulated patient has been exposed to radiation (how this exposure occurs will be discussed in the next section). Every time a radiation exposure occurs, a cell's `RadiationExposure` variable is increased. As with metastasis, when a cell reaches a "natural death" in the simulation, the likelihood that a secondary form of cancer will appear due to radiation exposure is calculated, using the cell's `RadiationExposure` as a basis.

5.1.1.3 *MUSHI-Lignancy* Implementation

The *MUSHI-Lignancy* simulation was implemented using Microsoft's C# language, and can be run from any wifi-capable computer. For the experiments, a Compaq tc4200 tablet computer was used to run the simulation, and a large 48" plasma screen was used to display the visualization. The simulation inherits from the MUSHI game engine, which manages the "game loop" by periodically interfacing with the networking layer (MUSHI-Chatter), handling the drawing operations for the display, and prompting the simulation's "sprites" to update themselves on a fixed schedule (see Figure 64). The "sprites" present in

MUSHI-Lignancy are the grid of automata and the entities that represent each player and his or her actions. The automaton grid is responsible for marshalling the *n* cells and *m* blood vessel segments in the current instantiation of the simulation. For the experiments reported on in Chapter 6, a grid size of 16 by 16 cells was found to achieve the optimal balance between the richness of emergent phenomena and pace (recall from Section 2.2.2 Group Behaviors at Science Museums, that groups typically stay at an exhibit between 1 and 4 minutes). The goal was to allow (in the absence of any intervention by the users) phenomena like tumor growth and metastasis to occur, and for the "patient" to come to a simulated "death" within a 2-3 minute window. (The "patient" would "die" when more than 50% of the cells in the grid became cancerous).

The networking layer (MUSHI-Chatter) was designed to allow multiple users log into the *MUSHI-Lignancy* simulation using wireless-capable devices. (The software design for these devices will be discussed in Section 5.1.2 User Roles in MUSHI-Lignancy). MUSHI-Chatter was built using a "lossy" custom UDP-based (User Datagram Protocol) network communication protocol. Unlike TCP (Transmission Control Protocol), UDP does not have any built-in mechanisms for confirming the arrival of the data packets. At the time MUSHI-Chatter was designed, wifi-capable handheld devices were notoriously unreliable at maintaining network connections, so rather than assume the overhead of establishing and maintaining TCP connections, we decided to pursue a "fire and forget" strategy with regards to datagrams. Thus, the server is the authority on all current simulation state data, and the devices submit input events and receive update events via UDP packets. Like many developers who attempt to build their own UDP-based protocols, the author ended up implementing many of the safeguards found in the TCP protocol, however (like keep-alive messages, and some rudimentary data-checking), so future versions of the MUSHI-Chatter network interface may rely on TCP instead of UDP. One advantage of the lossy UDP protocol was that it strictly enforces the decision to keep all simulation updates within the province of the server, and reduces the temptation to inappropriately offload certain types of update calculations to the handheld devices.

The visualization of *MUSHI-Lignancy* was implemented using the native C# System.Drawing graphics libraries, but to improve performance, any future implementations will use a graphics library like DirectX that allows for more control over the graphics pipeline (and, concomitantly, the opportunity to speed up rendering).

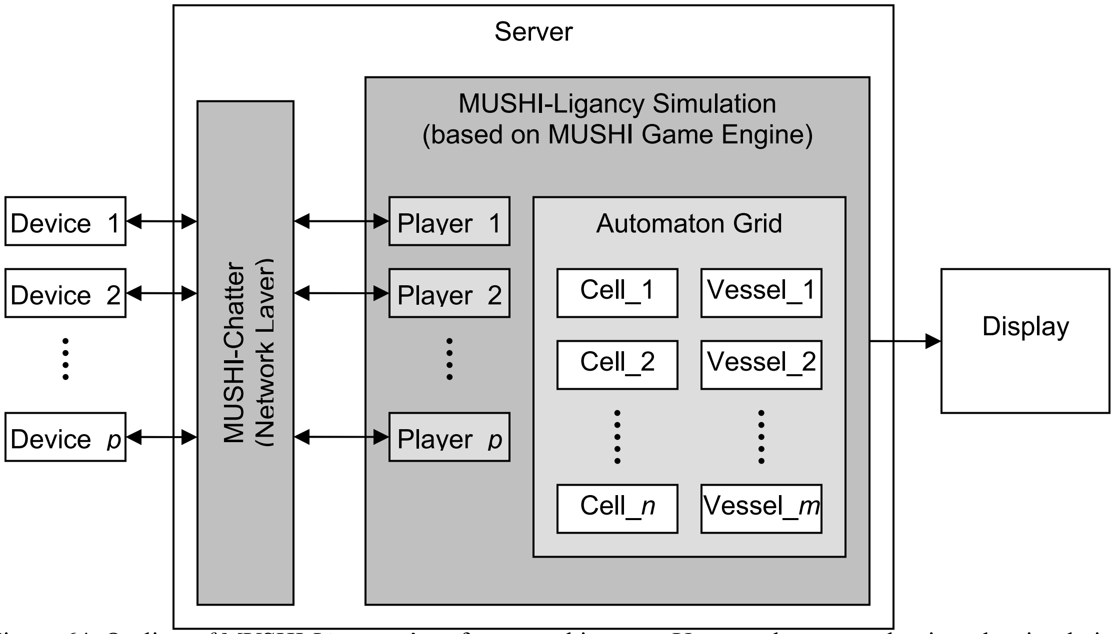


Figure 64. Outline *of MUSHI-Lignancy*'s software architecture. Up to *p* players can log into the simulation using wireless-capable devices. The simulation hosts *n* cell and *m* blood vessel segment automata, which update themselves on a schedule dictated by the game engine. The game engine obtains information about the automata (e.g., their current health) to visualize them for the display (in these experiments, a 48" plasma screen).

5.1.2 User Roles in MUSHI-Lignancy

The MUSHI platform was designed to allow individual users to wirelessly log into the simulation via handheld devices (see Figure 65). In this implementation, the devices user are Hewlitt-Packard iPaq h4100 handheld devices that run Windows Mobile 5.0, but these devices were intended to be used as a proxy for whatever devices a visitor may have with them when they attend the museum[23]. All user interfaces for the handheld devices were designed and built by the author using Microsoft's proprietary C# programming language and Windows Mobile libraries. Previous versions of MUSHI had been implemented with more of an eye towards eventual cross-platform use, using C++ and open-platform graphics libraries (like OpenGL for compact devices), but the performance that could be obtained by using Microsoft-native libraries and language made the switch worthwhile. (We also gained lower-level access to the IR port and the display). If the MUSHI system will ever move away from the experimental realm towards real-world deployment, large portions of the code should be portable, as it was designed to encapsulate and "hide" the native libraries as much as possible.

[23] A full implementation of *MUSHI-Lignancy* would need to accommodate a great many different devices, their native operating systems and versions, and perhaps even different communications protocols like Bluetooth. Because the research questions were more concerned with design than implementation, iPaq h4100s were supplied for visitors to use.

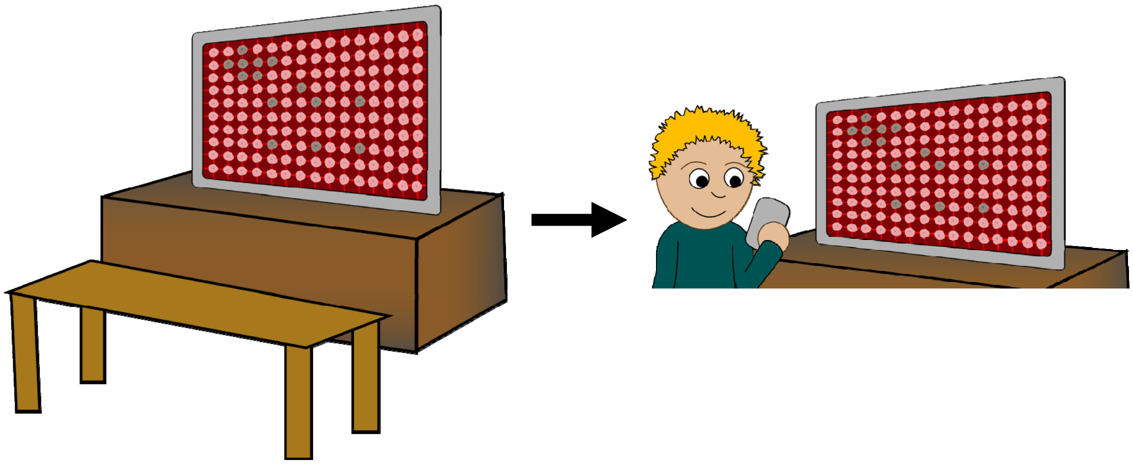


Figure 65. Illustration of the MUSHI-Ligancy setup. The simulation is displayed on a 48” plasma screen. A user logs into the ongoing simulation using a wireless-capable device. In this implementation, Hewlett-Packard iPaq h4100s were made use of for this purpose.

When a user logs into *MUSHI-Lignancy*, he or she is presented with a screen that gives them the option of selecting a role to play within the context of the *MUSHI-Lignancy* simulation[24]. The three roles available to the user are Surgery, Radiation, and Proton Beam. Each of these is a type of treatment used for cancer in real life. There was originally a fourth role, Chemotherapy, but initial testing showed that, from a “gameplay” perspective, the Chemotherapy role was, to put it frankly, too boring for visitors. It had the added complication of requiring much less frequent interaction, when compared against the other roles, and so it was eliminated from the lineup.

5.1.2.1 Surgery Role

A player who assumes the Surgery role is tasked with “excising” cancerous cells from the simulated patient, much as a real surgeon would use a scalpel to cut away tumors from a patient. After logging into the *MUSHI-Lignancy* simulation, a small color-coded rectangle, labeled with the user’s name, appears on the large plasma screen that depicts the gestalt of the simulation (see Figure 66). Players are able to distinguish one another’s representations on the shared screen using color-coding, a fairly standard approach in collaborative entertainment software (Bricker, Baker, Fujioka, & Tanimoto, 1998). The color-coded rectangle is an analogue to the “incision” that a surgeon can make into a patient, with the exception that this “incision” can be relocated by pressing the directional control pad buttons on the handheld device. When the surgery player makes “cuts” in the incision area (the two different user interfaces for this will be described shortly), the cells and blood vessel segments underneath the “cuts” take damage (the

---

[24] As Chapter 6 will describe, for some experimental sessions the ability to select a role was restricted to just one: the surgeon.

`PersistentDamage` variable gets incremented, see Table 4). This can decrease their "health," as represented on-screen by their diameters, or kill them outright.

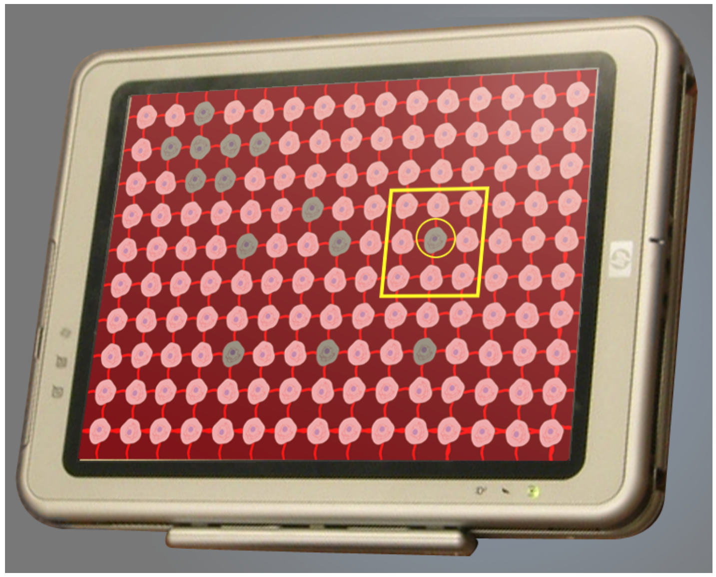


Figure 66. Depiction of the shared simulation screen after a player taking the Surgery role has logged in. (NOTE: this shows the simulation running on a tablet PC, not on the 48" plasma screen used in the experiments). The yellow rectangle indicates the extents of the Surgery player's ability to affect the simulation, an analogue to the "incision" a real surgeon makes in a patient. The yellow circle indicates the region that the player has chosen to excise.

Two distinct user interfaces were created for the handheld for the Surgery role, to support the controlled, lab-based experiment described in Section 6.6. One was designed to require a lesser degree of "attention," defined as the amount of visual attention and hands-eye coordination a user would need to give the user interface in order to operate it (see Figure 67). It was designed to operate much like a remote control, requiring no visual attention from the user. After logging in, the user is presented with a static screen that only depicts instructions for how to use the interface. Input is provided via the handheld device's hardware buttons. Input is limited to providing directions (up, down, left, and right) and a "cut" command. The "cut" command affects all cells within the scope of the "incision" rectangle, radiating out in lessening degrees of damage from the center.

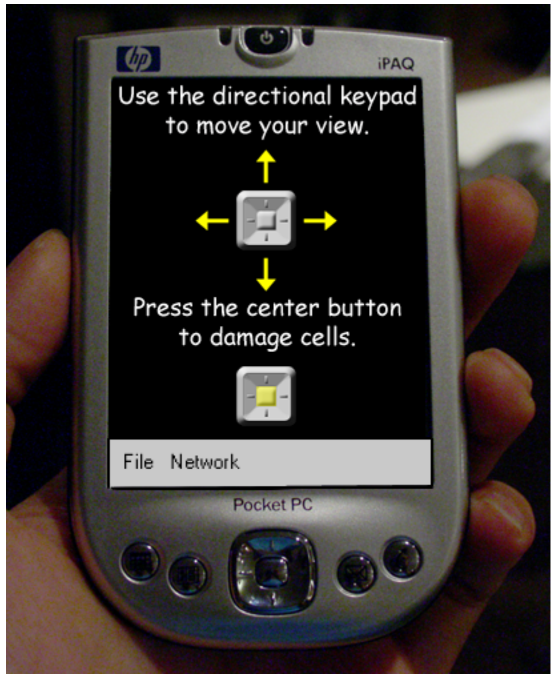


Figure 67. Photograph depicting the "Simple" version of the Surgery role's user interface. The image on the screen is only present to provide directions; it is static and does not change. The player makes use of the directional keypad on the device to provide all input.

The second version of the Surgery interface was designed to require, comparatively, a great deal of attention, both in terms of what the visitor needs to attend to on the device's screen as well as the amount of hand-eye coordination required to provide input (see Figure 68). After viewing an instruction screen, the player will be presented with a magnified view of the "incision" region depicted on the plasma screen as a yellow rectangle (see Figure 66). The player can move the "incision" rectangle around the "patient" as with the "Simple" version of the user interface, via the hardware button directional pad. Rather than pressing the center button of the directional pad to make a "cut," as "Simple" users do, users of this version of the Surgeon interface use the device's stylus to draw the desired incision on the handheld screen. The player can choose to cut through cells using straight lines, which damages them, or can carefully draw circles around the target cells, which excises them completely. The player can also cut through blood vessels, which acts like a temporary cauterization, decreasing the blood flow to the region. Providing input in this manner obviously requires much more attention and control than the "Simple" version of the UI needs.

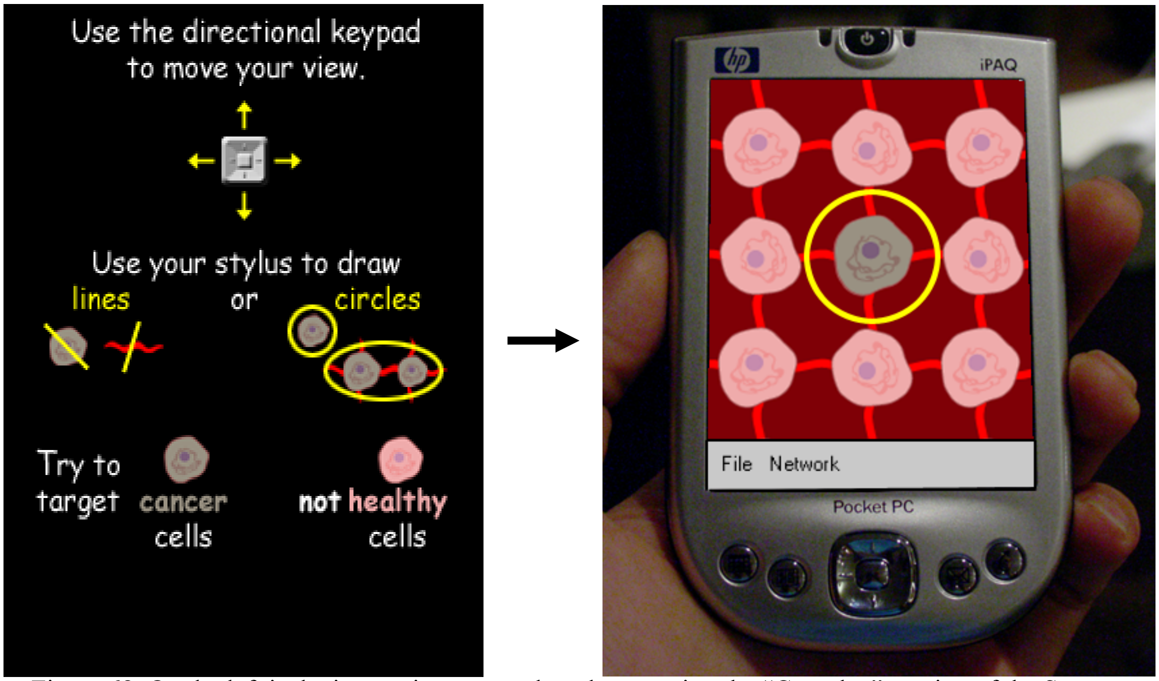


Figure 68. On the left is the instruction screen that players using the “Complex” version of the Surgery interface see when first logging into *MUSHI-Lignancy*. The image at right depicts the user interface seen while interacting with the simulation. Notice that the view is a “zoomed-in” representation of the “incision” region marked on the simulation screen by a yellow rectangle (see Figure 66). As the player moves the “incision,” the details on the handheld screen change to reflect the new incision region. The player has just circled the cancerous cell in the middle with his or her stylus, which acts as a “scalpel.”

5.1.2.2 Radiation Role

A player who assumes the Radiation role is tasked with administering radiation beams to the simulated patient. This amounts to choosing the vector and intensity of a beam of radiation to be applied to the simulated patient, and monitoring the amount of cumulative radiation to which the patient has been exposed. When a player assuming the “Radiation” role logs in, he or she does not see a rectangle on the shared screen like the Surgery players do, because Radiation players impact the entire patient. Rather, when they fire off a radiation beam, they see its path on the large shared screen (see Figure 69). Cells underneath the radiation beam will die as a result of their exposure, and the `RadiationExposure` value for those intersections on the grid (see Table 4) will be incremented so that the cells that re-grow in those spots will have a higher chance of becoming secondary cancer cells. Without going too much in-depth on radiation as a cancer treatment, it essentially “works” by disrupting the DNA of the cells in the path of the beam. Usually these disruptions are so catastrophic that the cells pretty much die in short order, unable to transcribe and manufacture the proteins needed to maintain their day-to-day functions, let alone cell division. Some cells, though, either by some fluke or by receiving a lighter dose of radiation by being on the margins of the beam, have their DNA damaged only slightly. Even though these may have been normal cells previously, the DNA damage may cause the same sorts of mutations (faster reproduction, delayed cell death) that are hallmarks of cancer. Thus, they are not the "children" of the original cancer cells, but are a distinct secondary cancer in their own right.

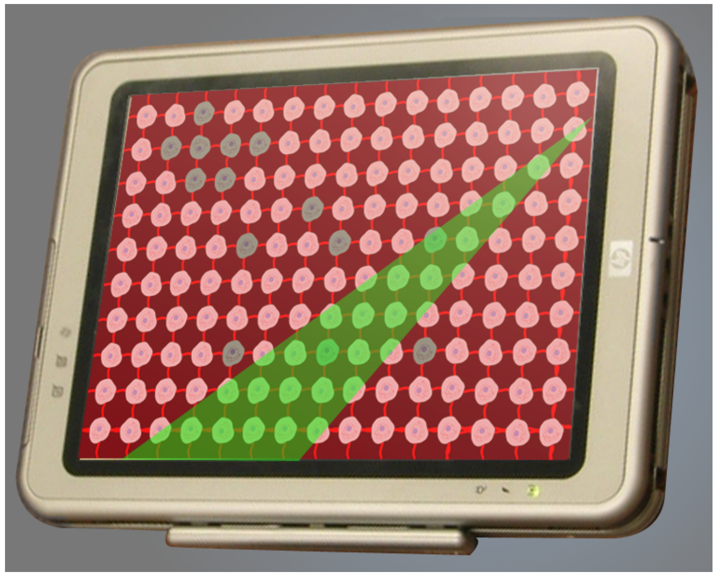


Figure 69. Depiction of the shared simulation screen after a player taking the Radiation role has logged in and fired a radiation beam. (NOTE: this shows the simulation running on a tablet PC, not on the 48" plasma screen used in the experiments). Notice that the beam, which entered the "body" on the top right, spreads in size as it passes through the patient, as real radiation beams do. The beam illustration will fade after a few moments, at about the same time the cells in its path begin dying off.

The user is presented with an instruction screen on his or her handheld device right after logging into *MUSHI-Lignancy*, which explains how to use the user interface and provides a rudimentary goal (not to damage too many normal cells). When the player switches to the main user interface screen, he or she is presented with several interactive elements (this user interface was not designed to be part of the user interface complexity experiment). At the top is a black field populated with small dots, each corresponding to a cell on the grid of the simulated patient (see Figure 70). On the borders of this grid are two drag-able icons, a green circle representing the entry point of the radiation beam, and a red square representing the exit point of the beam[25]. These icons are "pinned" to the borders of the grid, so the player is free to use the stylus to drag them around the margins of the grid, but cannot place one of the icons inside the grid (indicating an entry or exit point within the simulated body). Beneath the grid is a numbered slider, which the player can use to calibrate the strength of the radiation beam (in Grays, from 1 to 5). Beneath that is a button that, when tapped, will fire the radiation beam. Underneath that is a "progress bar" of sorts – it indicates the cumulative amount of radiation, in Grays, that the patient can be exposed to. The left end of the bar has a yellow smiley face icon, and the right end has a skull-and-crossbones icon (again, thanks to

[25] The author recognizes that the use of red and green violates one of the cardinal rules of user interface design (to wit, "Thou shalt make interfaces usable by the color-blind"). The colors were used as an oversight at first and were retained because formative testing confirmed their strong cultural significance (green = go, red = stop). The shapes were made distinct (a circle and a square) to support the small percentage of color blind users.

formative testing). The scale goes up to 80 Grays, because in real life, that is the maximum cap on the cumulative radiation that can be given to patients when being treated for cancer. (Patients to have varying tolerances, and organs respond differently, but generally, anything above 80 Grays will kill you fairly quickly). The cumulative radiation bar is updated via the server, as it maintains the final say (no pun intended) on the amount of cumulative radiation the patient has received.

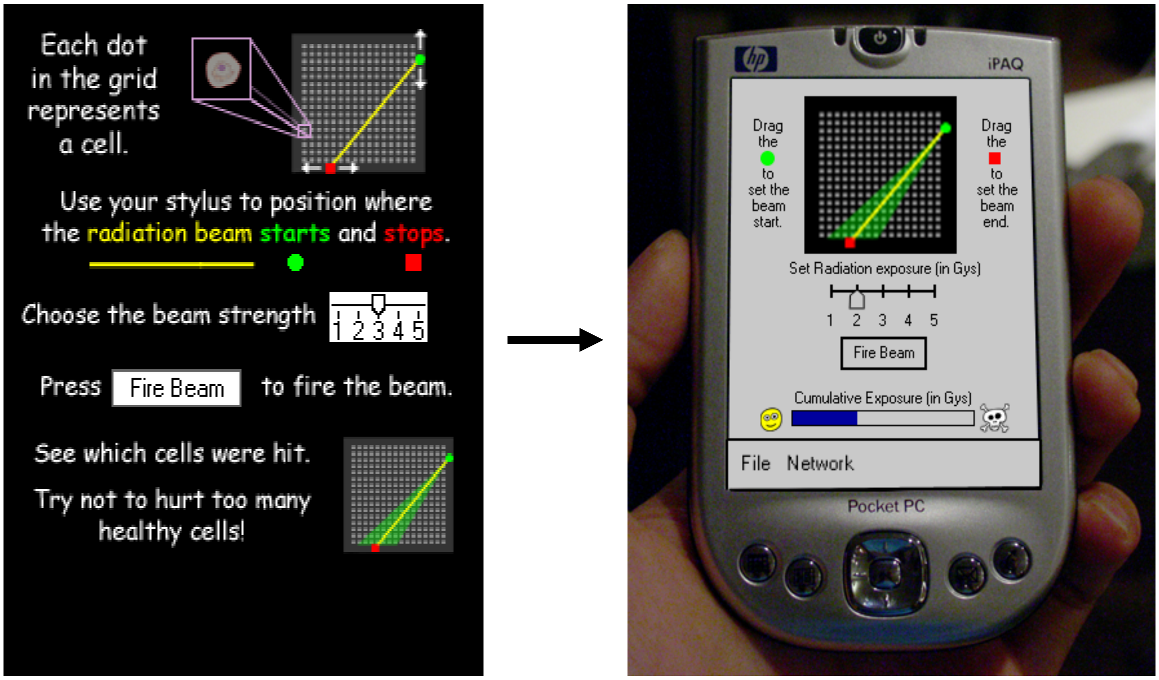


Figure 70. On the left is the instruction screen that players who choose the Radiation interface see when first logging into *MUSHI-Lignancy*. The image at right depicts the user interface seen while interacting with the simulation. The user is able to drag the green circle indicating the entry point of the beam and the red square of the exit point of the beam using his or her stylus. The grid underneath corresponds to the entirety of the simulated patient's cell grid. Underneath is a slider that controls the radiation strength, a "fire" button, and a bar that depicts the cumulative radiation exposure the patient has received.

5.1.2.3 Proton Beam Role

A player who assumes the Proton Beam role will find him or herself in a very similar situation as the Radiation role players. Proton Beam players will also choose the vector of a beam of protons to be applied to the simulated patient, and will also monitor the amount of cumulative radiation to which the patient has been exposed. When a player assuming the "Proton Beam" role logs in, he or she will similarly only be represented on the shared screen when a proton beam is fired, which will cause its path to appear on the large shared screen (see Figure 71). The differences between Radiation and Proton Beam players' roles stem from the differences between radiation and proton beam therapy.

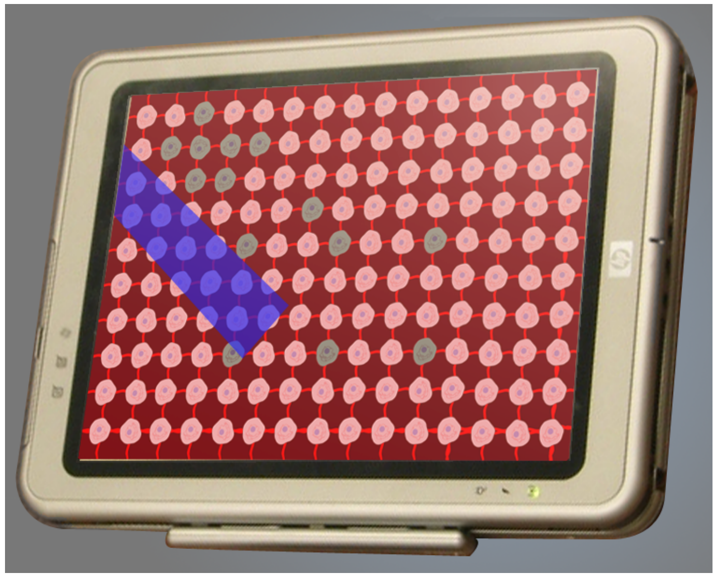
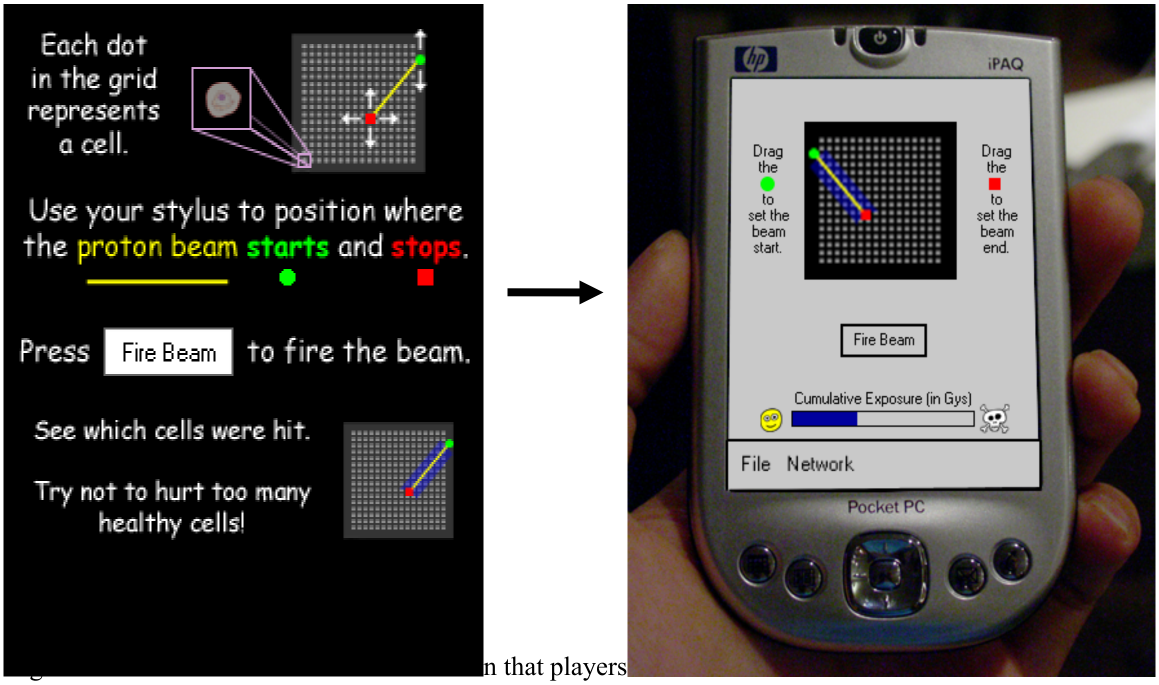


Figure 71. Depiction of the shared simulation screen after a player taking the Proton Beam role has logged in and fired a proton beam. (NOTE: this shows the simulation running on a tablet PC, not on the 48” plasma screen used in the experiments). Notice that the beam, which entered the “body” on the top left, does not spread in diameter, and only penetrates to a fixed depth, as real proton beams do. The beam illustration will fade after a few moments, at about the same time the cells in its path begin dying off.

In real life, proton beam treatment (where, quite literally, a beam of high-speed protons are shot at a patient’s tumors) is very similar to radiation therapy in application. They are also administered externally, and so an entry point for the beam must be chosen. The main differences, though, lie in the beam’s end point and spread. Radiation will continue to travel through the patient’s body until it exits, and thus can damage a great deal of non-cancerous tissue on its journey. Proton beams, however, can be tuned to only penetrate a desired distance into the patient’s body. This is accomplished by accelerating the protons to different speeds – just like a cue ball rocketing towards a cluster of pool balls, slower-moving protons just won’t penetrate as far as fast-moving protons. Another issue with radiation is that as it passes through a patient’s tissue, it expands to eventually form a sort of three-dimensional cone, a result of the beam colliding with the patient’s tissue and scattering. Proton beams, though, are less likely to “scatter” when they collide with tissue, and so they tend to maintain their tightly-focused diameter. The operating principle of proton beams is the same, however – as the protons crash into the patient’s cells, they disrupt the cells’ molecules, especially the DNA molecules. So proton beams also have the potential to cause secondary cancers – their compactness just exposes fewer cells to the risk.



n that players first logging into *MUSHI-Lignancy*. The image at right depicts the user interface seen while interacting with the simulation. The user is able to drag the green circle indicating the entry point of the beam and the red square controlling the penetration depth of the beam using his or her stylus. The grid underneath corresponds to the entirety of the simulated patient's cell grid. Underneath is a "fire" button and a bar that depicts the cumulative radiation exposure the patient has received.

The Proton Beam player, like the others, is presented with an instruction screen on his or her handheld device right after logging into *MUSHI-Lignancy*, which explains how to use the user interface and provides a rudimentary goal (not to damage too many normal cells). When the player switches to the main user interface screen, he or she is presented with several interactive elements similar to those seen by Radiation players. At the top is a black field populated with small dots, each corresponding to a cell on the grid of the simulated patient (see Figure 72). On the border of this grid is a drag-able green circle icon representing the entry point of the beam, which operates in the same manner as the Radiation player's green circle entry point icon. The Proton Beam player also has a drag-able red square icon, but unlike the Radiation player's version, it is not bound to the borders of the black cell grid. This red square icon represents the termination point of the Proton Beam, which the player can place anywhere within the black cell grid. Beneath the cell grid is a button that, when tapped, will fire the proton beam. Underneath that is the same cumulative radiation exposure bar that the Radiation user interface displays, ad which is updated by the server in the same manner.

5.2 Use Scenario of MUSHI-Lignancy

This section will provide a narrative that described how, under ideal circumstances, a group of visitors might use the *MUSHI-Lignancy* software.

5.2.1 Exploring the Context of Use

A group of three visitors, a parent and two children, wander into the Life Sciences–themed area where the MUSHI-Lignancy *exhibit is installed. They notice the bright colors and "video-game"-like appearance of the simulation, which is displayed on a large plasma screen. Unfortunately, a pair of other visitors are already using the simulation, so the family group watches the pair's interactions with the simulation on the large screen for a little while. After a bit, they wander over to some of the nearby exhibits, perhaps visiting one that uses a microscope and a slide of living cells to demonstrate cellular division. Another exhibit might have a glass case filled with late 19th century wax models used for medical education, which illustrate a few dramatic medical treatments. One in particular is a series of wax human heads, showing a large protuberant tumor on a patient's cheek, and several stages in the surgical operation used to remove the tumor, with the final head once again looking normal, aside from a black tracery of stitches. A third exhibit might have an old piece of medical equipment used to administer radiation, and a fourth might present different types of tools used for surgery.*

5.2.2 Initial Exposure

Before too long, the pair of visitors originally using MUSHI-Lignancy *wander off, and the family of three, noticing the vacancy, takes a seat on the low bench in front of the display. They pick up the handheld devices left on the table in front of the display, and follow the instructions on-screen to log into the simulation. When faced with the role-selection screen, they talk a bit about who will do what. The younger female child opts to adopt the Surgery role, and the older male child excitedly chooses Radiation, perhaps attracted by the name. The parent, a mother, notices that only one role is not represented, and so she selects the Proton Beam Role.*

The players spend a moment looking at their respective instruction screens. The younger child asks her mother fro clarification on the instructions, and shows her mother the handheld. The mother reads her daughter's handheld's screen, and summarizes the instructions in a form that the little girl will better understand, given their prior joint experience with a relevant exhibit:

"Well, this big screen here is a person, a patient, who has cancer. Remember, like that wax head model? There are some cells growing faster and bigger than they should, and you need to cut them out, just like in the model. Those are the darker ones. The pink ones are ok, don't hurt them."

The older boy, meanwhile, has already exited the instruction screen, and after a moment of toying with the user interface, realized that he could make big green swatch appear on the large shared display that would make many of the cells underneath shrivel and die. Encouraged by the magnitude of the effect, he experiments with cranking up the *exposure, and starts "blasting" radiation beams through the patient rather indiscriminately. Just as the mother and daughter are themselves exiting their instruction screens and trying to figure out how to engage with the simulation, they all get a message delivered on their handheld that declares that the patient has just died, owing to too much radiation exposure.*

Son: "What? What happened? Why'd he die?"

Mom, looking at the pop-up message on her handheld and then at the landscape of shriveled cells on the large shared screen, with one final green radiation beam fading to transparent, "He got too much radiation, I guess. Oh – were you playing just now? Did you do that?"

Son: "I dunno. I was just making the green stuff go everywhere."

Mom: "Well, the green stuff is radiation, I guess, so I think you just gave him too much."

Daughter: "But I didn't even get to try!"

5.2.3 Second Round

Just then, each handheld and the shared screen display a message that a new patient is about to appear. The messages disappear, and the family can see that they're once again looking at a patient with living cells. After a few moments, they notice that the grayish-green cells are starting to cluster in the lower-left region of the grid, and the blood vessels in the area are beginning to swell. At the same time, the other blood vessels are getting a bit smaller, as are the cells they supply with blood.

Daughter: "There's the tumor!" She quickly navigates her "incision" rectangle to be in the center of the gray-green cells, looking at the large shared display as she does the positioning. Once in place, she concentrates on the display on her handheld, and starts experimenting with making circles and slashes with her stylus.

In the meantime, the mother has been conversing with the son. Immediately after the new "patient" appeared, he started dragging his entry and exit icons around, preparing for another all-out assault.

Mother: "Wait, wait – this time don't go so fast."

Son: "OK." He fires a radiation beam straight down through the patient, laying waste to a large swath of normal cells."

Mother: "Wait, wait, wait – don't hurt the pink ones. They're good. The other ones are cancer."

Son: "Oh, ok. I didn't realize that."

The mother finally turns her attention to her own device, and starts trying to figure out how to use it.

Mother: "Huh, what's going on here…"

The son notices his mom's difficulty, and stops his targeting to help. "You drag the little things around, and then hit the button."

Mother: "The things?"

Son: "Yeah ,here." He holds his own device out so that his mother can see, and he drags the entry point icon to the left. "See? Now the beam's gonna come from the left."

Mother, looking back and forth between the devices, drags her own entry point icon. "Ohh, OK." Both mother and son turn back to their respective devices, and begin targeting. Both fire their beams, which intersect very near where the daughter had her "incision" rectangle.

Daughter, startled: "Wait, what just happened? Everything just disappeared!" She looks up and sees the green and blue traces of her brother's and mother's beams, respectively, on the large shared display. "What's that stuff?"

Son: "We just shot those cancer cells up! Awesome! Look at that! They're all dead!"

Mother, to the daughter: "That's us, honey. We just used our beams."

Son: "Wait, there's still one left!" Points at a lone cancer cell on the margin of the former tumor site, which has largely been wiped out, on the large shared display. Directs sister, "Get it!"

The daughter looks on the large shared screen where her brother is pointing, down at her device, which is showing an array of dead cells, and back at the large shared screen. She uses the buttons on her device to navigate to the cancer cell, while looking at the large shared screen. Once she arrives, she looks back at her device, and makes a decisive circle with her stylus. Continuing to look at her device, she waits for the cell to die. Triumphantly, "Got it!"

Son: "Wait wait wait – what's going on? Why's it coming back?" He has noticed that near the edges of where he and his mother targeted their beams, cells are growing back – cancerous cells. "We killed it! I don't get it!"

Mother: "Maybe it's the radiation? Remember when grandpa had cancer, and got radiation, and he was ok for a while, but then he got cancer again a year later? The doctor said that it was the radiation that caused it the second time."

Son: "But if radiation makes people get cancer, why do they use it to kill cancer? I don't get it."

Mother: "Well, if grandpa hadn't had the radiation, he would have died pretty quick – they said a few months. But he got radiation, and so we got to spend a whole extra year with him. He got to see your cousin get born."

5.2.4 Commentary on Use Scenario

This section will relate the software use scenario to some of the themes explored in the chapter on the museum context (Chapter 2), existing computer-based exhibits in museums (Chapter 3), and the design of computer-supported collaboration (Chapter 4).

The episode illustrated above demonstrates several design features that promote effective group learning. The "mother" character assumes a mediator role, as described in Section 2.2.3 Identity and Roles in Museum Settings, and attempts to link what they are doing in the simulation to other exhibits they had seen that day, and to personal experiences shared by the family. The presence of more "authentic" exhibits relating to cancer allows the mother to relate the "inauthentic" on-screen representations back to those more tangible exhibits (see Section 2.1.1 Authenticity and the Object-Based Epistemology on authenticity in museums). Although the software is virtual by definition, it attempts to capture the processes underlying real cancer growth (one of the recommendations from Section 2.1.1.2 Implications for Computer-Based Exhibits), which the mother recognizes and points out when the radiation beams cause new cancer to form during the second round of play.

All three family members make use of the large shared display as an anchor for their discussions (see Section 3.2.2 Multi-User Kiosks: Large Display Kiosks in Museums, and Sections 4.2.1 and 4.3.1 on Single-Display Groupware in CSCW and CSCL, respectively). By making the actions, outcomes, and attentional foci of the players visible on the shared display, the visitors are able to better coordinate their actions (see Section 2.2.2 Group Behaviors at Science Museums, Section 2.2.3 Identity and Roles in Museum Settings, and Section 4.3.1 SDG in CSCL). Because the software is an open-ended simulation, the visitors can try it until they feel they have mastered it, engaging in Active Prolonged Engagement (see Section 2.1.2 Interpretation and Authority). The visitors each have equal access to the simulation via their handheld interfaces, promoting equal engagement (see Section 2.2.2 Group Behaviors at Science Museums, Section 3.2 Kiosks in Museums, and Section 4.3 on Computer Supported Collaborative Learning for further discussions of equal access). The different interfaces support the visitors in assuming different occupational roles, which together form a positively interdependent, jigsawed activity structure that encourages more on-task discussion (see Section 2.2.3 Identity and Roles in Museum Settings, Section 3.3 which touches on the use of jigsawing in museums, and Section 4.3 on jigsawing in CSCL).

CHAPTER 6

Experimental Design and Methods

This chapter opens by situating and stating the research questions driving this study (Section 6.1), followed by definitions of the variables, and a description of the measures used to assess the dependent variables (Section 6.2). The general research paradigm employed by this research will be described (Section 6.3), and the site of the research will be described in Section 6.4. Section 6.5 outlines the scope and purpose of the formative research performed in the first phase of the study. The latter portion of the chapter, Section 6.6, covers the experimental design for the lab-based portion of this study.

6.1 Research Questions

A brief review of the problem space is needed to set the stage before spelling out the research questions, in case the reader has come directly to this chapter.

6.1.1 Recap: Situating the Research

This research is predicated on the goal of supporting museum visitors as they use their own personal devices as Opportunistic UIs (O-UIs) to join a collaborative learning activity hosted on at least one large, shared display (see Section 3.2.2 for a summary of the use of large, shared displays in museums). By using O-UIs in conjunction with a shared display, the form-factor resembles that of a Multi-Machine User Interface (MMUI) paradigm. Altogether too little is known about MMUI design for museums, but Section 3.4 provides a summary of the sparse accounts of MMUIs in use in museums. One flagged area of concern is how visitors may (or may not) divide their attention between the different devices in a MMUI-based exhibit. Although the shared displays employed as a component of MMUIs can support simultaneous use by groups of learners in museums (see Section 3.2.2), theoretically improving an exhibit's prospects of supporting collaborative learning; on the other hand, there is evidence that the use of mobile devices in museums might interfere with group learning processes (see Section 3.3).

The concern is that if visitors use O-UIs to join in a shared collaborative activity visualized on shared public display(s), the O-UIs could have the potential to draw so much of the visitors' visual attention that the public display(s) – and perhaps their companions – become superfluous. Shared public display(s) can support collaborative learning (e.g., by providing grounding for conversation and supporting shared task monitoring needed for joint attention management), but only if visitors attend to them.

Likewise, O-UIs might allow exhibits to scale to accommodate groups of visitors, but if the visitors don't speak to one another, such an exhibit could only nominally be considered supportive of collaborative learning. There is some tentative evidence that "simplifying" the user interface may reduce the monopoly O-UIs can have on visitor attention (see Section 3.3.1.2 Output from Single-User Handheld Device Activities). Unfortunately, it can be hard to design rich opportunities for learning with only "simple" user interfaces – "complex" user interfaces allows for more information to be transmitted to users, and more nuanced input to be obtained. These concerns are outlined in Section 3.5.2 Open Questions for MMUIs in Museums.

The research questions presented here take a first slice, looking to see if the individual and collaborative activities of visitors using MMUIs are impacted when the *complexity* of the private user interfaces (the Opportunistic UIs) is manipulated[26].

6.1.2 Questions

Phase I: **Formative Research – no research questions**

Phase II: Does increasing the **user interface complexity** of Opportunistic-UIs in a MMUI-centric museum exhibit:

1. Produce the heads-down phenomenon?
2. Affect the potential for the activity to support collaborative learning?

   As measured by:

A. The **individual**:
   a. The **individual division of attention** paid to private and public devices
   b. The **individual engagement** with the activity
   c. The **individual task performance**

B. The **group**:
   a. The **group division of attention** paid to private and public devices
   b. The **group's engagement** with the activity
   c. The **group's task performance**

6.2 Variables and Measures

The research questions of Section 6.1.1 were stated as succinctly as possible, which means that there is room for multiple interpretations of the exact definitions of the independent and dependent

[26] For the purposes of this research, the mobile devices used as O-UIs were supplied, to better concentrate on the problem space at hand (and not get caught up in writing code that supports twenty-odd different mobile devices).

variables. Section 6.2.1 will describe the independent variable named in the research questions: Opportunistic User Interface complexity. Section 6.2.2 will describe each of the dependent variables.

6.2.1 Independent Variable: Opportunistic User Interface Complexity

None of the prior work on mobile devices in museums was explicitly designed to examine how the heads-down phenomenon might be affected by UI design, but a common implication was that "simpler" graphical user interfaces were to be preferred over "complicated" ones. One study of the use of hypermedia informational kiosks in museums tried to come up with objective measures of user interface complexity, but they chose features that are perhaps more relevant to browsing information than engaging in computer-mediated activities, like shallowness, downward compactness, and navigability (Yamada et al., 1995). For the purpose of this research, there is a need for a measure that addresses the demands on visual attention that a user interface will impose – the heads-down effect's primary symptom is a handheld device's monopoly on visual attention.

Many users, when asked, can easily tell the difference between a visually "simple" and a "complicated" user interface, but much like the debate over "erotica" versus "pornography," making distinctions in user interface complexity seems to be a highly personal affair. From time to time HCI researchers have tried to come up with methods of objectively quantifying user interface complexity, but no particular method has risen up and captured the allegiance of a majority of the HCI community, perhaps because no single definition would work for all tasks and all users. This section will discuss how the literature was reviewed for relevant information, and then present a few of the more prominent methods for defining complexity, before presenting the approach that will be used in this research.

In search of a working definition of UI complexity, HCI literature (primarily the ACM SIGCHI archives and the HCI journal archives) was reviewed to look for how complexity is defined in the context of designing graphical user interfaces (especially mobile GUIs). Literature related to attention, working memory, and visuospatial cognition, especially as applied to educational software, was also reviewed. Quite a bit of early CAI/CBI (Computer-Aided Instruction/Computer-Based Instruction) literature was concerned with the cognitive impacts of different types of user interface components (particular topics of interest included studying the impact of the variability of the UI, and the value of imagery and animations). The sources that yielded the most information were the journals *Cognition & Instruction*, *Learning & Instruction*, *Journal of Educational Psychology*, *Educational Technology Research and Development*, and *Journal of Science Education and Technology*. The journals *International Journal of Computer-Supported Collaborative Learning* and *Journal of the Learning Sciences*, and their associated conferences, *Computer-Supported Collaborative Learning* and *International Conference of the Learning Sciences*, did not yield very much information on user interface complexity – perhaps because authors in these venues tend towards "scruffy," as opposed to "neat," research methodologies (Crevier, 1993).

One of the more procedural methods that can be used to compare the UI complexity of competing user interfaces is the Goals-Operators-Methods-Selection Rules (GOMS) approach (John & Kieras, 1996). Because GOMS analyses attempt to represent the actual symbolic goings-on within a user's brain as they operate a user interface, they are as "true" a definition of user interface complexity as one is likely to achieve, down to the milliseconds required to perform UI tasks. GOMS was designed for use in high-stakes design verification (e.g., military missile launch systems), with the underlying assumption that the users would be experts (e.g., highly-trained servicemen). User interface experts, especially military-trained experts, behave very differently than novices, however. Since it is a safe assumption that computer-based museum exhibit users will be first-time users, a GOMS analysis is not appropriate – there is likely to be too much variation in first-time users' approaches to a new interface to be able to assign timing data and construct selection rules that would accurately describe user behaviors.

This research is concerned with varying the degree to which O-UI interfaces would visually distract users from other elements in the context, so definitions of UI complexity that spoke to how much of a first-time user's attention resources would be occupied while using the UI were sought. Some HCI researchers have attempted to automatically compute UI complexity by looking at factors like element size (smaller objects take longer to register visually), local density of elements (it takes longer to process dense arrangements), alignment (aligned elements are easier to scan), and grouping (clustering elements into functional groups reduces eye-travel time) (Miyoshi & Murata, 2001; Parush et al., 1998). These particular measures are all based on research into how humans cognitively process visual stimuli and how attention is regulated. People do not have infinite capacities for attending to visual stimuli; only a certain amount can be stored in what is dubbed "visual working memory" by cognitive psychologists (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; E. E. Smith & Jonides, 1997). The basic notion behind working memory is that humans have a finite pool of short-term memory that can be devoted to processing stimuli, and that different types of stimuli are in turn likely to occupy different amounts of "space" in working memory. A term for describing the amount of working memory "space" taken up by a stimulus is the "cognitive load" of that stimulus (van Merriënboer & Ayres, 2005). In the absence of any relational properties, the ability for humans to store independent visual stimuli in working memory degrades as the numbers of stimuli are increased (Luck & Vogel, 1997), which is probably what underlies the recommendation for lower element density in user interfaces. The exact capacity of visual working memory depends greatly on the specific visual expression of the stimuli, however – humans have the capacity to attend to larger numbers of visual features when the features are presented in a conjunctive manner (Luck & Vogel, 1997). This may underlie the reason why alignment and grouping have been seen to improve user interface usage. As far as element size goes, human attention is affected by bottom-up processes that filter for the salience of stimuli (Knudsen, 2007), which may be why humans are more prone to noticing larger and closer objects. The visuospatial workings of the mind are still very much under exploration, however, and as with all experiments dealing with cognition, it is best to resist the impulse to draw strong conclusions from the findings without taking careful consideration of the exact conditions of the studies in question. For example, "salience" can have many definitions, depending on the context.

Table 5. Thus far the terminology used to frame this research has been rather casual, in that it is derived from observational studies of visitors in museums. Educational psychology has the power to define these "observational" concepts in more precise terms based on how the mind processes information. This table shows the educational psychological interpretation of several "observational" concepts.

"Observational" Concept	Equivalent Concept in Educational Psychology
"Complexity" of user interface	Cognitive load induced by user interface
Extra work / overhead in completing a task	Extraneous cognitive load
The "heart" or "meat" of a task	Germane cognitive load
Suffering from distraction	Split Attention Affect

A better guide for understanding how O-UI complexity might affect visitors may come from the educational psychology literature, since many experiments in this field are constructed to study cognition in the context of learning exercises. Educational psychologists have found that when the working memory of learners is highly occupied, i.e., when it carries a "high cognitive load," learning can in fact be impeded. They have also found that when a learner must divide his or her visual attention between stimuli located in different spatial locations, cognitive load is increased and learning is impeded, a phenomenon known as the "Split Attention Effect" (Sweller et al., 1990). Thus far the terms "simple user interface," and "complex user interface" have been used rather loosely, but by using more precise concepts from educational psychology these ideas can be recast as "low cognitive load user interface," and "high cognitive load user interface" (see Table 5). So what this study aims to do, then, is to bracket the design space for Opportunistic User Interfaces, contrasting, on one end, a design that will impart the *least* amount of cognitive load on a user's visual working memory, against a design on the other end of the spectrum that imparts a relatively large cognitive load on the user's visual working memory. The choice to employ user interfaces with as large difference in cognitive load as is practical was made so as to exaggerate whatever effects O-UI "complexity" may have on MMUI usage. (Given that this is the first systematic investigation of the topic, this seemed prudent – if an impact is found, further research can be done to refine these broad notions of "simplicity" and "complexity" into component parts).

The "cognitive load" concept from educational psychology permits the use of existing empirical work to select user interface elements that are more or less likely to induce cognitive load. On one end of the spectrum, there are devices that require virtually no visual attention at all: the "remote control" hardware button-based paradigm (see Figure 73). Individuals tend to engage in "orienting" behaviors, wherein they direct a sensory system (e.g., eyes) towards the focal point of their attention (Knudsen, 2007). If a device does not compete for visual attention, all other things being equal, it is less likely to impose a high cognitive load, owing to the lack of a split attention effect.

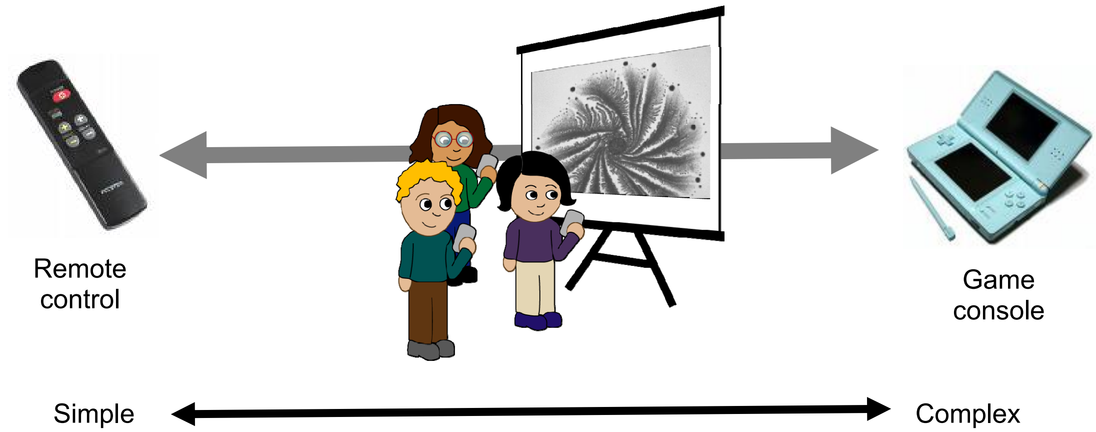


Figure 73. Illustration of the spectrum of user interface “complexity” used for the O-UIs of the MMUI system in this research, where complexity is defined by the amount of visual attention required to operate the device. Remote controls are typically operated without any visual attention whatsoever, and so they exemplify the “simple” end of the spectrum. A game console like the Nintendo DS, which has two display screens, one of them touch-sensitive and operated with a stylus, is an example of a handheld device that requires a great deal of visual attention to operate

To represent the “complex” end of the spectrum, then, what is needed are user interface elements that innately induce the greatest amount of cognitive load. Although the characteristics mentioned earlier (small element size, large local density, no alignment, and no grouping) would undoubtedly have an effect, by designing an *arbitrarily* complex user interface, this study could be cast as a straw man comparison. There is a distinction between cognitive load that results from arbitrarily distracting or irrelevant elements, called *extraneous* cognitive load, and cognitive load that is a natural outgrowth of the innate characteristics of the task, known as *germane* cognitive load (van Merriënboer & Ayres, 2005). Thus, to represent the “complex” end of the spectrum, a decision was made to design a user interface that would employ dynamic graphics in a *germane* manner. Dynamic imagery is known to impart a significant cognitive load on viewers (Bodemer, Ploetzner, Feuerlein, & Spada, 2004). The dynamic graphics in the “Complex” O-UI used in this study are made *germane* by explicitly tying them to the means of providing input: users must draw shapes and lines with a stylus, and these shapes and lines must be placed with respect to the currently-depicted imagery. Thus, it is ensured that visitors *must* attend to the dynamic graphics and will consequently have a higher visual cognitive load, but it is also the case that the higher cognitive load is still relevant to the task at hand (as opposed to being arbitrarily-induced). An existing analogue for this sort of interface can be found in the commercial Nintendo DS game system (see Figure 73). With these two contrasting models of complexity as models, the remote control and the Nintendo DS, the user interfaces used in this study were designed (see Section 5.1.2.1 Surgery Role), and subsequently refined in Phase I of the study (see Section 6.5 Phase I: Formative Research an Section 7.1 Phase I: Formative Study Results).

Table 6. Conditions used for the *O-UI Complexity* independent variable.

O-UI Complexity Condition	Description
"Simple" O-UI	Reduces the amount of cognitive load placed on working memory by removing graphical display from O-UI and allowing input to be given via hardware buttons, which decreases risk of split attention effect
"Complex" O-UI	Maximizes the amount of germane cognitive load placed on the working memory by adding a dynamic display to the O-UI and requiring that the stylus-provided input is tightly coupled to the dynamic imagery on the O-UI display. Presence of secondary display runs the risk of inducing the split attention effect

6.2.2 Dependent Variables

The dependent variables were chosen on the basis of their ability to serve as bellwethers for collaborative learning processes (see Table 7 for an overview). An initial question the reader may have is why, if *learning* is of interest, none of the measures presented in Table 7 attempt to measure learning directly. For example, why not use pre- and post-tests to try to measure knowledge gains?

There are several reasons for the decision not to use pre- and post-tests. One explanation is rooted in the differences between cognitive and sociocultural theoretical perspectives on learning. The use of pre- and post-tests arises from a cognitivist view that learning is evidenced by changes to a learner's inner mental organization, like mental models (Johnson-Laird, 1983). A core idea behind the sociocultural perspective on learning, however, is that knowledge must be expressed externally first, via interaction patterns like action or speech, before it can become internalized to help learners change their mental models (Sawyer, 2006). This externalization is usually thought of as taking place in face-to-face scenarios, but can also take the form of artifacts like bulletin board postings or written texts, with the interaction occurring when that artifact is interpreted or read. "Solo learning," as when a student reads a textbook, is not seen as being solo at all when it is viewed though a sociocognitive lens: there is a textbook author on the other end of the interpretive process, who has in turn been influenced by his or her interactions with many other socially-constructed understandings and artifacts. While individuals can and do maintain their own private, personal understandings of a scenario, these private understandings alter and are altered by the surrounding social context (Stahl, 2006b). Thus, if one takes a sociocultural stance towards understanding learning, markers of learning can be found in the dynamic *processes* of the scenario (like the steps a learner takes to complete a task, or conversational turns) that form the activity system (Greeno, 2006). The reader will notice that most of the dependent measures in Table 7 are, in fact, process-oriented.

Studying processes as a method of gauging learning in no way countermands the use of methods that try to assess mental model changes (like pre- and post-tests), so an attentive reader may still be wondering about the lack of pre- and post-tests in this research. This brings up the second explanation: practicality. First of all, longitudinal studies of museum visitors have shown that it can take months, or years for the changes in mental models resulting from museum experiences to occur (Crane, Nicholson, Chen, & Bitgood, 1994; Falk & Dierking, 2000). Secondly, by administering a pre-test, visitors are being primed to attend to certain features of an exhibit that they may not have otherwise taken notice of, predisposing them to learn things they ordinarily would not have – the "test effect" (Diamond, 1999; Peart, 1984). Third, by administering a "test," visitors may be more likely to behave as they would in the expectation-laden formal environment of a classroom, and not as they normally would in the informal learning environment of a museum – some would argue the whole point of an informal learning environment is to provide learners with "no limits, tests, or lectures," such that they are free to engage in idiosyncratic knowledge-building rather than try to adhere to some external standard of "correctness" (Crane et al., 1994). Finally, it is simply very difficult to recruit a representative sample of museum visitors to take part in an activity that will last more than 15 minutes, and the addition of pre-and post-tests can easily eat up 10 minutes or more, leaving very little time for them to experience the exhibit under study. For these reasons, many museum researchers have turned to process-oriented measures, like conversational analysis, to assess the suitability of an educational intervention (Allen, 2002, 2004; Borun, 2002; Borun et al., 1996; K. Crowley & Callanan, 1998; Kevin Crowley & Eberbach, 2005; Diamond, 1986; McLean, 1999; Scott. G. Paris & Hapgood, 2002; Serrell, 1997). Much can still be learned about whether an activity is likely to be, in colloquial terms, a "nonstarter," just by looking at process-oriented measures, especially activities involving multiple participants. Assessing learning via cognitivist, direct-measurement approaches like pre- and post-tests is just one method of many of coming to understand how an activity may or may not support collaborative learning. By way of contrast, if *only* a cognitivist perspective is used, other facets important to measuring and understanding collaborative learning will be missed. While knowledge tests can be useful for performing summative evaluations, they often miss nuances that can help theorists better understand how people collaborate, and can help designers improve their design efforts.

The dependent measures used in this study, although not measuring learning directly, measure aspects of the learning scenario that indisputably affect learning, like attention. The review of prior work in museums exposed the fact that visitor attention is a valuable resource, and when misdirected, it can work against the ability for users to learn. So, the manner in which the visitors divided their attention among the different elements present in the different experimental conditions was certainly a concern, and the specific measures used to gauge *Individual Division of Attention* will be described in greater depth in Section 6.2.2.1 Individual Division of Attention. Because the MMUI format has not been previously studied in museums to any great extent, to verify the assumption that certain styles of attention division impact measures for *Individual Engagement* and *Individual Task Performance*, data on those aspects were also collected (See Sections 6.2.2.2 Individual Engagement and 6.2.2.3 Individual Task Performance, respectively).

This research is not just concerned with the effects of the independent variables on individuals, however. The main purpose of this research is to study design strategies for a form factor intended to support collaborative, small-group learning, so group-level phenomena also need to be assessed. Using the individual as the unit of analysis is a mainstay of traditional education, but to truly understand a collaborative learning context, the level of granularity needs to be that of the group, and measures that take into account group-level processes and structures need to be used (Greeno, 2006). Some educational researchers have begun using the concept of performance "equity" as a measure of the success or failure of the collaborative aspect of an activity (Kapur & Kinzer, 2007), a concept that has also been labeled "mutuality" (Barron, 2000). Certainly, one wouldn't wish for a group activity that encouraged only one member of the group to pay attention, be engaged, or perform well – this is the paradigm employed by the single-user kiosks of Section 3.2.1 Single-User Kiosks in Museums that were deemed so unsatisfactory for group learning. Sections 6.2.2.4 Group Division of Attention, 6.2.2.5 Group Engagement, and 6.2.2.6 Group Task Performance describe the secondary analyses of individual performance metrics used to determine if visitors are equally attentive, equally engaged, and performing at equal levels, respectively. There are also a few measures that can *only* be described at a group-level of granularity – one is the degree to which groups take advantage of the shared display as a grounding tool (see Section 6.2.2.4's measure of attention "synchronicity"). Another is how successful, overall, the collaboration was at addressing the overarching joint task, which will be addressed in Section 6.2.2.6 Group Task Performance.

Table 7. Descriptions of dependent variables used in this study. See Table 20 for a listing of the instruments used to capture data for these variables.

	Dependent Variables	Descriptions
Individual	Individual Division of Attention	**Proportion**: The time spent looking at the public display vs. private display vs. companions **Frequency**: The number of times gaze switches between public display vs. private display vs. companions **Duration**: The length of time of the gazes at each display and companions **Awareness**: self-reported measure
	Individual Engagement	**Conversation**: number of utterances made by the individual **Participation**: number of input events initiated by the user
	Individual Task Performance	**Individual score**: number of individual's targets that were appropriate targets, less the number that were inappropriate, divided by total targets **Ownership**: degree to which user's actions do not replicate the actions of others
Group	Group Division of Attention	**Synchronicity** of group members in gazing at public display, measured in proportion of time spent in synchronized gazing **Attention Inequity**: dissimilarity of group members on attention measures of proportion, frequency, duration and awareness
	Group Engagement	**Participation Inequity**: dissimilarity of group members in number of input events initiated **Conversation Inequity**: dissimilarity of group members in number of utterances
	Group Task Performance	**Outcome** (i.e., success or failure) **Score Inequity**: dissimilarity of group members in individual scores **Ownership Inequity**: dissimilarity of group members in ownership

6.2.2.1 Individual Division of Attention

Attention, as it is used here, is primarily used to denote visual attention, as the two are usually synonymous for normally-gifted users. By way of example, in studies of collaborative user interfaces that tracked users' eye movements, users were found to move their gaze to the item they were attending to, even before mentioning it or gesturing towards it (Ou, Oh, Yang, & Fussell, 2005). So, to establish users' divisions of attention, the items they are gazing at from moment to moment can be tracked, and that data used to compute several measures that will describe the behavior.

Table 8. Formulas for computing the various *Individual Division of Attention* measures.

<table>
<tr><td rowspan="7">Individual Division of Attention</td><td rowspan="3">Proportion</td><td>O-UI</td><td rowspan="3">$\forall g : Target(g) == T, \frac{\sum_g Duration(g)}{t}$
where:
g = gaze
t = length of session, in seconds
T = given target (either O-UI, Shared Display, or Other Players)
$Target(g)$ = target of gaze g
$Duration(g)$ = duration of gaze g, in seconds</td></tr>
<tr><td>Shared Display</td></tr>
<tr><td>Other players</td></tr>
<tr><td rowspan="3">Duration</td><td>O-UI</td><td rowspan="3">$\forall g : Target(g) == T, \frac{\sum_g Duration(g)}{\sum_g 1}$
where:
g = gaze
T = given target (either O-UI, Shared Display, or Other Players)
$Target(g)$ = target of gaze g
$Duration(g)$ = duration of gaze g, in seconds</td></tr>
<tr><td>Shared Display</td></tr>
<tr><td>Other players</td></tr>
<tr><td>Frequency</td><td colspan="2">$\frac{g}{t}$
where:
g = # gazes
t = length of session, in minutes</td></tr>
</table>

The first measure that speaks to how individuals are dividing their attention is the *Proportion* of attention devoted to different gaze targets. After obtaining a moment-by-moment accounting of a participant's gaze targets, the proportional amount of time spent looking at the private interface, versus the public interface, versus other important elements of the learning context, like their companions, can easily be computed (see Table 8). If the proportion is very high for the O-UI, it is likely that the case is a candidate for being classified as an example of the heads-down phenomenon. Another measure that provides insight into a participant's attention is the *Frequency* with which he or she switches gaze between targets in the learning context (see Table 8). To the extent that one can assume that participants look at the public, Shared Display in order to ground themselves in a common context, a higher *Frequency* of gaze shifts may indicate that a user is more engaged in the collaborative aspect of the activity – as would a higher *Frequency* of gaze shifts to his or her companions. A related measure is that of the *Duration* of the participant's gaze at different targets in the environment, or, in other words, how long a user continues to look at a target once his or her gaze shifts to it (see Table 8). Attention *Duration* is used in conjunction with the attention *Proportion* and *Frequency* measures to more accurately diagnose heads-down problems: a canonical case would be when a participant has a high *Proportion* of attention devoted to his or her O-UI, and the attention to the O-UI is typically of long *Duration*, with a very low *Frequency* of shifting attention to wither the public display or his or her companions (see Table 9).

Table 9. Individual gaze behaviors emblematic of the heads-down effect.

Proportion of gazes	**High proportion** of gazes devoted to O-UI
Frequency of gaze shifts	**Low frequency** of gaze shifts
Duration of gaze	**Long duration** of gazes devoted to O-UI **Short duration** of gazes devoted to other targets

The attention measures discussed thus far have all been "objective," in the sense that the data is captured and analyzed from a perspective that does not take the user's own perspective into account. Much of the prior work on mobile device usage in museums that was responsible for the recognition of the heads-down phenomenon relied on the personal reports from the users themselves, captured from either questionnaire or interview data. To obtain a more subjective *Awareness* measure of attention division, variations on a question originally used by (Wagner et al., 2006) in their work: "I knew exactly what the others were doing" were employed (See questions 1-14 in Appendix B). In general, museum questionnaires yield more useful information when centered on visitors' attitudes, and not recall of specific information, and so the questions were predominately structured in this manner (Diamond, 1999).

6.2.2.2 Individual Engagement

Engagement is comprised of the "functional aspects of activity" that are so important to understanding the overall participation structures at play during a collaborative activity (Greeno, 2006). Of particular interest were measures that would indicate how much agency, or *Participation,* a user was bringing to the activity. To measure *Participation,* all of the input events initiated by users that would result in actions (e.g., a state change) in the shared activity were logged (see Table 10 for how these logs are interpreted). A high level of *Participation* will show whether or not a user was engaged in the activity. Because engagement with a learning activity is a necessary (although not sufficient) precursor to actually learning (Hudson & Bruckman, 2004; O'Donnell & Dansereau, 1992), a significantly lower level of *Participation* for one of the two experimental conditions would indicate that it is less able to help visitors learn.

Table 10.Formulas for calculating the *Participation* measures of *Individual Engagement.*

Moves	$\frac{m}{t}$	where:: *m* = # of moves initiated by the player in a session *t* = length of session, in minutes
Damage Events	$\frac{d}{t}$	where:: *d* = # of damage events initiated by the player in a session *t* = length of session, in minutes
Total Participation	$\frac{m+d}{t}$	where:: *m* = # of moves initiated by the player in a session *d* = # of damage events initiated by the player in a session *t* = length of session, in minutes

The *Participation* measure gauges how engaged an individual is with the exhibit, but also needed is a measure of how engaged an individual is in the collaborative learning endeavor. As mentioned in the introduction to this section, Section 6.2, researchers who assume a sociocultural or Vygotskian perspective often use properties of group *Conversation* as indicators of collaborative learning (Sawyer, 2006). The idea is that knowledge can only be internalized by first externalizing ideas in spoken form to the group. For that reason, the number of utterances made during group *Conversation* will be counted. The analysis of overall individual *Conversational* engagement will use these raw counts to compute the per-minute utterance frequency, which indicates to what degree the participant is engaged in the social context. To get a more nuanced understanding of how conversation might be used for learning, though, the utterances will also be coded by category.

The first category applied to the utterances is “on-task”/“off-task,” a binning often seen in collaborative learning research. “On-task” behaviors are generally considered to be those that are relevant to the collaborative activity at hand, and can be further discriminated into interactive and non-interactive categories depending on whether they are targeted at another participant (Hertz-Lazarowitz, 1992). For the purposes of this research, only *interactive* utterances will be considered for “on-task”/“off-task” binning – participants using software often direct their remarks at the software itself (e.g., “Come on, go away.”) or to themselves (e.g., “Oops.”) – and what is of interest here is the degree to which participants are socially engaged in an on-task manner. The number of “on-task” utterances will give some sense as to how invested participants are in the collaborative activity, and separate out participants who are actively engaged from those who are merely socially engaged (i.e., chatty). The rough “on-task”/“off-task” binning tells something about base levels of *Engagement* in the shared task, but there are two further methods for examining participant utterances. One is to determine how the utterances relate to the context of the execution of the joint activity (a functional perspective), and the other approach is to determine how the remarks relate to the context of the group’s shared knowledge building (an interactional perspective).

The functional perspective, in the case of this activity, attempts to single out utterances for their potential to make a tactical or strategic impact on the group’s behavior. By way of contrast, some remarks might be merely observational in nature – commenting on the existence of certain elements *without* proposing direct actions to be taken (tactics), or describing larger patterns that can be capitalized on for gain (strategy). Other remarks may serve a function in an interactional sense (e.g., back-channel acknowledgements of speaking turns, like “uh-huh”, or of comprehension, like “I see”) but do not alter or contribute to the group’s socially-constructed body of tactical or strategic knowledge. Because many utterances contain both tactical and strategic content, a single category, *Tactical/Strategic*, was used to encompass both. The *Tactical/Strategic* category is very similar to the interpreting/applying category used in (Borun et al., 1996), where museum conversations among families were coded as belonging to one of three levels of learning: identifying, describing, and interpreting/applying. More details will come up in Section 7.2.2.3 Evidence for Interaction between Visitors, when the analysis of the results of this coding scheme is discussed.

Table 11. Formulas used to compute the various *Conversation Individual Engagement* measures.

<table>
<tr><td>Frequency</td><td colspan="2">$\frac{u}{t}$, where u = # of utterances,
t = length of session in minutes</td></tr>
<tr><td rowspan="3">Proportion</td><td>On Task
possible codes:
[on, off]</td><td>$\frac{o}{u}$, where
u = # of utterances,
o = # on-task utterances</td></tr>
<tr><td>Functional
possible codes:
[none, Tactical/Strategic]</td><td>$\frac{s}{u}$, where
u = # of utterances
s = # tactical/strategic utterances</td></tr>
<tr><td>Interactional
possible codes:
[New Statements, Responses, Continuations]</td><td>$\frac{r+c}{u}$, where
u = # of utterances
r = # Responses
c = # Continuations</td></tr>
</table>

The interactional perspective is one often considered by collaborative learning researchers. In particular, they often look for interaction patterns within the conversational exchanges, to see if, for example, learners build upon each other’s ideas (Palincsar & Herrenkohl, 1999). In a study that contrasted groups that succeeded collaboratively against those that failed, both building upon one another’s comments and echoing one another’s remarks were seen as precursors to effective collaborative learning (Barron, 2000). Explanation giving is another conversational behavior often linked to effective collaborative learning (Bargh & Schul, 1980; Webb, 1984). The quality of explanations has been found to correlate positively with learning outcomes in small-group learning in classrooms (Webb, 1989). Perhaps for these reasons, conversational elaboration is often used as a measure of learning in museums (Scott. G. Paris & Hapgood, 2002). For the purposes of this study, only a simplistic interactional coding scheme is used: on-task utterances are classified as either *New Statements*, meaning that a conversational antecedent could not be found, *Responses*, meaning that the speaker is replying to or referencing an utterance made by a companion, and *Continuations*, meaning that the speaker is following up on an utterance he or she made previously. A participant's *Responses* and *Continuations* can be taken together to provide a rough metric for how much collaborative conversational building in which he or she is taking part.

6.2.2.3 Individual Task Performance

An advantage to adopting a game-like activity structure for a learning activity is that there are often built-in measures of task performance. If an educational software designer is careful to design the in-game tasks such that the users must attain some of the learning goals in order to be able to successfully complete the tasks, the task performance measures can also serve as a sort of proxy for measuring learning (Lyons & Pasek, 2005, 2006).

In the case of the software activity used in this research, users were asked to try to eliminate cancer from the simulated patient. One of the more modest learning goals for the application was to illustrate that while inflicting too much damage to otherwise normal cells would kill the patient as effectively as the cancer might, collateral damage may be necessary and unavoidable. The collateral damage tradeoff underlies virtually all treatment options for cancer: oncologists and other who treat cancer tend to use approaches that yield the greatest likelihood of eliminating cancer cells while only just keeping the collateral damage below the level where it kills the patient. With that in mind, the software keeps track of the damage done by each user to both normal and cancerous cells, which allows *Individual Score* measures to be computed that take this tradeoff into account.

There are several different ways of computing the *Individual Score*; the one which best mirrors the real-life tradeoffs is the *Weighted Efficacy* measure (see Table 12). In this measure, a damaged cancer cell is "worth" ten times the amount of a damaged normal cell, meaning that if a participant damaged 1 cancer cell and 10 normal cells, the *Weighted Efficacy* measure would be 0. Killing a cancerous cell incurs four times as much value as damaging it, so if a participant killed 1 cancer cell and damaged 4 normal cells, the *Weighted Efficacy* measure would be (1-(0.025*4))/(1+ (0.025*4)), or 0.82. With this measure, participants who maximize the number of cancer cells killed are rewarded, even if some normal cells are damaged or killed in the process.

Table 12. Formulas used to calculate the *Individual Score* measure of *Individual Task Performance*.

Individual Score	**Weighted Efficacy (WE)**	$\frac{c\_dead-(0.75\cdot n\_dead)+(0.25\cdot c\_dmg)-(0.025\cdot n\_dmg)}{c\_dead+(0.75\cdot n\_dead)+(0.25\cdot c\_dmg)+(0.025\cdot n\_dmg)}$ where *c_dead* = # cancer cells killed *c_dmg* = # cancer cells damaged *n_dead* = # normal cells killed *n_dmg* = # normal cells damaged To compensate for inherent O-UI scoring differences: **Adjusted Weighted Efficacy**("Simple") = WE – 0.82 **Adjusted Weighted Efficacy**("Complex") = WE – 1
	Unweighted Efficacy (UE)	$\frac{(c\_dead-n\_dead)+(c\_dmg-n\_dmg)}{c\_dead+n\_dead+c\_dmg+n\_dmg}$ where *c_dead* = # cancer cells killed *c_dmg* = # cancer cells damaged *n_dead* = # normal cells killed *n_dmg* = # normal cells damaged

Section 7.1.1 Case Study: Balancing Impact of Surgery O-UIs on Simulation describes the formative testing done to balance the impact of the two O-UIs on the simulation that every time a "Simple" participant presses the center button, the cell in the center of his or her incision rectangle is eliminated, but the surrounding four cells are damaged. Taking this into account, in the best case, a "Simple" participant's incision rectangle contains only cancerous cells, which would give a *Weighted Efficacy* measure of 1 when the button is pressed. In the worst case, a "Simple" participant must eliminate a cancerous cell surrounded by normal cells. In this case, the *Weighted Efficacy* would be 0.82, as computed above. With *Weighted Efficacy*, then, there would be no way for "Simple" users to ever score as high as "Complex" users, no matter how conscientiously they "operated." To compensate, the figure actually used in this analysis is the *Adjusted Weighted Efficacy* measure, which is computer differently for each of the different conditions. It is essentially a measure of how close a participant's *Weighted Efficacy* (*WE*) score has come to the best possible score in the worst possible situation for scoring. So, for all "Complex" users, 1 is subtracted from their *WE*, because in all situations, it is possible for them to score a perfect 1. For "Simple" users, 0.82 is subtracted from their *WE*, because their worst scenario is a cancerous cell surrounded by normal cells. The *Adjusted Weighted Efficacy* measure, then, is a large negative number when a participant is far from the best-of-the-worst score, and a smaller negative number when closer. It is possible, though, for "Simple" users to score greater than 0, if they have had the luxury of "operating" within a large tumor for a majority of a round, so this scoring method slightly favors "Simple" users.

Lest one wonder if all of these numerical manipulations have unduly skewed the *Individual Score*, an *Unweighted Efficacy* measure is also computed to provide an alternate perspective, one which values the damage and death dealt to cancerous and normal cells equally (see Table 12). This score does not capture the notion of necessary collateral tradeoffs; rather, it values the selective targeting of only cancerous cells. Naturally, then, the *Unweighted Efficacy* score innately favors "Complex" participants, because they have the option of being as precise as they would like.

The second measure related to Individual Task Performance is that of *Ownership*. The degree of *Ownership* is the proportion of a player's actions that are not replicated by his or her partners. The *Ownership* degree indicates something about how effective a participant is at coordinating his or her individual task execution with the task executions of his or her partners. Because the collaborative task division in this joint activity is largely one of territory division, in this case, the *Ownership* degree is measured in terms of the number of visits to the simulation's grid areas (see. So, imagine that Player North's first move placed her at grid location (4, 7); Player East had been at (4, 7) at some time prior; during her visit, Player South would also occupy (4, 7); and after her visit, Player East would return to (4, 7). In this case, if Player North made one and only one move, the *Ownership* degree for Player North would be 0.25 – meaning that she can lay claim to 25% of the moves that landed players into the single grid square she occupied over the course of the game. Another example: imagine that Player West made 25 moves, visiting 20 unique grid squares (and revisiting 5 of them twice), but *no* other players visited any of those 20 grid squares, his degree of *Ownership* would be 1, indicating complete ownership of the territory he visited during the game.

Table 13. Formula used to calculate the *Ownership* measure of *Individual Task Performance*. Notice that *Ownership* is a concept that only has meaning in the context of group activities wherein tasks need to be divided amongst participants. In this activity, task division is territory-based.

Individual Task Performance	**Ownership**	$\sum_g \frac{visits(p,g)}{visits(p,g)+\sum_c visits(c,g)}$ where *p* = the player *g* = each grid space in the simulation *c* = each of the player's companions *visits(p, g)* = the number of visits paid to grid position *g* by player *p*

6.2.2.4 Group Division of Attention

It can be informative to see if all members of a group divide their attention in similar ways while using the software – while educational software should first and foremost be effective at helping people learn, an important second goal is that the software be designed so that a wide variety of participants are able to use it consistently well. This is especially true for museum exhibits. If a design change needs to be made to improve the software's educational potential, then, it can be useful to know to if that design change is likely to be a simple universal fix, or will need to be specialized to accommodate participants with different behaviors. Returning to the context of *Group Division of Attention* for an example: if it is decided that some sort of "heads-up display" (HUD) is needed to help keep group members apprised of their collective status, knowing if visitors are consistent about frequently checking the Shared Display will help designers decide if the HUD should be located on the Shared Display, or on participants' O-UIs.

The first of the two group division of attention measures, *Attention Inequity* (see Table 14), is really just a further processing of the *Proportion*, *Frequency*, and *Duration* measures computed for each individual participant. Essentially, the *Inequity* is obtained by computing the standard deviation of each of these measures for each group. A quick check of the *inequity* of group members on the *Proportion*, *Frequency*, and *Duration* measures of attention will uncover any inconsistencies in gaze behavior. A low standard deviation indicates that the participants are similar in their behaviors, whereas a high standard deviation indicates dissimilarity. The standard deviation is taken across groups (and not across all participants in a condition) because group members are all experiencing the same simulation (and thus the same Shared Display stimuli), but the simulation, being emergent, can differ from group to group.

Notice that the population standard deviation, σ, is used in lieu of the sample standard deviation, *SD*, as the means of calculating the amount of deviation, or inequity. This differs from (Kapur & Kinzer, 2007), who used *SD* to calculate a participation inequity measure. The argument is that σ is a more appropriate calculation to employ because the group of participants *is* the entire group for whom we are interested in calculating a deviation, and because the deviation is used here as a *descriptive* statistic (and not as an *inferential* statistic used to form assumptions about a larger population)[27].

The three measures of attention used for individuals, *Proportion*, *Frequency*, and *Duration*, tend to flatten the nuance present in actual use scenarios, however, reducing the dimensionality of the data. Among other things, one lost data dimension is the specific temporal features of user attention behaviors. Temporality plays an important role in at least one behavior that is very important to group learning: namely, whether or not the group members are utilizing the Shared Display as a grounding space for their actions and conversations. This can be measured by determining the *Synchronicity Degree* of group members in gazing at the Shared Display. For the purposes of this study, a granularity of 2s intervals was used: essentially, the number of players gazing at the Shared Display during each 2s interval was tallied. These tallies were then used to compute measures like dyad synchronicity, triad synchronicity, quad synchronicity, although ultimately for the purposes of this study, only a single composite metric was used (see Table 14). The *Synchronicity Degree* is a measure that sheds light on the collaborative utility of the shared public display. If the groups in the "Complex" O-UI condition demonstrate significantly lower *Synchronicity Degree* values than "Simple" participants, the likely conclusion is that the complexity of the O-UI is interfering with a process known to aid collaborative learning.

[27] In any case, the differences between *SD* and σ are relatively minor – if anything, the *SD* tends to exaggerate the deviation when calculated for small samples. This exaggeration is intentional, as the sample *SD* calculation was originally designed for inferential statistics to compensate for the fact that we must use the sample mean as a stand-in for the population mean. The substitution of the sample mean results in a lower calculated deviance than would be seen if the true population mean was used, so the *SD* formula is constructed to produce larger deviation values.

Table 14. Formulas for calculating *Group Division of Attention* measures.

<table>
<tr><td>Synchronicity Degree</td><td colspan="2">$\frac{2 \cdot ((2 \cdot d) + (3 \cdot t) + (4 \cdot q))}{g \cdot l}$, where
d = the number of 2s intervals where two players were gazing at the Shared Display
t = the number of 2s intervals where three players were gazing at the Shared Display
q = the number of 2s intervals where four players were gazing at the Shared Display
g = group size
l = length of the session, in seconds</td></tr>
<tr><td rowspan="3">Attention Inequity</td><td>Proportion</td><td rowspan="2">$t \in [O-UI, SharedDisplay, OtherPlayers]$
$\sigma(\forall p \in G, [A(p,t)])$
where:
p = the player
G = the group
A(p,t) = one of the *Individual Division of Attention* measures for player *p* (Proportion or Duration - see Table 8) for a given target *t* (the O-UI, Shared Display, or Other Players]</td></tr>
<tr><td>Duration</td></tr>
<tr><td>Frequency</td><td>$\sigma(\forall p \in G, [F(p)])$
where:
p = the player
G = the group
F(p) = the gaze shift frequency measure for player *p*</td></tr>
</table>

6.2.2.5 Group Engagement

A group learning activity does not lie up to its potential if only a select number of the participants become engaged with it. *Participation Inequity* can easily be determined by a similar method to that presented in Section 6.2.2.4: computing the population standard deviation amongst a group on their individual *Participation* measures. Low standard deviations indicate equitable participation, and thus, presumably, equitable opportunities for learning. This method has been adapted from the PI, or "Participation Inequity" measure used in (Kapur & Kinzer, 2007), with the substitution of population standard deviation, σ, in lieu of sample standard deviation, *SD*. A similar method was used in (Morris et al., 2004), where the concern was the "democratic" use of a shared tabletop device, and where the standard deviation of participation was referred to as a "dominance" score.

The amount of conversation present in a group is also a sign of the group's overall degree of engagement – even if one participant is quiet, it may be the case that he or she is engaged in listening. That said, it is better for learning purposes if there is more equity in the amount of conversation between group members. Learning through conversation is known to occur by engaging in what are known as "transactive dialogues" that indicate a mutuality in conversational participation (Barron, 2000). We can assess the degree of mutuality by taking the *Individual Engagement: Conversation* measures (see Section 6.2.2.2 Individual Engagement) and once again computing a standard deviation for each group. Low standard deviations indicate that participants are similarly engaged in conversation – although one must be careful not to rely on standard deviation-derived measures alone for making interpretations, to detect those instances where the visitors were just consistently and mutually *un*-communicative (i.e., sitting around silent as stones).

Table 15. Formulas used to calculate *Group Engagement* measures.

Participation Inequity	**Moves**	$\sigma(\forall p \in G, [Moves(p)])$, where: *p* = the player *G* = the group *Moves(p)* = the # of moves made by player *p*
	Damage Events	$\sigma(\forall p \in G, [Damages(p)])$, where: *p* = the player *G* = the group *Damages(p)* = the # of damage events made by player *p*
	Total Participation	$\sigma(\forall p \in G, [Moves(p), Damages(p)])$, where: *p* = the player *G* = the group *Moves(p)* = the # of moves made by player *p* *Damages(p)* = the # of damage events made by player *p*
Conversation Inequity	$\sigma(\forall p \in G, [C(p)])$ where *p* = the player *G* = the group *C(p)* = one of the *Conversational Individual Engagement* measures for player *p*: Frequency, or Proportion (On-task, Functional, or Interactional); see Table 11	

6.2.2.6 Group Task Performance

Performance Inequity is once again a standard deviation of the *Individual Scores* (see Section 6.2.2.3 Individual Task Performance). It indicates if the group members were more or less equally able to perform well in the context of the shared activity. *Ownership Inequity* is similar: it answers the question of whether or not the group members are taking equivalent ownership of their portions of the joint task. We have one additional measure that sheds light on group performance, however: the ultimate *Outcome* of the activity. In this case, the *Outcome* is whether or not the group was able to succeed in eliminating the cancer from the simulated patient. Groups typically played multiple "games," so the *Outcome* can be computed across all of the "games" (see Table 16).

Table 16. Formulas for computing the measures of *Group Task Performance*.

<table>
<tr><td rowspan="1">Outcome</td><td colspan="2">Percentage of successful outcomes
$\frac{success}{success + fail}$
where
success = # simulated patients where cancer was eliminated
fail = # simulated patients who “died”</td></tr>
<tr><td rowspan="2">Score Inequity</td><td>Weighted Efficacy Inequity</td><td>$\sigma(\forall p \in G, [WeightedEfficacy(p)])$
where
p = the player
G = the group
WeightedEfficacy(p) = the Weighted Efficacy Individual Score for player p</td></tr>
<tr><td>Unweighted Efficacy Inequity</td><td>$\sigma(\forall p \in G, [UnweightedEfficacy(p)])$
where
p = the player
G = the group
UnweightedEfficacy(p) = the Unweighted Efficacy Individual Score for player p</td></tr>
<tr><td>Ownership Inequity</td><td colspan="2">$\sigma(\forall p \in G, [Ownership(p)])$
where
p = the player
G = the group
Ownership(p) = the Ownership Individual Task Performance measure for player p</td></tr>
</table>

6.3 Research Approach: Design-Based Research

Chapter 2 acquainted the reader with an accounting of the physical, social, and historical-cultural facets of the learning experiences to be had in informal museum learning environments. It is necessary to have a solid grasp of these three components in order to make use of a Design-Based Research (DBR) approach, as the research presented here does. While it might seem facile to note that people's behaviors are affected by their surroundings (the *physical* context), their companions (the *social* context), and learned social conventions (this *historical-cultural* context), it took researchers in the educational community many decades to acknowledge that there were inherent limitations to what could be understood about learning if these contexts were ignored. For a good portion of the 20th century, educational researchers tried to study learning in controlled (and often laboratory) settings. It is important to understand the changes educational research has gone through to understand the importance of context to the Design-Based Research (DBR) methodology, and how knowledge of that context is used by DBR, so the history underlying the creation of DBR will now be briefly reviewed. Then how a DBR perspective has been applied to this work will be described.

6.3.1 Inception of DBR

Education research has long been influenced by psychological research methods, and like the field of psychology, it went through its own behaviorist and cognitivist phases in the mid-20th century. More context-sensitive attitudes came to the fore in the latter quarter of the century, perhaps influenced by the rediscovery of L.S. Vygotsky's work studying the crucial role of social relationships in learning. Historical-cultural psychology (Vygotsky, 1978), situated cognition (J. S. Brown, Collins, & Duguid, 1989; Suchman, 1987), situated learning (Lave & Wenger, 1991), everyday cognition (Lave, 1988; Rogoff & Lave, 1984), and cognitive apprenticeship (Collins, 2006) are just a few of the schools of thought that have collectively formed what is termed the *socio-cultural perspective* on learning.

Heavily influenced by socio-cultural theories of learning, one of the major drivers behind the formulation of DBR as a methodology was the recognition that many educational interventions developed in laboratory (or otherwise tightly controlled) environments seemed to flounder when implemented in more naturalistic environments (Ann L. Brown, 1992). The laboratory-derived interventions, resulting as they did from controlled experiments, often would not have "ecological validity" when applied to other environments (like actual classrooms) – too many factors (or the wrong factors) were left out of the experimental conditions (A.L. Brown & Campione, 1996). DBR, on the other hand, embraces the notion of situated cognition (J. S. Brown et al., 1989; Suchman, 1987) and acknowledges that a context is inextricably linked to educational outcomes.

6.3.2 Defining DBR

DBR was influenced by some extent by the same anthropological perspective that characterizes much of the work on situated learning, but is not just a rooted descriptive methodology. Lest one think that DBR produces just narrow accounts of behavior in certain, very specific locales, it is worth noting that what is now called "Design-Based Research" was originally dubbed "design experiments" (Ann L. Brown, 1992), and experimentation in the traditional sense is still a part of DBR. The goal for DBR research is dual: to (1) formulate and test new theories of learning, while simultaneously (2) creating functional educational interventions that produce positive educational effects (Barab, 2006). Thus, it has an "intermediate theoretical scope" (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003) that places it squarely in "Pasteur's quadrant" of scientific research that is both concerned with advancing the frontiers of science while simultaneously being concerned with how the advancements will be applied (see Table 17). The goal of DBR is not to produce interventions that will work in only one, singular (naturalistic) context; researchers using DBR attempt to uncover generalizable theories of learning useful to other researchers, which can be encoded into actionable "design guidelines" useful to practitioners.

Table 17. Pasteur’s Quadrant, a method of characterizing research by the degree to which it (a) advances the frontier of understanding, and (b) takes into consideration how the fruits of the research will be used. Adapted from (Stokes, 1997). Design-based research falls in “Pasteur’s quadrant,” the class of research that is both dedicated to advancing the frontier of scientific understanding while also taking into account how the research will be applied.

		Consideration of Use?	
		No	**Yes**
Advancing Frontier of Understanding?	**Yes**	Pure basic research (Bohr)	Use-inspired basic research (Pasteur)
	No		Pure applied research (Edison)

Controlled experiments are still performed in DBR; the key difference being that the experiments arise from a bottom-up process, as opposed to a more traditional top-down process. As Cobb et al. describe it, “one of the distinctive characteristics of the DBR approach is that the research team deepens its understanding of the phenomenon under investigation while the [research] is in progress,” meaning that formal experiments, rather than being devised from the outset, emerge as researchers become exposed to interesting emergent phenomena (Cobb et al., 2003). DBR research iterates through two phases: one that is prospective, and involves trying out new ideas predicted by the current theories of learning; and one that is retrospective, which involves following up on interesting phenomena observed during the prospective phase, often by means of devising traditional controlled experiments to isolate variables of interest. DBR allows researchers working with highly complex environments to make larger strides than they would be able to via traditional experimentation alone – in a sense, DBR is very similar to a hill-climbing maximization algorithm, with random restarts (the “prospective” phases) to allow researchers to get out of local minima. What might be curious to those originally trained in engineering disciplines is that educational researchers view the iterative nature of DBR as being a derivation of engineering approaches (Barab, 2006; Cobb et al., 2003), like the spiral model used in software engineering (Boehm, 1988). Like software engineering, In DBR prospective and retrospective phases are applied iteratively to improve a tangible expression of the emergent theoretical understanding: an educational intervention.

6.3.3 Applying DBR

A DBR methodology was chosen for this work because, when properly executed, it seems to offer special advantages for those attempting to integrate new technology into existing social contexts. It is altogether too easy to create a technological intervention for learning environments that, while theoretically sound, simply fails to work as intended. Conversely, it is possible to create a technological intervention that works very well in a specific environment, but because the designers are unable to articulate (or empirically verify) theoretical relationships between certain design elements and desirable outcomes, the innovations are hard to translate to new contexts.

DBR, as with any methodology, contains many variants under its umbrella, many of which are summarized in (F. Wang & Hannafin, 2005). The variant closest to the approach used in this research is “development research”, which begins with literature review, expert consultation, analysis of examples, and case studies of current practice. The fruits of those processes are reported in Chapters 2, 3, and 4. The next stage in “development research” is to interact and collaborate with research participants to approximate interventions – essentially, to engage in formative evaluation. Empirical testing of interventions is then recommended, so that the knowledge generated by the research can be encoded into heuristic statements, or design guidelines, that carry some empirical heft. Such an ordering was followed by this research, as the next two sections will describe.

6.3.3.1 The “Prospective” Phase: Phase I: Formative Research

This research began with a “prospective” phase (see the prior Section 6.3.2 Defining DBR), that originated in a close reading of how mobile devices are currently being deployed in museums (see Sections 3.3 Handheld Devices in Museums and 3.4 Multi-Machine User Interfaces in Museums). From that a form-factor was identified (the Multi-Machine User Interface approach) that, based on existing theories of learning, should promote collaborative learning (see Section 3.4 Multi-Machine User Interfaces in Museums), and thus was deemed worthy of further iterative exploration. From the review of existing technology in museums a phenomenon worthy of further experimentation (the heads-down effect, see Section 3.3.1.3 Access to Companions while engaged in Single-User Handheld Device Activities) was identified, as was an independent variable (the “complexity” of the mobile device interface, see Section 6.2.1 Independent Variable: Opportunistic User Interface Complexity) that might impact the phenomenon.

Continuing the “prospective” phase, an educational intervention – *MUSHI-Lignancy* – was constructed, based on the MMUI paradigm. Two competing user interfaces for the Opportunistic User Interfaces (O-UIs) were devised, which would allow for experimentation with the independent variable (O-UI “complexity”) to see if it did indeed have an impact on the observed phenomenon (the heads-down effect). This period of formative research is described in Section 6.5, and was the first phase of the research study. The purpose of this formative phase was to conduct iterative *in situ* design revisions on the competing interfaces for the O-UIs, to ensure that none of the interfaces had major design flaws that would make the experiment of Phase II a straw-man test. While engaged in iterative formative testing, Design-Based researchers are encouraged to “generat[e] a comprehensive record of the ongoing design process”

(Cobb et al., 2003), because the changes that get made, an their rationales, will likely provide the fuel for further prospective and retrospective cycles. Examples of the insights gleaned from this process are presented in Section 7.1 Phase I: Formative Study Results.

6.3.3.2 The "Retrospective" Phase: Phase II: Lab-Based Experiment

The "retrospective "phase of this research study took a more traditional experimental form than Phase I's formative testing. Conducted in a controlled, lab-like environment outfitted with A/V recording equipment, this portion of the study pitted one condition (the *MUSHI-Lignancy* MMUI with "Simple" O-UIs) against the other (the *MUSHI-Lignancy* MMUI with "Complex" O-UIs), to determine if O-UI "complexity" (see Section 6.2.1 Independent Variable: Opportunistic User Interface Complexity) did indeed correlate with the heads-down phenomenon.

6.4 Research Site: The Exploratorium

The Exploratorium is a very large science center located near the Golden Gate Bride in San Francisco, CA. It was founded in 1969 by Frank Oppenheimer (brother of Robert of atom bomb fame) who had spend several years developing hands-on teaching apparatuses while instructing physics at the University of Colorado. One hallmark of these teaching apparatuses is that they would allow users to directly interact with or manipulate the scientific phenomena under study. He envisioned a museum that would contain similar apparatuses as exhibits, and would allow visitors, especially children, to engage in what would later variously be known as hands-on, free-choice, discovery, or inquiry learning. It was created in an era of increased attention to social justice issues, and the power-to-the-people mentality of the time strongly shaped the way exhibits were designed for interaction and interpretation (see Section 2.1.2 Interpretation and Authority for a lengthier discussion of this issue). The Exploratorium, as an early adopter of and continued leader in promoting exhibits based around constructivist theories of learning, is an ideal institution for testing software-based exhibits that similarly adopt a constructivist perspective (as opposed to the didactic perspective still seen in many modern museums, exemplified by the A/V Guides of Section 3.3.1 Single-User Handheld Device Activities in Museums).

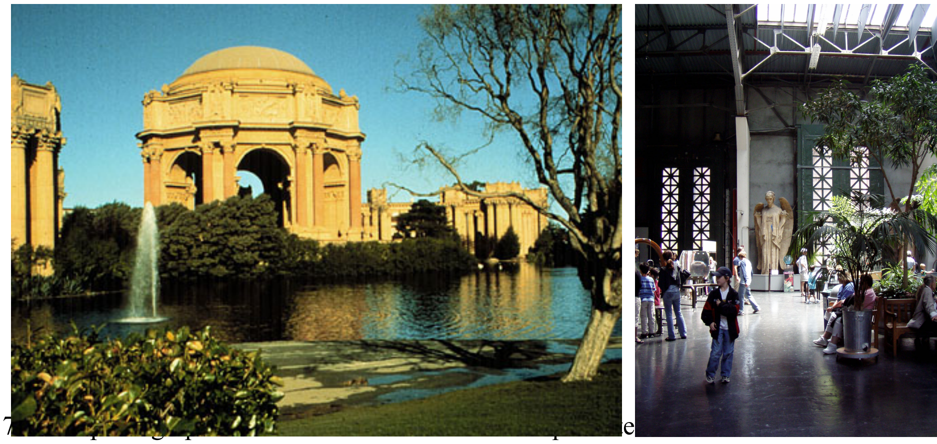


Figure 7 … e … s, which houses the Exploratorium, taken from (Exploratorium, n.d.-b). The photograph on the right, taken by the author under the skylights, shows the height and dimness of the interior space.

6.4.1 Physical Context

Oppenheimer secured a space in the run-down former Palace of the Fine Arts, left over from the 1915 Panama-Pacific International Exposition, which was torn down to its steel superstructure and rebuilt. The interior space is very large, taking up 110,000 square feet, and houses around 400 distinct exhibits at any one time (Exploratorium, 2008). Despite the elegant exterior finish, the interior is very bare-bones, resembling an aircraft hangar (see Figure 74), and in fact was used to store military vehicles during World War II. The architecture is important to mention because it imposes constraints that shaped the course of this research.

The building is in the shape of a long crescent, with both the entrance and exit located on one end of the crescent (see Figure 75). The enclosed laboratory space used in Phase II of this research was located at the opposite end of the building from the entrance, so even though the museum opened at 10 A.M., visitors would not begin to reach the far end of the museum until around 1 P.M. (A good proportion never even reached the opposite end of the museum during their visits). This made recruiting participants on the floor a challenging experience – groups had to be recruited between 1 and 3 P.M., because after 3 visitors began to get antsy and would exhibit “museum fatigue” (Maximea, 2002a), already plotting their escape from the museum.

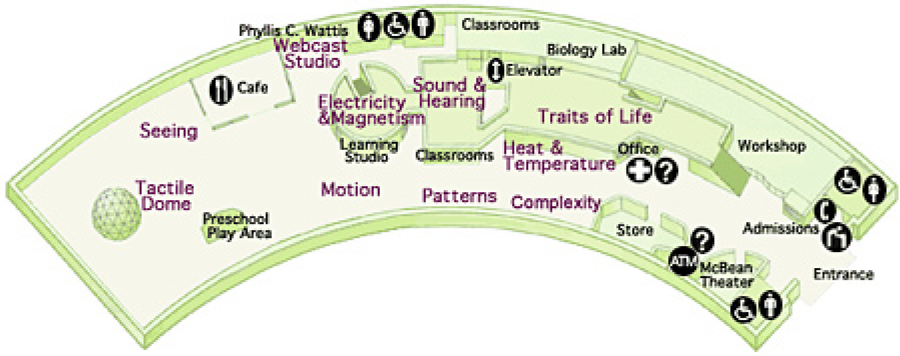


Figure 75. Floor map of the Exploratorium, taken from (Exploratorium, n.d.-a). Phase I of this research took place in front of the museum store, located on the inner arc of the floor. Notice that both the entrance and exit are located on one end of the museum (the right side of the image), and that the lab space used in Phase II of this research is located on the opposite end, behind the *Tactile Dome* exhibit. The mezzanine level in the middle that houses the *Traits of Life* collection of exhibits is where Phase III of this research was located.

There is a second, not-enclosed mezzanine level that is reached via stairs or an elevator (it has the feel of a free-standing balcony). The center of the mezzanine level is home to the *Traits of Life* collection of exhibits, the only region of the Exploratorium to deal with the life sciences. This region would have been the ideal contextual setting for the experiment to take place, but it came with its own challenges. Owing to its second-floor status, this area is not as well-trafficked as the rest of the museum. The main stairway to the mezzanine level faces the entrance to the museum, and so many visitors opt to stay on the main floor as they enter. Much like the far end of the museum, the mezzanine tends to have highest attendance only during the middle-to-late portion of the day, so recruiting could only be done from around noon until 3 P.M. The Palace of the Fine Arts has very high ceilings, cement floors, and an open floor plan that result in a cacophony of noise and activity. This can make clear audio recordings very tricky to obtain. Moreover, because there are no windows in the structure, aside from a single set of skylights in the middle of the arc (see Figure 74), the interior space is very dim, making photography difficult. These data collection challenges prompted the positioning of the controlled-experiment phase of this research, Phase II, in the laboratory space, despite its suboptimal physical context.

6.4.2 Social Context

Over 562,000 visitors attend the Exploratorium each year, with 50% being local Bay Area residents, and only 20% attending from out of state or from another country (Exploratorium, 2008). The Exploratorium's employees had the perception that their museum's average age skewed higher than many other science museums, but their ratio of 51:49 adult:child attendees (Exploratorium, 2008) was very close to the 50:50 split seen at other science centers (see Section 2.2.1 Demographics at Science Museums).

The Exploratorium hires around 75 high school students each year to serve as "Explainers," a role similar to that of docents found in other museums. They answer questions, help maintain exhibits, and give demonstrations for visitors. The Explainers are a useful resource for researchers, because they can provide a dual insider/outsider perspective on new exhibits (Diamond, 1999). Explainers have been used to good effect as participants in other technology studies at the Exploratorium (Hsi, 2003).

6.5 Phase I: Formative Research

Section 6.3.3.1 The "Prospective" Phase: Phase I: Formative Research describes how the first phase of this research study takes the form of a "prospective" phase of DBR research, wherein the impact of an *in situ* educational intervention is explored. In particular, the "Simple" and "Complex" user interfaces to be used in the second, experimental phase of this research needed to be refined, to be sure that each functioned reasonably well given its design constraints. The goal was to construct interfaces that operated in an equivalent fashion on the shared simulation, but had drastically different UI "complexity." The "Simple" version could only make use of the built-in hardware buttons on the handheld to provide input. The "Complex" version needed to provide dynamic, on-screen output, and require that the user interact with that output using a stylus (see Section 6.2.1 Independent Variable: Opportunistic User Interface Complexity for a deeper discussion of why these two conditions would adequately represent "Simple" and "Complex" user interfaces in this context).

A list of potential roles (surgery, radiation therapy, proton beam therapy, and chemotherapy) were derived from common real-world treatment options. To be a candidate for Phase II, a role had to be plausibly implemented in both a "remote-control," buttons-only fashion for the "Simple" O-UI condition, and in a dynamic, stylus-interactive fashion for the "Complex" O-UI condition. By "plausibly implemented," the intention is that an interface would not seem like some sort of retrofitted Frankensteinian creation that exists solely to serve the purposes of the Phase II experiment, but would rather be an interface that could conceivably exist on its own "in the wild." Phase II needed to be able to avoid the charge of being set up as a straw-man test by virtue of user interfaces that were either implausibly simple or implausibly complex for the educational activity. To determine the plausibility of the interfaces, formative testing needed to be conducted *in situ*, on the floor of the Exploratorium, with real visitors.

6.5.1 Physical Context

MUSHI-Lignancy was set up on a standard 6'x2.5' folding table near the Exploratorium's gift shop. This is a high-traffic area likely to draw the attention of visitors both entering and exiting the museum (see Figure 76). Seating was provided in the form of two long benches: one was placed in front of the table for visitors to sit at, and another was placed behind for the researchers to sit on. In terms of computer equipment, the HP Compaq tc4200 that was used to run the *MUSHI-Lignancy* simulation was also used to display the "large" public view (see Figure 77) – in this fashion, Phase I differed from the other phase, which presented the public view on a truly large display, a 4' plasma screen. This was done primarily for convenience – the lightweight setup allowed for rapid deployment after each design revision. Thus, an *in situ* test could be conducted immediately after making an incremental change to the user interfaces, making this phase of the study extremely iterative. The handheld devices were the same as used in other phases, Hewlitt-Packard iPaq h4100, and were placed in cradles on the table, near the tablet computer.

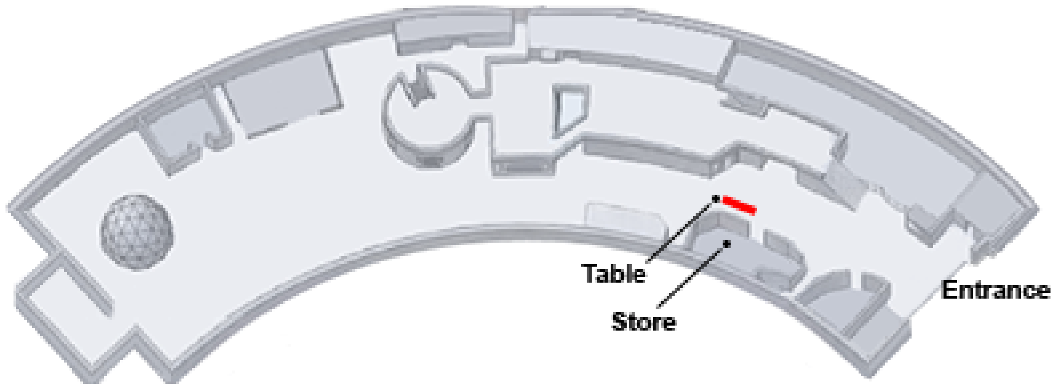


Figure 76. The location of Phase I. A folding table was set up outside the gift shop, located near the entrance to the museum.

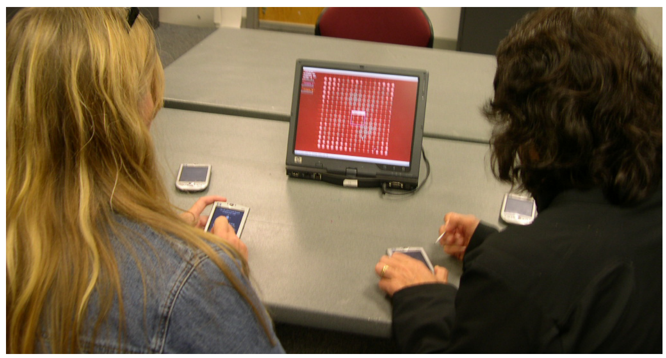


Figure 77. This photograph, taken b the author, is not of the actual setup used in Phase I, but is similar in terms of equipment used (photographs were not taken of participants during Phase I). The Shared Display is just the screen of the tablet computer, rotated to face the users.

6.5.2 Recruiting Participants

A sign was taped to the front of the table, reading "Would you like to help us test a game?" The author has since learned that some museum professionals avoid using the word "test" when conducting formative evaluations, because it has the possibility of making visitors feel as if *they* are the ones being tested (S. M. Taylor, 1991). This can reduce the amount of useful feedback from visitors, because if they feel that the misunderstanding is their fault, they are less likely to articulate which aspects of the design were giving them problems. The sign did not seem to inhibit visitors, however – while some people did just look at the exhibit and pass on, those who did stop to try were not shy at all about sharing their views. Visitors were free to gather around, sit down, play with the interfaces, and leave according to their own desires.

6.5.3 Procedure

This phase was formative research, and so had no formal "procedure" as such, although the typical way the formative sessions played out can be described. Two researchers (the author, and an Exploratorium staff member or volunteer) would sit behind the table that *MUSHI-Lignancy* was set up on. (Visitors were more likely to approach the table if two researchers were present, instead of a solo researcher). Museum visitors were free to approach the table where the *MUSHI-Lignancy* simulation was set up. If a group of visitors approached the table, the author would ask them if they would like to try out a new game, to get their opinions on how to make it better. The term “game” was used both because of the simulation’s superficial resemblance to a video game, and because it was a term more easily understood than the term “complex systems simulation.” If the group seemed interested, the author would explain further that the game allowed them each to play a different type of doctor, and work together to try to fight cancer in a simulated patient.

Once a group decided to participate, the author would explain what they were seeing on the shared, public screen – and what each of the different components of the simulation were (the normal cells, the healthy cells, and the blood vessels). No user seemed to have a difficult time understanding what any of these elements would be, or what the public screen of the simulation was showing, possibly because the “growth” of the cancer is very animated and visible in the simulation. The author would explain the different roles to whatever depth visitors desired, starting with relatively shallow, functional descriptions rooted in the “game” experience (e.g., “you shoot a beam of radiation into the patient to try to kill as many cancer cells as possible”) and progressing to more in-depth explanations (e.g., “radiation kills cells by disrupting their DNA”) if the visitor requested or otherwise indicated a deeper interest. A fair proportion of visitors did have trouble with understanding exactly what the proton beam was, even after an explanation, although chemotherapy, surgery, and radiation seemed to all be familiar concepts to most visitors.

The visitors were then instructed to pick up the handheld devices, choose a role, and log in to the simulation. In the few instances when a visitor had a problem with this part of the procedure, the author would help them, since the log-in interface design wasn’t really a part of this investigation. For the first part of the formative evaluation, the focus was on developing only the “Complex” O-UI interfaces. While the visitors were using the interfaces, the researchers would take notes, sometimes while standing behind them or to the side to see what they were doing with their O-UIs. Before standing behind them, the researchers would ask the visitors if they were comfortable with them looking over their shoulders. No visitor refused, nor did any seem discomfited, probably because throughout the formative evaluations the researchers were careful to adopt a cheerful, casual tone, and were engaged in an active conversation with the visitors throughout.

Throughout the visitors’ experiences they would be asked questions, and their replies were recorded (as were any comments offered up unasked). If someone was having difficulty, the author would ask them what they were trying to do, and how they were trying to do it, before showing them how they could accomplish the task with the current implementation of the user interface. After giving them a few moments get comfortable with the demonstrated usage, the visitor was asked if, now that they knew what they were doing, they could think of either a better way to do the same task, or a better way of making how to do the task clearer. This proved to be a very good way of eliciting useful design feedback: with their former confusion fresh in their minds, but with a newfound understanding of the task, visitors in this situation were able to give very directed feedback on the user interface design (see Section 7.1 Phase I: Formative Study Results for some of these results).

After a group played through an instance of the simulation (i.e., after the patient lived or died) visitors were given the option of trying again, and the added option of changing the role they would play in the game. Some groups chose to go 3 or even 4 rounds so they could try each role. Other groups physically traded handhelds in the middle of a "game," and got experience with the different roles that way. Visitors were not required to assume multiple roles, but the majority of visitors were curious enough to do so. When the group concluded their experience, they were asked three questions: what they liked, what they didn't like, and what surprised them about their experience with the "game." This latter question seemed the best at provoking comments that were helpful for correcting user interface flaws, because it revealed instances where there were mismatches between the interface design and visitor expectations (see Section 7.1 Phase I: Formative Study Results for some of these results).

Those visitors who had tried different roles were asked which roles they liked the best, and why. Surgery was the hands-down favorite, although visitors were not very articulate about why ("it's the most fun," while valid feedback, doesn't help an interface designer much). Chemotherapy was, conversely, the dud of the different interfaces: throughout several iterative redesigns, it consistently received the worst reviews ("boring," "not fun at all," and "not enough to do" were some of the comments). Unlike the other three roles, which impacted user-selected parts of the simulated patient, chemotherapy operated systemically on the entire patient, and so there was less of an interesting relationship between the private device and the public display. The pace of interaction was also a lot slower – the user would have to choose a type and a dosage of chemotherapy, administer it, and then wait a while to determine the effects before administering it again. In a sense, the feedback loop was a lot "looser" for chemotherapy than the other roles. The problems with the role could not have been fixed without misrepresenting how chemo is applied in real life, so the chemotherapy role was dropped from the lineup (and which is why its interface is not described in Chapter 5).

Once the remaining O-UI interfaces reached a stable form (after many formative iterations), it was time to decide which role or roles to construct a "Simple" interface for, for use in Phase II. (Phase II was limited to a single role to increase the power of the analysis). The Surgery role seemed the most natural to adopt: it was both the most highly-interactive of the three "Complex" O-UIs (thus the most likely to draw visitors' attention to the handheld, and get them caught up in the heads-down phenomenon), as well as the most popular with visitors. Thus the "Simple" version of the Surgery O-UI was implemented, and more *in situ* formative testing was conducted with it. No changes needed to be made to the initial assignment of functionality to hardware buttons (a directional pad for movement, and the center button for triggering an "excision" of the tissue under the incision rectangle – see Section 5.1.2.1 Surgery Role for details). Subsequent iterations focused on tuning the amount and scope of the damage produced when visitors would press the "excise" button, so that the "Simple" O-UI was about as efficient as the "Complex" Surgery O-UI at eliminating cancer cells (see Section 7.1.1 Case Study: Balancing Impact of Surgery O-UIs on Simulation for a case study)

6.6 Phase II Lab-based Experiment

Controlled, lab-based experiments, once a research mainstay, no longer make up the bulk of research on technology targeted at children. In a survey of ten years of published research on technology for children, lab experiments came in third (37%), with field studies (53%) and "action research" (a category that includes DBR) coming in second at 42%[28] (Jensen & Skov, 2005). Much can still be gained from lab-based experiments, however, especially in circumstances where the variables under investigation are highly likely to be influenced by other factors. In the case of this phase of the research, the goal was to understand how visitors' attention would be affected by alternate O-UI designs. One truth about museum floors in general, and the Exploratorium in particular, is that they are full of visually and auditorally distracting elements. Therefore, the ability to control for external distractions was deemed to be an advantage that outweighed the "ecological validity" that testing in an *in situ* context would have provided.

6.6.1 Physical Context

The lab used for this phase of research was located in a small, controlled, (nearly) soundproof room located at the end of the museum opposite from the entrance (see Figure 78). It was created for use by Sue Allen's in-house Visitor Research and Evaluation group[29], and the author was fortunate enough to be given access to the space while her group was in a data analysis phase of their research. It was outfitted with state-of-the-art recording equipment, and a separate room that housed the computer system and hard drive array used to collect video and audio signals.

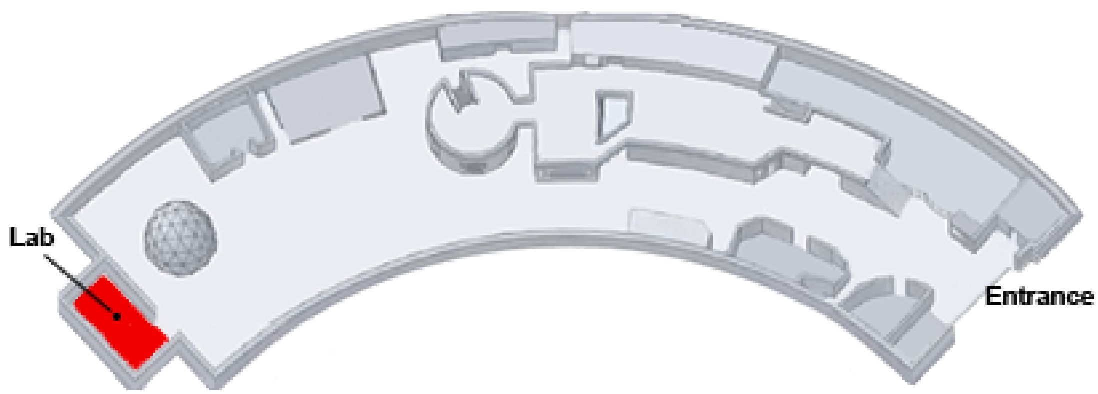


Figure 78. The location of the controlled lab used in Phase II. It was located behind a closed door at the far end from the entrance to the museum.

---

[28] The percentages do not add up to 100 because some studies use multiple methods.

[29] A reader will have noticed that Dr. Allen's work has been referenced elsewhere in this paper; she is one of the preeminent visitor studies researchers.

The interior of the lab is shown in Figure 79. Unlike the rest of the Exploratorium, it is well-lit, (relatively) quiet, and empty of other people. The extra objects visible in the background (and, indeed, the large table *MUSHI-Lignancy* was set up on) are exhibits used by Dr. Allen's group in their research. During this study, the other exhibits were covered with large blankets so that they would not distract visitors. The ceiling-mounted microphones are clearly visible, as are the wall-mounted and tripod-mounted video cameras[30].

Figure 79. Photographs of the *MUSHI-Lignancy* setup used in Phase II taken by Sherry Hsi of the Exploratorium. The image on the left shows the double doors that were opened during participant recruiting, and the alternate video camera (the position of this camera was jockeyed about to give the best views of participant faces). The image on the right shows the 4' plasma screen and handheld computers used by participants, as well as the stools used for seating. The hanging microphones used to capture audio are visible, and the structure mounted on the wall behind the plasma screen is the primary video camera.

Visitors would sit on low stools in front of the 4' plasma screen on which *MUSHI-Lignancy* was displayed. The handheld devices were arrayed in front of each spot, and each had unique identifiers. Cardinal directions were used to identify each position – North, South, East, and West – to avoid any ordinal prejudices (e.g., "player #1" has certain connotations). The identifiers were also placed on name badges worn by participants, both so that they did not need to use their names to refer to one another to preserve anonymity, and also to make the later matching of video data to log data easier to perform.

6.6.2 Recruiting Participants

Two methods were used to recruit participants for this phase of the study. An announcement was placed in the Exploratorium's member newsletter, announcing the study and providing contact information should anyone wish to participate. Participants recruited in this fashion were scheduled for early-morning time slots, so early afternoons were free, as that was the optimal time for recruiting on-the-spot participants. Recruiting from these two pools (members and average visitors) ensured that the findings would apply to

[30] It should be noted that, out of respect to Dr. Allen's ongoing research project, the author was not able to rearrange the video cameras entirely to suit the needs of the experiment. This is the reason why the number of subjects varies in some of the visual attention measures – the wall-mounted camera could not be repositioned to guarantee that the participants could not move out of frame.

both “serious” and “casual” visitors Members tend to be more in search of learning experiences (in other words – concerned about whether or not an experience provides new or interesting knowledge) than the average visitor, who tends to be more sensitive to the affective components of the visit (in other words – whether or not an experience is enjoyable). Both perspectives are valuable to obtain (although later analysis revealed no differences between the groups on the measures used in this study). A truly representational sampling was not obtained, as only 10,000 out of 562,000 annual visitors to the Exploratorium (roughly 2%) are members, whereas 6 out of 41 participants (15%) were members in this study.

When recruiting participants from the floor, a more “systematic” sampling procedure than a representative sampling procedure was used, meaning that an attempt was made to balance certain demographic characteristics (e.g., age and gender) (Diamond, 1999). (Owing to the repeated-measures design, each group was exposed to both conditions, so the systematic sampling was performed to prevent, say, all of the older females from experiencing the “Simple” condition first). In the background research, both age and gender were identified as factors that could influence how visitors responded to computer-based exhibits (see Section 2.2.1.2 Implications for Computer-Based Exhibits). Unless there was a severe imbalance, though, *any* group of three or four visitors passing near the laboratory where Phase II was conducted were recruited. During active recruiting, the double doors to the lab were left open, showing the running *MUSHI-Lignancy* simulation. Sometimes visitors approached the researchers and asked to participate.

6.6.3 Conditions

This experiment had two conditions, one wherein a visitor group would use of a “Simple” O-UI, and one wherein the visitors would use a “Complex” O-UI (see Section 6.2.1 Independent Variable: Opportunistic User Interface Complexity for a longer discussion of O-UI “complexity”). To make the most of the small number of participants, an “incomplete” repeated-measures design with rotation was used, meaning that each group experienced *both* of the experimental conditions, but the order of exposure was varied across subject groups (see Table 19 for the simple Latin Square used). The order in which the conditions were assigned was initially selected randomly, and subsequently rotated, to counterbalance any practice effects (Shaughnessy, Zechmeister, & Zechmeister, 2008). Each the first group of the day was randomly assigned to a sequence, and alternate sequence assignments were used for the rest of the day. In both conditions, the players were assigned the same role – Surgery.

Overall, 6 groups started with the “Simple” condition, with an average group size of 3 (18 participants total). Seven groups began with the “Complex” condition, with an average group size of 3.3 (23 participants overall). One group of four participants and one dyad from the first sequence (simple → complex) opted not to continue on into the second condition, and two dyads from the second sequence (complex → simple) chose not to continue on for the second condition – reducing the number of participants in the second condition by 10 (see Table 19). Thus, the total number of groups was 11 per condition, with 37 participants in the “Simple” condition and 35 in the “Complex” condition (see Table 18).

Table 18. Total numbers of participants in each condition (note – because this is a repeated measures study, some participants experienced the "Simple" condition first, followed by the "Complex," and vice versa).

O-UI Complexity Condition

"Simple"	"Complex"
n participants = 37 n groups = 11 average group size = 3.4	n participants= 35 n groups = 11 average group size = 3.2

6.6.4 Procedure

After entering the laboratory room and having a seat, the lead investigator would read an introductory speech (found in Appendix A). This speech explained the participants' rights, outlined the course of their experience if they chose to participate, and briefly explained *MUSHI-Lignancy* and the role they would be playing. After distributing the consent (and, if minors were involved, assent) forms, the visitors were each issued a name badge sticker with the letter corresponding to their position (N, S, E, or W) written on it. They were also each given the corresponding handheld labeled with N, S, E or W. (Internally, each of these handhelds had a text-based configuration file that would allow the position letter to be set in advance, so the data logs would correspond as well).

Table 19. Table illustrating the procedure, and the number of participants and groups in each of the two Latin Square sequences. Notice that between the applications of conditions 1 and 2, one dyad and one group of 4 participants dropped out before trying the Complex condition in the first sequence, and two dyads both dropped out from the second sequence (and one of these groups also declined to complete the questionnaire).

	Time →			
Condition Sequence	**Condition 1**	**Questionnaire**	**Condition 2**	**Interview**
Simple → Complex	n = 18 groups = 6 group size = 3	n = 18 groups = 6 group size = 3	n = 12 groups = 4 group size = 3	n = 12 groups = 4 group size = 3
Complex → Simple	n = 23 groups = 7 group size = 3.3	n = 21 groups = 6 group size = 3.5	n = 19 groups = 5 group size = 3.8	n = 19 groups = 5 group size = 3.8

The participants were then allowed to engage with the simulation. If any participant was observed having trouble logging in to the simulation, the lead investigator or a helper (an Exploratorium employee or volunteer) would step in and assist. At the conclusion of each "game" (marked by the participants either eliminating the cancer from the simulated patient, or the patient dying) the participants were asked if they would like to try again. The groups, on average, chose to play 4.6 games in their sequence's first condition, although those groups who were assigned to the simple condition played a few more rounds (5.3) than those initially assigned to the complex condition (4).

When the members of a group indicated that they were content with the number of games that they had played thus far, written questionnaires were distributed to the participants (see the next section for a more detailed explanation, and Appendix B to view the actual questionnaire). The questionnaires were distributed at this point in the procedure to ensure that participants' responses pertained only to the first condition, and their perceptions wouldn't be altered or confused by being exposed to multiple conditions. At this point, some of the participants decided to decline participating in the second half of the sequence, as noted in Table 19.

Those groups that continued on to the second condition of the sequence chose to play a similar number of rounds (4.1 on average) as in the first condition. When the group indicated that they were finished with the second condition in the sequence, they participated in an open-ended interview. The interview was administered at this point in the sequence so that the participants could share their opinions on both of the contrasting conditions. The intent was to try to elicit comparisons that could be used in future, prospective phases of research.

6.6.5 Instruments

This phase of the study collected five main groups of data: video recordings of participants using the software, data logs of the moment-by-moment state of the simulation and the users' interactions with the software, written questionnaires, and video recordings of open-ended interviews with the participants. Each of these will be discussed in turn. Table 20 provides an overview of which of the different instruments were used to capture data for the dependent variables in this study.

Table 20. Instruments used to capture data for the dependent variables in this study. See Table 7 for more detailed descriptions of the components of the variables.

Dependent Variables	Variable Components	Instruments
Individual Division of Attention	**Proportion Frequency Duration**	Video records of software use (see Table 8 for details)
	Awareness	Questionnaires
Individual Engagement	**Conversation frequency**	Video records of software use (transcriptions of audio portion) (see Table 11 for details)
	Participation	Software Logs (see Table 10 for details)
Individual Task Performance	**Individual score**	Software Logs (see Table 12 for details)
	Ownership	Software Logs (see Table 13 for details)
Group Division of Attention	**Attention Inequity Synchronicity**	Video records of software use (see Table 14 for details)
Group Engagement	**Participation Inequity**	Software Logs (see Table 15 for details)
	Conversation Inequity	Video records of software use (transcriptions of audio portion) (see Table 15 for details)
Group Task Performance	**Group Score Score Inequity Ownership Inequity Outcome**	Software Logs (see Table 16 for details)

6.6.5.1 Video Records of Software Use

Video recordings are the primary source of data used to answer the first set of research questions put forth in Section 6.1.2 Questions. The photographs in Figure 79 depict the wall-mounted camera and mobile, tripod-mounted camera used to collect video data. The wall-mounted camera feed was the primary source of data for this analysis, and the tripod-mounted camera used in instances when a visitor's head became obscured. All audio was taken from the ceiling-mounted microphones. See Figure 80 for a top-down diagram of how the equipment was arranged.

For the purpose of consistency, prior to analysis each of the video files originating from the wall-mounted camera were cropped to begin at the moment the all players logged into the simulation (when their "incision" rectangles became visible), and to end at the moment the last "game" in the session ended (when a popup message indicating the outcome became visible). The "incision" rectangles and popup were visible to the wall-mounted camera, as they were displayed on the Tablet PC tucked behind the plasma screen, which was running the simulation and splitting its video output to both its native LCD screen (which faced the camera) and the plasma screen (see Figure 80). The video files originating from the tripod camera were cropped to have the exact same starting and ending points as their wall-mounted counterparts, by identifying an unmistakable visual artifact visible to both cameras (such as the moment a particular participant first touched a handheld), and reassigning the embedded timing data of the tripod camera file to match that of the wall-mounted camera file. For this purpose, the *Adobe Premiere* video editing application was used.

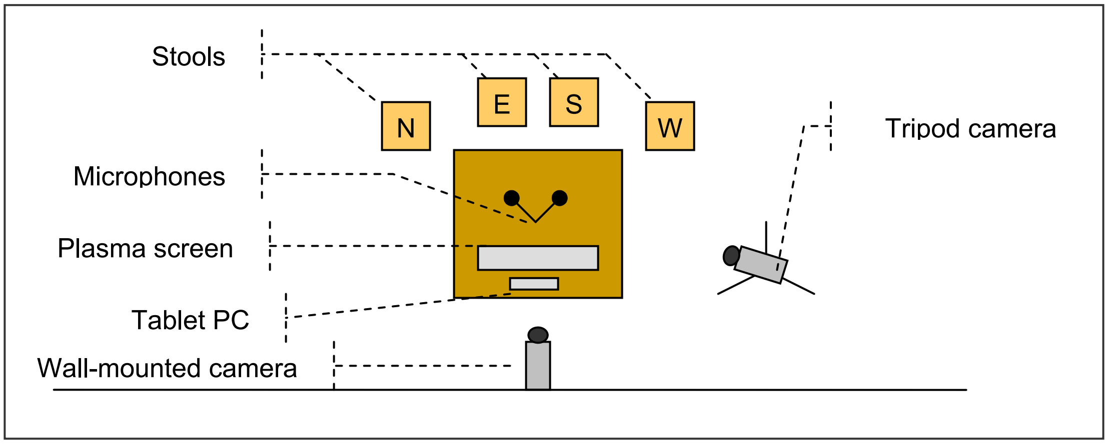


Figure 80. Top-down illustration of the audio-visual equipment setup in the lab used in this phase of the study.

The video data was primarily used to determine the moment-to-moment focus of attention for each participant. In order to capture this behavior, video transcription software was used (*InqScribe*). The different gaze targets (O-UI, Shared Display, Companions, and Other) were assigned to different keys, which, when pressed, would enter in the time code (hh:mm:ss;ff) and the target name into the transcript. The coding of gaze targets was performed by a single person (the author), who watched the video recordings of experimental sessions multiple times, concentrating on coding the gaze targets for a single participant on each viewing. Although another coder was not available to perform a check on the coding accuracy, periodic self-checks (i.e., *intra*-coder reliability checks) found that the coding was accurate to within 5 frames (i.e., given a 30 fps playback rate, 1/6th of a second). The transcription software provides a great deal of control over playback speed, allowing users to step frame-by-frame to get extremely accurate timing information. The 5-frame variation in coding accuracy is entirely the result of human judgment: for example, there is generally a small fraction of time between the lift of a participant's chin and the lift of a participant's eyes as they shift gaze from an O-UI to the Shared Display, and if the participant is blinking, the exact moment of gaze transition is open to interpretation. A difference of 5 frames (1/6th of a second) should be well above adequate for the purposes of this study, however.

The secondary function of the video data was to provide an audio record of participants' conversations. The author once again used the *InqScribe* transcription program to transcribe the audio portion of the video files. Periodic self-checks revealed a 20 frame, or about 2/3 of a second, range in accuracy on capturing the exact moment a speech utterance was produced. This figure should still be more than adequate for this analysis.

6.6.5.2 Software Logs

The *MUSHI-Lignancy* application was designed to output text-files to log each session of use. Two main types of data, regular state updates and asynchronous user-generated events, were encoded into two different file types: a system log, and user logs, respectively. In each cycle of the "game loop," after the simulation computed state updates for each of the cellular automata, these state updates were appended to the end of the system log. The data includes all parameters needed to later re-create the step-by-step unfolding of the simulation, should such a recreation be necessary. Section 5.1.1.2 MUSHI-Lignancy as a Cellular Automat describes these parameters, and they are summarized in Table 4. Also included are the current positions of the users' "incision" rectangles at the time of the state update.

When a user logs into the *MUSHI-Lignancy* simulation, a unique identifier is passed from the handheld device to the server (in this case, an ever-increasing unique ordinal number, combined with the letter "direction" of the user: N, S, E, or W). A new log file is created for the user, and this is where each action taken by the user is recorded. Every time a user moves his or her "incision" rectangle, this move is recorded in his or her log. Each time a user performs surgery, the action, and its consequences (in terms of the amount of damage done to whatever normal and cancer cells were affected) are likewise entered into the user's log. The timestamps were all generated within the simulation, so determining when the user actions occurred *vis-à-vis* the different state updates in the state update log is quite simple. There were a few occasions where, due to user error or a glitch with the handheld wifi (which can be finicky), a user was effectively logged out of the simulation and had to log back in, thus earning a new unique identifier, but these instances were easily rectifiable prior to data analysis.

To establish which log files corresponded with which video files, the author created a small program to run through the log files and tentatively assign the logs' unique game numbers to experimental session numbers using the date and timestamps in the logs. The author then engaged in a manual rectification process to handle the border cases, sometimes referring back to the video files to check for logged simulation events that would be visible on the video to establish the offset of the log's timing data from the video's timing data. Once the games played in each session were firmly established, the author wrote another small program to comb through the system and user logs and extract the information needed for computing the engagement and task performance measures.

6.6.5.3 Questionnaire

The questionnaire was primarily designed to probe each participant's awareness of his or her group context, and his or her opinions of same (see Appendix B). The first two pages are devoted to eliciting feedback on the collaboration, using different written strategies like Likert questions (questions 1-9, and 12 and 13), short answer (questions 10 and 11) and a single multiple-choice question (question 14).

Likert questions, which force a participant to select an item from a pre-determined scale, are good at eliciting opinion-oriented data, which museum visitors are better at providing than, say, quantitative estimates of their own behavior, like time spent on a given task (Diamond, 1999). Sometimes museum visitors will select a middle, neutral value for Likert questions it more out of force of habit than because it represents their true opinion (Diamond, 1999). Instead of removing the neutral option as Diamond suggests, it was placed it in its own column at the end of the scale, so that selecting a neutral values is an intentional act on the part of the respondent. The Likert questions were designed to capture participant's opinions about their own contributions towards the group effort (questions 1, 2, 6 and 9) and about their group members' contributions (3, 4, 5, 7 and 8).

The next section of the questions attempted to get at visitors' opinions regarding the general difficulty and appropriateness of the *MUSHI-Lignancy* simulation as a museum exhibit (page 3). Seven Likert questions addressed issues like user-perceived challenge (questions 18 and 20), length of play (question 21), alternatives to collaborative play like solo play (question 15) and competitive play (question 16), and perceived entertainment value (question 17) and educational value (question 19). The remaining questions asked them to rate the appropriateness of *MUSHI-Lignancy* for the Exploratorium and elaborate in short answer form(questions 22 and 23), and to do the same for their opinion of the appropriateness of computer-based exhibits for a museum like the Exploratorium (questions 24 and 25).

The final page of questions was an attempt to assess any misconceptions visitors might have about the underlying behaviors of the *MUSHI-Lignancy* simulation after having used it (page 4). Time constraints and concerns about priming visitors to attend to certain elements of the simulation ruled out the use of a written pre-test to measure more traditional "content learning." Such difficulties are not new to those doing research in museum contexts; many museum professionals have had mixed experiences with what can be gleaned by comparing pre- and post-tests, and generally favor other methods (like dialogue analysis or content recognition post-tests) to understand visitor learning (Allen, 2002; Diamond, 1999). In a sense, the questions on page 4 are a form of content recognition, although they probe participants' abilities to recognize the underlying rules or emergent patterns present in the simulation, in lieu of using visual or terminological (i.e., word) recognition.

6.6.5.4 Interview

The interviews were conducted by the primary researcher at the end of the second condition of the sequence. The questions stared off essentially the same, from a printed list, and the interviewer would follow-up responses as seemed prudent. Every group was initially invited to share their impressions about the two different interface designs, from both an emotional ("which did you prefer") and utilitarian ("which was easier to use") perspective. Because the purpose of the interview was not to collect data to make comparisons across groups, the focus was not on duplicating the interview process for each group; but rather, on querying the group members about circumstances unique to their experience with both versions of the O-UI (e.g., "what were you trying to accomplish by doing X?"). This information was used primarily to inform future O-UI designs, and to help the researchers disambiguate strange patterns in the collected quantitative data (although this proved to be not needed).

CHAPTER 7

Results

This chapter provides an overview of the results obtained by each phase of the study. The formative study phase, Phase I, was used to refine the designs used in the experimental phase, Phase II. Owing to the qualitative and revisionist nature of Phase I, the results will be presented in a narrative case-study format. The purpose of including these case studies is to provide a more nuanced understanding of the design decisions that filtered into the O-UI design.

7.1 Phase I: Formative Study Results

Overall, visitors did not seem to have too many problems learning to operate the O-UIs to be used later in Phase II of this study. For Phase II to be a controlled experiment, the two different surgeon O-UIs needed to be equivalent in terms of their impact on the simulation. The first case study details the changes made to ensure that neither the “Simple” nor the “Complex” O-UI design provided an overwhelming and non-germane advantage in executing the joint task in Phase II. The second case study concerns the delivery of damage feedback. The damage feedback (and whether or not users attend to it) affects the quality of the learning and the collaboration users engage in, and the formative redesign process also shed some light on future directions for research.

7.1.1 Case Study: Balancing Impact of Surgery O-UIs on Simulation

The purpose of Phase II’s experiment was to study the impact of the different O-UI designs on collaborative task execution. For that reason, it was extremely important to ensure that neither O-UI have a *non-germane* advantage in task execution. A *germane* advantage would be one owing to the innate differences in O-UI complexity- for example, a more complex O-UI should by all rights provide more detailed output, and should allow users to be able to provide more nuanced input. If the “Complex” O-UI’s detailed output and nuanced input lead to a competitive advantage *vis-à-vis* the “Simple” O-UIs, this is a legitimate and germane tradeoff, and one that needs to be considered when designing O-UIs. The essential features of the task should not change between O-UI interfaces, however, and the “power” of each interface to affect the shared simulation should be roughly equivalent. An example of a *non-germane* advantage would be if, for example, one type of O-UI user would be able to perform drastically more surgery events than another, or perform surgery events that inflicted more “damage” on the targeted regions, thus providing an artificial advantage to that O-UI design.

7.1.1.1 Granularity of Surgery O-UIs

The two Surgery O-UIs need to be balanced in terms of how much of the simulated patient they are able to impact – if one O-UI was able to take out large swathes of tissue at a time, when the other was restricted to knocking out one cell at a time, this would result in very different outcomes. The most obvious design is to limit both to a “single-cell” granularity, allowing each to “operate” on only a single cell in the patient at a time, but this design also artificially weights the task execution in favor of the “Simple” O-UIs: “Simple” O-UI users need only to press a button to perform damage on the cell, whereas “Complex’ users would need to perform a comparatively much more complicated, and slower, stylus movement. Initial tests with this design (which were performed, it should be noted, with graduate students at a university who were adept at using handheld devices, and not actual museum visitors) bore out this imbalance: groups using the “Complex” interface almost uniformly lost their patients, whereas “Simple” O-UI groups almost uniformly succeeded.

The “Complex” users must have access to more than one simulation element at a time to demonstrate the value of a “Complex” O-UI, otherwise there is no room for the nuance in input that only “Complex” O-UIs can provide. For this reason, a 3x3 granularity was chosen, giving access to 9 cells at a time. But by upping the granularity of the “Complex” O-UI without also increasing the granularity of the “Simple” O-UI, though, a different imbalance was created. The “Simple” O-UI users had to make nine times as many moves to cover the same amount of territory on the Shared Display for every single move made by the “Complex” O-UI users. This imbalance reversed the performance seen in the trials prior: now the “Complex” O-UI users were handily beating back the cancer in less than a third of the time “Simple” O-UI users were taking.

Just raising “Simple” O-UIs to a 3x3 level of granularity wasn’t the full answer, either. By this point, the informal formative trials were being conducted on the floor of the Exploratorium, with actual visitors. Giving 3x3 granularity to both O-UIs allowed “Complex” users the ability to conduct surgery with nuance, so it was certainly an improvement over single-cell granularity. In this case, though, “Simple” O-UI users were forced to “knock out” nine cells at once, resulting in a lot of collateral damage amongst healthy cells. The results were once again one-sided: “Complex” users were succeeding, whereas “Simple” users were losing patients, this time to collateral damage, not cancer. To establish equivalency between the O-UIs, in terms of their relative ability to impact the simulation, the amount of damage meted out by each interface needed to be balanced.

7.1.1.2 Damage Balancing of Surgery O-UIs

The paradigm used for the damage dealt by the “Complex” O-UIs is this: if a cell is circled in its entirety by the stylus, it is considered to be “excised,” and the cell at that location in the simulation was “killed,” meaning that it is stripped of its properties (its health and cancer/normal status) and that grid location is available to be populated by neighboring dividing cells (or metastasizing cancer cells elsewhere). Likewise, if a line is drawn completely across a blood vessel, the vessel is considered to be "severed," and its blood taken away and redistributed to the body (the speed at which a vessel regrows is determined by the blood flow of neighboring vessels). Other damage events (incompletely drawn circles or lines, or marks that "nick" or do not completely intersect a simulation element) are processed to the degree to which they overlap the simulation elements they intersect. These actions are simplistic analogues of what occurs in a real human body, and demonstrate how the nuance of a more "Complex" O-UI can implicitly communicate lessons about the subject matter. (This is the main reason educators would want to risk employing more complicated O-UIs). For this reason, it seemed more sensible to try to adjust the damage done by the "Simple" O-UI to bring it into equivalence with the damage dealt by "Complex" O-UIs than to engage in the reverse.

Establishing damage equivalence is a tricky thing, however – technically, if a "Complex" O-UI user were to circle *all* of the cells on their interface, they would do about the same amount of damage as the "Simple" O-UI users did by pressing their button. In practice, though, museum visitors using the "Complex" O-UI almost never tried this approach. Even when all of the nine cells on the "Complex" O-UI screen were cancerous, visitors in the formative trials might circle two or three or even four at a time, but never the whole screen. As a result, visitors using the "Simple" O-UI were killing their simulated patients, and killing them quickly, as compared to the "Complex" users. One could argue that this imbalance is a germane one – the "Complex" users were merely taking advantage of the nuance allowed to them by their O-UI. Unfortunately, one could also argue that by expanding the granularity for the "Simple" O-UIs from single-cell granularity to a 3x3 grid, a straw-man comparison was set up where participants in the "Simple" O-UI condition were sure to fail at the joint task.

The resolution to this conflict came from a careful observation of the usage patterns of the "Complex" O-UI users. As mentioned above, even when operating on a screen full of cancer cells, they would almost never try to circle all or even a majority of the nine cells present on the O-UI screen at one time. If anything, they tended towards circling the centermost cells on the "Complex" O-UI. If there were more cancerous cells above the center cell, they would move their "incision" rectangle upwards, centering their new targets in the O-UI, and circling them there. This strong centric preference was either an artifact relating to the physical dimensions and form-factor of the handheld device itself, or was borne out of a desire to see the full immediate context of the cell being operated on. Those visitors who could articulate the reasons for their centric tendencies had various responses that would conform to either hypothesis: "[it] felt more comfortable," "it seemed more natural to do it that way," "I like seeing all the other [cells]," or "I like it better in the middle." Regardless of the reasons behind the centric tendency, if this damage pattern was replicated for use by the "Simple" O-UI, the damage meted out to the patient would be balanced.

The new damage model designed for the "Simple" O-UI still worked at a 3x3 level of granularity, but rather than obliterating all 9 cells in the incision rectangle when the center button was pressed, each center-button press corresponded to an imaginary circle drawn around the centermost 5 cells – the equivalent of a "Complex" user circling those five cells with a stylus. (If damage had been limited to just the centermost cell, the effect would have been to just re-implement the single-cell level of granularity that was previously found to be problematic). After testing this version with a few visitors, it became apparent that the level of damage was still too high, *vis-à-vis* the "Complex" O-UI, and so the size of the imaginary circle was contracted to encompass the blood vessels surrounding the centermost cell, and just brush against the four neighboring cells. With this version, visitors were finishing each "game" with the same amount of success, and taking roughly the same amount of time to complete each "game." These impressions formed during the formative study were confirmed in the experimental study in Phase II: the averages were roughly between 2 and three minutes in length per game, and around an 80% success rate for both (see Table 21).

Table 21. Average game length and success rate in Phase II (the experimental study) showing a rough equivalence between the "Simple" and "Complex" O-UIs. (Notice that each condition has quite a bit of variance on both measures).

	"Simple" O-UI Groups n = 11	**"Complex" O-UI Groups n = 11**
Average game length (s)	*M* = 126.89 s *SD* = 194.37	*M* = 165.29 s *SD* = 185.91
Outcome: Percent successful games	*M* = 85.35% *SD* = 0.25	*M* = 74.24% *SD* = 0.38

7.1.2 Case Study: Damage Feedback

The nominal educational goal for the simulation was, broadly, to help participants understand the systemic effects of different cancer treatment options, and specifically, to help them understand the tradeoffs inherent in pursuing more or less aggressive treatments (the more aggressive the treatment, the larger the potential for collateral damage and even death, but the less aggressive the treatment, the chance is greater that the cancer will grow and spread). This study, while not directly designed to measure educational impact, was formed as part of a larger DBR agenda, and so the software needed to be designed as if it were to be used for its intended purpose. For that reason, it was important to ensure that the participants were made aware of how their treatment actions impact the simulation – without making the cause-effect relationship clear, the overarching educational goals would be unobtainable.

The simulation itself, being a dynamic entity, does respond to participants' treatment actions. The original intent behind employing an ongoing, dynamic simulation was to allow visitors to engage in trial-and-error experimentation, and to allow them to hone the observational skills needed to perceive change in a complex system. Unfortunately, the first formative trials revealed that visitors were having a hard time differentiating changes they had wrought in the simulation from changes their companions had engendered or from emergent changes that were unrelated to their actions. Part of the issue seemed to be related to speed: the nature of the simulation introduced an inherent delay from, say, slicing through a blood vessel and then seeing adjacent cells "wither on the vine." This delay caused problems with both individual action attribution (e.g., “Did *I* do that?”) and companion action attribution (e.g., “Did *you* do that?”) which made effective collaboration impossible. In the first few trials, visitors gamely enacted their treatment options (e.g., surgery, chemotherapy, radiation, and proton beam radiation) but reported that they did so without any real understanding, or without any real attempt to understand how their actions, and those of their partners, were affecting the larger context of the simulation.

Excessive latency, and specifically of the latency of shared visual information, has been shown to negatively impact small-group collaboration where a shared visualization is used (Gergle, Kraut, & Fussell, 2006). It had been hoped that visitors would show more patience in waiting to observe results, but the delays involved (on the order of

3 to 10 seconds) were above the delay shown to negatively impact lower-level action coordination (around 900 ms) (Gergle et al., 2006). Visitors simply didn’t take the time to observe the impact of their (or their companions’) actions before initiating new actions and, much like the surface of a pool that is having many pebbles thrown into it, it soon became impossible for a viewer to distinguish the origin of the different “ripples” in the simulation. Even worse, sometimes visitors would repeat the same action again and again until a result, any result, became apparent, thus causing much more collateral damage to the simulated patient (and often bringing about its demise, especially where the radioactive treatment options were concerned).

The first attempt to address this problem of providing adequate feedback to users was enacted simply on the floor – after the first two sessions when the problem became apparent, some of the simulation parameters were tweaked to provide a much more rapid response on the part of the cells and blood vessels to damage events. For any complex system (the wing design of stealth fighter planes comes to mind) there is an inherent tradeoff between sensitivity and stability, however. Increasing the sensitivity of the simulation components to damage did not work as intended – the increased sensitivity just decreased the stability of the simulation, and anything more than the slightest amount of treatment could bring about sudden death, as if the patient being simulated was a hemophiliac or had a compromised immune system. Treating such patients was not the educational purpose of the exhibit, however, and more opportunities for trial and error were needed to make it an engaging (rather than frustrating) experience.

The second approach was to make a player’s impact more explicit by directly stating, on the Shared Display, how much damage was just done immediately after an action was committed, even if the full impact on the simulation would not be visible for a few moments. Initially this information was presented in the form of small numbers on a 100 point scale (e.g., -10, which is minor damage, and -100, which indicates that a cell will die) which would appear next to the damaged simulation elements and fade away after 3 seconds. Younger visitors complained that the numbers lingered too long, and older visitors complained that the numbers didn’t linger long enough, and they were too small to see well. A majority of the visitors, young and old, reported that the numbers just made the screen look cluttered, with so much already going on in the dynamic simulation, and that they ignored the numbers anyhow. Indeed, incorporating the damage feedback in the midst of the dynamic display probably was pushing against visitors' innate visuospatial cognition limits, overloading their visual working memory, since they had to sort through many stimuli when looking at even a small region of the display.

The third approach was to keep the damage information on the shared screen (so that all players were privy to the impacts of their partners' actions) but off to one side, away from the main body of the simulation display. To effect this, the simulation was shrunk down a bit and moved to one side, and on a column with a plain background, each player's name and damage tallies were displayed. After each action, the tallies (the amount of damage done to normal and cancer cells, and the number of each that were killed outright) were updated. Visitors completely ignored this source of information, however – many expressed surprise when they were asked how often they had attended to it during their play, as it had gone entirely unnoticed. The relatively static (and numerical) display simply failed when placed in competition with the dynamic simulation display for visitors' visual attention.

A fourth approach to deliver feedback information to the players in a manner that would attract their attention was to provide the information directly to the players via their O-UIs. After receiving treatment action input from a player, the simulation would generate a feedback message that would immediately be delivered to the player's O-UI, detailing the amount of damage dealt to the cancerous and normal cells. Programming convenience and standard practices suggested that this message should be delivered in the form of a pop-up message, but there was an additional reason for using a pop-up message (complete with an "OK" button that needed to be pressed before it would disappear). If one takes a Learner-Centered Design approach, one must sometimes consider complicating tasks in order to force (strongly encourage?) learners to attend to certain features of the learning environment. In this case, the thought was that by forcing visitors to tap the "OK" button, they would also be forced to attend to the feedback message, even if only briefly. Unfortunately, this rationale was an abject failure.

Virtually all of the visitors tested with the popup mechanism – around 10 in all – reported that not only did they fail to read the content, they "hated" the pop-ups, finding them "annoying," "in my face," and many "wish[ed] they would just go away." The popup content failed to spark any sort of reflection, and the pop-ups were a hindrance to be disposed of as quickly as possible before getting back to the main business of interacting with the simulation. They also posed particular problems for "Simple" O-UI users, who didn't think to visually attend to their O-UIs very often, and would get "stuck" before realizing that the reason their incision rectangle was ignoring input was that there was a popup requiring their attention on their O-UI. Even though the arrival of the popup was audibly heralded by a beep from the O-UI, visitors would almost never look down to the O-UI screen until they noticed a freeze of their actions on the Shared Display, which inevitably caused frustration.

The fifth (and final) approach was deceptively simple. The different methods for using O-UIs as a channel to provide damage feedback were unpopular among virtually all of the users, and the original intent was in any case to make each player's actions apparent to all other players so as to better aid collaboration. Locating the feedback on the Shared Display, and tying the feedback to the location where visitors were conducting their actions (i.e., not asking them to divide their attention between two locations on the Shared

Displays) seemed prudent, but visitors had complained about the visual clutter of the numbers. Rather than provide detailed information on the damage done to simulation elements in numerical format, then, the new approach was to simply flag *which* elements had been damaged.

The method used was extremely simple – the saturation of the colors of the damaged element was dramatically decreased and the luminosity was increased for a brief moment, so the element would appear to "flash." Additionally, the element would be "flopped" on its longest axis (recall 2-dimensional sprites were being used, so this was a simple matter of changing the mapping of source image coordinates to the on-screen sprite coordinates) so that it appeared to "wiggle." When combined with the on-screen representations of user actions (circles or lines indicating surgical actions, or colored triangles or columns indicating the path of beams) visitors reported having little to no trouble attributing the damage events to the originating player. The simple flash-and-wiggle was dynamic enough to attract their attention, but simple enough that it didn't overwhelm their ability to process the feedback. This simple feedback event was enough to bridge the latency gap that was originally so problematic. It had the additional effect of actually inducing many visitors to pause and watch what would happen "next" to the elements that had flashed-and-wiggled, which was the original hope – that visitors would engage in detailed observations of the impact of treatment actions.

This case study is interesting in that it suggests a direction for future CSCL/CSCW research. In collaborative software-based activities where visual latency is a problem, simple visual *markers* or *placeholders* may be used to maintain the attention of users so that when the desired data actually arrives, the user is prepared to make sense of it. It is also worth testing whether the use of markers/placeholders might actually improve the collaborative performance, perhaps through some sort of priming mechanism that implicitly helps users prepare for the information to come. In educational terms, these placeholders might operate as "opportunities for reflection," which are valuable to learning processes, but altogether too hard to come by in many learning environments.

7.2 Phase II: Impact of O-UI Complexity

Phase II of this experiment set out to contrast the individual and group behaviors of visitors who used a "Simple" Opportunistic-UI as an interface to a software-based exhibit against those who used a "Complex" Opportunistic -UI. There is reason to believe that visitors may have a difficult time dividing their attention between a "Complex" O-UI and other elements of the collaborative context (e.g., the Shared Display, or their companions). There is reason to suspect that an inability to effectively divide attention will adversely impact collaborative learning processes. In Section 6.6, the physical context for the Phase II experiment (Section 6.6.1 Physical Context6.6.4 Procedure), the means of recruiting participants (Section 6.6.2 Recruiting Participants), the different conditions (Section 6.6.3 Conditions), the procedure (Section 6.6.4), and the instruments used (Section 6.6.5 Instruments) were all described. This section will report on the results of that experiment.

7.2.1 Establishing the Existence of the Heads-Down Effect

The first research question concerns the "heads-down phenomenon." Does its existence indeed correlate with the "complexity" of the O-UI interface? The heads-down phenomenon, as described by museum professionals, is characterized by both visual attention management behaviors (the tendency to stare fixedly at a handheld device for long periods) and by the consequences of those behaviors (a lack of awareness of the surrounding context). Specific hypotheses for each of these will be discussed in turn.

The primary evidence used to study visual attention behaviors comes from studying where participants direct their gazes. If the heads-down phenomenon occurs, one would expect to see the following patterns:

H1: *Individual participants devote a larger* **Proportion** *of their time to viewing the O-UI than the Shared Display (or other gaze targets like companions).*

H2: *Individual participants would spend larger unbroken spans of time gazing at the O-UI, in other words, the average gaze* **Duration** *devoted to the O-UI would be larger than the gaze duration devoted to other gaze targets.*

H3: *Individual participants would have a much lower* **Frequency** *of shifting their gaze from target to target.*

If the heads-down phenomenon occurs, it should also impact some behaviors at the group level of granularity. With respect to *Group Division of Attention* measures, if group members be suffering from the heads-down effect, they should also exhibit problems with joint attention management, exhibited by a lower gaze *Synchronicity Degree*:

H4: *Groups should demonstrate a lower gaze* **Synchronicity Degree**

Gaze-related patterns of behavior are just the visible symptom, however; of the heads-down phenomenon. As described by museum practitioners, the secondary effect of the phenomenon is a marked lack of attention to other elements in the visitors' environment, which in this case includes both the visitor's companions and the Shared Display. If participants exhibit the heads-down phenomenon, one would also expect to see the following secondary symptoms that indicate a lack of awareness of companions:

H5: *Individual participants would report lower levels of* **Awareness** *of their partners' actions.*

H6: *Individual participants would engage in less* **Conversation** *with their partners.*

Participants must refer to the Shared Display in order to decide where they will move next, and so to the extent that the heads-down effect occurs, one would expect to see:

H7: *Individual participants would commit fewer* **Moves** *per unit time*

Assuming that the three secondary symptoms listed above are in fact caused by poor visual attention management behaviors, we would expect to see the following correlations:

H8: **a:** *Self-reported* **Awareness** *would correlate* **positively** *with more* **"heads-up"** *visual attention behaviors*

b: *Self-reported* **Awareness** *would correlate* **negatively** *with more* **"heads-down"** *visual attention behaviors*

H9: **a:** *Amount of* **Conversation** *would correlate* **positively** *with more* **"heads-up"** *visual attention behaviors*

b: *Amount of* **Conversation** *would correlate* **negatively** *with more* **"heads-down"** *visual attention behaviors*

H10: **a:** *Frequency of* **Moves** *would correlate* **positively** *with more* **"heads-up"** *visual attention behaviors*

b: *Frequency of* **Moves** *would correlate* **negatively** *with more* **"heads-down"** *visual attention behaviors*

The first four predictions and can be answered by an analysis of the coded video data of participants' gaze targets. It will be the subject of the next section. The secondary symptoms of the heads-down effect will be described in 7.2.1.2 The Heads-Down Effect and Awareness of Shared Context, and the correlations between the two with be detailed in 7.2.1.3 Demonstrating Relationship of Heads-Down Visual Attention Behaviors and Awareness

7.2.1.1 The Heads-Down Effect and Visual Attention Management

The coded video data should reveal if there is indeed an emergence of an externally-observable heads-down phenomenon as the O-UI "complexity" is increased. It should be pointed out that this study made use of an incomplete repeated-measures design with rotation (meaning that each group would be sequentially exposed to first one condition, and then the other, in an order that changed between groups). For that reason, a repeated-measures (also known as a direct-difference, within-subjects, or paired) t-test was used for the comparisons presented in Table 22. Each of the measures used to assess the *Individual Division of Attention* will be discussed in turn.

It is obvious that participants in the "Simple" and "Complex" conditions, while devoting about the same *Proportion* of time to gazing at their companions (around 2% of the play time), have diametrically opposed viewing habits regarding O-UIs (and, consequently, the Shared Display). Participants in the "Simple" condition spend only around 14% of their time gazing at the O-UIs, and the vast majority of their time, 83%, gazing at the Shared Display. On the other hand, participants in the "Complex" condition spend only 33% of their time looking at the Shared Display, reserving twice that amount, 65%, for their O-UIs. The differences between the conditions are strongly significant, and suggest that O-UI complexity does indeed promote a heads-down phenomenon, following prediction **H1**.

Table 22. Results of the analysis of the coded video data-based *Individual Division of Attention* measures used to establish visual attention management behaviors relevant to the heads-down effect. (See Figure 6 and Figure 7 for visualizations).

Dependent Variables		"Simple" O-UI Participants $n = 30$	"Complex" O-UI Participants $n = 30$	Significance (1-tail, paired t test) $n = 60$, $df = 29$
		Conditions		
Proportion (% of time devoted to gaze target)	O-UI	$M = 0.14$ $SD = 0.08$	$M = 0.65$ $SD = 0.15$	**$t(29) = 15.45$ $p < 0.0000001$**
	Shar ed Display	$M = 0.83$ $SD = 0.09$	$M = 0.33$ $SD = 0.15$	**$t(29) = 14.80$ $p < 0.0000001$**
	Othe r players	$M = 0.02$ $SD = 0.008$	$M = 0.02$ $SD = 0.07$	none
Duration (average gaze duration, in seconds)	O-UI	$M = 2.55$ $SD = 1.47$	$M = 12.52$ $SD = 32.45$	**$t(29) = 1.81$ $p < 0.040$**
	Shar ed Display	$M = 20.56$ $SD = 24.82$	$M = 2.81$ $SD = 1.00$	**$t(29) = 4.25$ $p < 0.0001$**
	Othe r players	$M = 0.97$ SD = 0.78	$M = 1.06$ $SD = 0.69$	none
Frequency (gaze shifts per minute)		$M = 9.50$ $SD = 4.47$	$M = 14.99$ $SD = 5.87$	**$t(29) = 6.15$ $p < 0.000001$ two-tailed**

Gaze *Proportion* doesn't tell us the whole story, however: it could be the case that participants in the "Complex" O-UI condition, while on the balance spending more time looking at their O-UIs, still adequately monitor the goings-on displayed on the Shared Display. This supposition may be supported by the fact that "Complex" users exhibit significantly more *Frequent* gaze shifts, about 15 per minute in contrast to the roughly 10 gaze shifts per minute exhibited by the "Simple" participants (see Table 22). This countermands prediction **H3**, which suggests that if players are suffering from the heads-down effect, they will shift their gaze less *Frequently*. If the "Complex" participants are in fact using their more *Frequent* gaze shifts to monitor the Shared Display, one should expect the average *Duration* of the gazes they devote to the two different gaze targets to closely mirror the 65/33 ratio between the O-UI/Shared Display *Proportions* discussed above. The reasoning is as follows: adequate monitoring requires regular "check ins" with the Shared Display, as longer *Durations* devoted to the O-UI increase the chance that participants are missing important information visualized on the Shared Display. The only way to maximize the monitoring would be to limit the average *Duration* of O-UI gazes such that the ratio of their length to the *Duration* devoted to the Shared Display is 65/33. This is not the case, however.

Although participants in the "Complex" condition do indeed shift their gazes between the two displays significantly more than participants in the "Simple" condition, it seems that the "Complex" condition participants still stare at the O-UIs for significantly longer lengths of time, bearing out prediction

H2 (see Figure 7). Their gaze *Duration* when looking at their O-UIs averages 12.52 uninterrupted seconds (see Table 22), more than *four* times as long as they tend to look at the Shared Display (2.81 s), a far cry form the roughly two-to-one ratio that the 65/33 *Proportion* data would suggest. These numbers reflect a practice, observed by this researcher, where the majority of the participants in the O-UI condition tend to spend very long *Durations* looking exclusively at the O-UI, bookended with more rapid fusillades of attention-shifting between the O-UI and the Shared Display. The example illustrated in Figure 81 highlights this behavior.

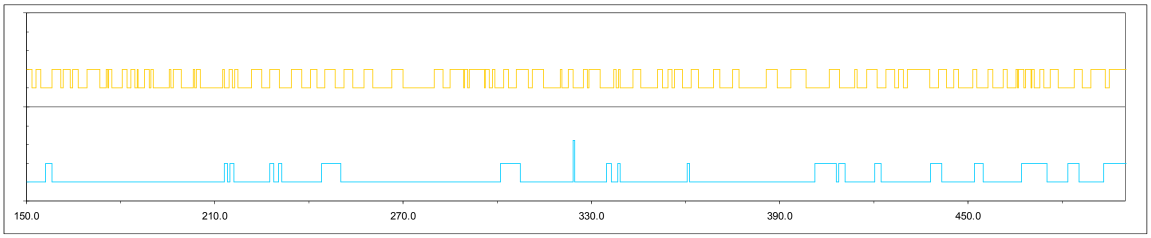


Figure 81. Visualizations of two "Complex" participants' moment-by-moment gaze targets, on the y-axis, with time (in seconds) on the x-axis. The gaze targets are in ascending order on the y-axis: O-UI, Shared Display, and companion(s). The top, yellow line is very atypical; most participants exhibited viewing behavior similar to the lower blue line.

The participant on the top of Figure 81, represented by the yellow trace, quite successfully divides his attention between the O-UI and the Shared Display. This person is clearly not suffering from any heads-down problems, which is reflected in the participant's average gaze *Durations*: 3.09s for the O-UI, and 2.11s for the Shared Display. The second participant, however, represented by the blue line on the bottom, clearly *is* exhibiting the heads-down phenomenon. He spends most of his time looking at the O-UI for long *Durations*, interrupted by small episodes where he looks at the Shared Display and back at the O-UI. This, too, is reflected in the gaze *Durations*: 13.05s for the O-UI, and 2.65s for the Shared Display. Given that these latter figures (13.05s versus 2.65s) more closely resemble the overall averages for the "Complex" condition participants (12.52s versus 2.81s), one can infer that this heads-down pattern in how participants divide their visual attention is fairly common amongst "Complex" O-UI users. So it seems that while prediction **H3** was countermanded by the evidence, it was because of a faulty assumption that the instances when participants shifted their gaze would be distributed relatively evenly across the session. Without an even distribution of gaze shifts, measuring the gaze shift frequency indicates little about the quality of a participant's monitoring practices.

The final prediction, which was related to *group* visual attention management, was H4: Groups should demonstrate a lower gaze *Synchronicity Degree.* This prediction held very true: the average *Degree of Synchronicity* for "Simple" groups was around 0.71, compared to 0.16 for "Complex" groups (see Table 23, and Figure 8). To the extent that the gaze *Synchronicity Degree* is an indicator of joint attention management, it seems to be showing that "Complex" groups do not engage in much joint attention management.

Table 23. Results of the analysis of the coded video data-based *Group Division of Attention* measure, *Synchronicity Degree*, relevant to the heads-down effect. (See Figure 8 for a visualization)

	Conditions		
Dependent Variables	**"Simple" O-UI Groups** *n* = 9	**"Complex" O-UI Groups** *n* = 9	Significance (1-tail, paired t test) *n* = 18, *df* = 8
Gaze Synchronicity Degree	***M* = 0.71** ***SD* = 0.19**	***M* = 0.16** ***SD* = 0.10**	$t(8) = 15.75$ **$p < 3E\text{-}16$**

Taken together, the results show that the "Complex" participants do demonstrate visual attention management behaviors that are very much in line with what the heads-down phenomenon would predict. Despite the hoped-for potential of the dynamic graphics of the Shared Display or the actions of companions to be able to attract participants' visual attention, supplying visitors with graphically "complex" O-UIs does produce visual attention behaviors consistent with the heads-down phenomenon. The next step is to examine whether or not the deleterious secondary symptoms occur.

7.2.1.2 The Heads-Down Effect and Awareness of Shared Context

Heads-down visual attention behaviors should result in decreased awareness of the other elements of the shared context. The most direct, although not necessarily reliable, way to assess awareness is to ask the participants directly. After the first condition, participants were given a questionnaire which asked them to rate the awareness they felt they had of their partners' actions (Q7), and their group's performance (Q9) on a 5-point Likert scale. The self-reported measures are of little help, however – the results are inconclusive for **H5**. Neither of the two Likert questions showed any significant trend (see Table 24), Both "Simple" and "Complex" participants responded very similarly to Q9, which queried group awareness, and both reported slightly lower partner awareness (Q7). It is interesting to note, though, that the drop-off for "Complex" participants was significant, $t(37) = 3.05$, $p < 0.0066$ (two-tailed). For some reason, "complex' users reported more awareness of the group than they did of their partners – a point which will be returned to later.

Table 24. Results of the analysis of the questionnaire data-based *Individual Division of Attention* measure. Because the questionnaire was administered after the first condition in the series, there was only one reply from each participant, so an independent t test was used. (See Figure 9 for a visualization).

	Conditions		
Dependent Variables	**"Simple" O-UI Participants** ***n* = 18**	**"Complex" O-UI Participants** ***n* = 21**	**Significance (1-tail t test)** ***n* = 39** ***df* = 37**
Awareness average of self-reported Likert measure: *Q7: "I was aware of how well my partners were doing at all times"* where 5 = strongly agree	*M* = 3.11 *SD* = 1.49	*M* = 2.71 *SD* = 1.15	none
Awareness average of self-reported Likert measure: *Q9: "I was aware of how well our group was doing at all times,"* where 5 = strongly agree	*M* = 3.72 *SD* = 1.23	*M* = 3.76 *SD* = 1.22	none

Rather than relying solely on participant's recall of their awareness, other measures can be used to determine the degree of awareness exhibited by participants. One is to measure the frequency of *Conversation* – it can be generally assumed that if a person is speaking to another person, the converser is paying some modicum of attention to the conversee, and at the very least is showing some awareness of their conversational target's existence (chatty airplane seatmates notwithstanding). Thus, **H6** predicted that "Complex" participants will show lower *Conversation* frequency, which is measured by counting the number of distinct utterances made by each participant and dividing by session length (in this work an utterance is considered to be an uninterrupted speech turn, because sentences can be hard to distinguish). The other indirect indicator of participant awareness of the shared context is to measure the frequency of *Moves* made, since making moves requires that participants attend to the Shared Display. **H7** thus predicts that "Complex" participants will show lower *Move* frequencies. Examining the data in Table 25, it is clear that participants in the "Complex" condition are significantly less likely to make *Conversational* utterances and to *Perform Moves* than "Simple" participants, confirming both **H6** and **H7**.

Table 25. Results of the analysis of the *Individual Engagement* measures relevant to predictions **H6** and **H7**, *Conversation* frequency and *Move* frequency, respectively. (See Figure 12 and Figure 14 for visualizations).

	Conditions		
Dependent Variables	**"Simple" O-UI Participants** *n* = 31	**"Complex" O-UI Participants** *n* = 31	**Significance** (1-tail paired t test) *n* = 62, *df* = 30
Conversation (utterances per minute)	*M* = 6.07 *SD* = 2.73	*M* = 4.51 *SD* = 2.62	$t(30) = 4.32$ $\mathbf{p < 0.00008}$
Participation: Moves (moves per minute)	*M* = 57.23 *SD* = 26.99	*M* = 21.23 *SD* = 8.89	$t(30) = 7.96$ $\mathbf{p < 0.000000003}$

7.2.1.3 Demonstrating Relationship of Heads-Down Visual Attention Behaviors and Awareness

The prior Section, 7.2.1.2 The Heads-Down Effect and Awareness of Shared Context, showed that at least two of the predicted secondary symptoms (lower **H6** *Conversation* and **H7** *Move* frequency) were indeed present, but the true test is if these secondary symptoms correlate with the heads-down visual attention behaviors. Even though no significant differences were noted for **H5** between conditions, self-reported *Awareness*, in Section 7.2.1.2, the self-reported measures still correlate more or less as predicted by **H8a** and **H8b** (see Table 26).

In particular, **H8a**, which predicted that awareness would correlate positively with the Proportion and Duration paid to the Shared Display, significantly held for "Simple" participants on Q7, which rates awareness of partners. "Complex" participants trended as predicted by **H8a** on Q7, although the correlations were not statistically significant. On Q9 (which rated group awareness) both "Simple" and "Complex" participants trended as **H8a** predicted, although the only significant correlation was for "Complex" participants on the Shared Display *Proportion*. With respect to **H8b**, which predicted that *Awareness* would correlate negatively with heads-down attention behaviors like the *Proportion* and *Duration* of attention paid to the O-UI; it held significantly only for "Complex" participants on Q9, which assesses group awareness. "Simple" participants trended as **H8b** predicted, however, on the *Proportion* correlations. It can be said that overall, **H8a** and **H8b** were moderately well supported by the evidence. Again, it is interesting to note that the correlations for "Simple" participants were significant with respect to the question that addressed awareness of partners, whereas the correlations were significant for "Complex" participants only on the question that addressed group awareness. This will be referenced later.

Table 26. Correlations between the gaze-based *Individual Division of Attention* measures and two of the self-reported *Awareness* measures. (See Figure 10 and Figure 11 for visualizations).

		Gaze Proportion O-UI	Gaze Proportion Shared Display	Gaze Duration O-UI	Gaze Duration Shared Display
Q7: "I was aware of how well my partners were doing at all times"	**"Simple"** (n = 18, df = 16)	-0.367 (not significant)	**0.459** (t(16) = 2.07, p **< 0.0276**)	0.098 (not significant)	**0.412** (t(16) = 1.81, p **< 0.045**)
	"Complex" (n = 21, df = 19)	-0.246 (not significant)	0.303 (not significant)	- 0.141 (not significant)	0.202 (not significant)
Q9: "I was aware of how well our group was doing at all times"	**"Simple"** (n = 18, df = 16)	-0.273 (not significant)	0.380 (not significant)	0.212 (not significant)	0.355 (not significant)
	"Complex" (n = 21, df = 19)	**- 0.451** (t(19)= 2.26, p **< 0.018**)	**0.453** (t(19) = 2.27, p **< 0.017**)	**- 0.429** (t(19) = 1.77, p **< 0.046**)	0.368 (not significant)

The correlation between the *Conversation* frequency and the visual attention management measures is only somewhat as **H9a** and **H9b** predict (see Table 27). **H9a** predicts that *Conversation* frequency will correlate positively with the visual attention *Proportion* and *Duration* devoted to the Shared Display. While there is a moderate (although not significant) positive correlation for "Simple" participants, there is pretty much no correlation for "Complex" players – the correlation coefficient approaches 0. With respect to **H9b**, which predicts that *Conversation* frequency will correlate negatively with the visual attention *Proportion* and *Duration* devoted to the O-UI, for "Simple" participants, the predicted trend only occurs for gaze Proportion, but not gaze Duration. "Complex" players have the converse pattern: the predicted trend only occurs for gaze Duration, but not got gaze Proportion. It seems to be the case, then, that while "Simple" players talk more when they look more at the Shared Display, "Complex" player conversations are not affected by attention to the Shared Display. "Complex" player conversations *are* negatively impinged upon, however, when participants spend long unbroken durations gazing at their O-UIs, suggesting that they may be "getting lost" in the heads-down effect. Of course, because none of these correlations is significant, **H9a** and **H9b** can claim only weak support.

Table 27. Correlations between *Individual Division of Attention* measures and *Conversation* frequency. (See Figure 13 for a visualization).

	Gaze Proportion O-UI	**Gaze Proportion Shared Display**	**Gaze Duration O-UI**	**Gaze Duration Shared Display**
"Simple" (*n* = 36, *df* = 34)	-0.199 (not significant)	0.158 (not significant)	0.096 (not significant)	0.066 (not significant)
"Complex" (*n* = 34, *df* = 32)	-0.018 (not significant)	-0.040 (not significant)	-0.235 (not significant)	0.017 (not significant)

The prediction of **H10a** was that the frequency of *Moves* would correlate positively with visual attention paid to the Shared Display, and this holds significantly true for both "Simple" and "Complex" participants for the gaze *Proportion* measure. Gaze *Duration* does not seem to be as important a factor, although it trends as predicted for "Simple" participants. The same is true for Duration and prediction **H10b** – it trends as predicted (the gaze Duration devoted to the O-UI correlates negatively with Move frequency) for both conditions, but is not significant for either. **H10b** does hold significantly, for both conditions, when it comes to gaze *Proportion*, however – there is a very strong negative correlation between devoting more visual attention to the O-UI and making *Moves*. Overall, there is moderately strong support for **H10a** and **H10b**.

Table 28. Correlations between *Individual Division of Attention* measures and *Move* frequency. (See Figure 15 for a visualization).

	Gaze Proportion O-UI	**Gaze Proportion Shared Display**	**Gaze Duration O-UI**	**Gaze Duration Shared Display**
"Simple" (*n* = 36, *df* = 34)	**-0.462 (t(34) = 3.03 p < 0.0046)**	**0.524 (t(34) = 3.59, p < 0.001)**	-0.121 (not significant)	0.146 (not significant)
"Complex" (*n* = 34, *df* = 32)	**-0.486 (t(32) = 3.14 p < 0.0036)**	**0.505 (t(32) = 3.31 p < 0.0023)**	-0.133 (not significant)	-0.184 (not significant)

The results for predictions **H8a**, **H8b**, **H9a**, **H9b**, **H10a**, and **H10b**, especially when considered together, provide convincing evidence that the visual attention behaviors described Section 7.2.1.1 The Heads-Down Effect and Visual Attention Management do in fact impact the secondary measures described in Section 7.2.1.2 The Heads-Down Effect and Awareness of Shared Context. It seems to be the case that when participants engage in more heads-up behaviors (regardless of the experimental condition) they have a greater awareness of the shared context, and that when participants engage in more heads-down behaviors (regardless of the experimental condition) they have less awareness of the shared context. Considering the weight of evidence, the variety of data sources it is derived from, and the internal consistencies of the data, it is safe to conclude that **Research Question 1 has been verified: increasing the user interface**

complexity of Opportunistic-UIs in a MMUI-centric museum exhibit does indeed produce the heads-down phenomenon.

7.2.2 Impact on Potential for Collaborative Learning

7.2.2.1 What Should Collaborative Learning Look Like in a Museum?

There is no one definition for what good support of collaborative learning is, just as there is not one definition for what collaborative learning itself is. Attempting to construct a definition that spoke to all of those researching collaborative learning, the best Pierre Dillenbourg was able to come up with, in an oft-cited quote, is that collaborative learning is “a situation in which two or more people attempt to learn something together” (Dillenbourg, 1999). Some have sought to avoid the problem of a universally applicable (and thus universally useless) definition of “collaboration” by breaking it down into a suite of other concepts, like “coordination” and “cooperation, but here is still a great deal of argument in the literature about how each of these differs from one another (Rummel & Spada, 2005).

Part of the reason for the difficulty of defining collaboration is that it is an emergent process, and as such it is very powerfully shaped by all aspects of the context in which it is developed. Are the learners debating ideas, or building an artifact? Are the learners at the same level of competency, or have varied degrees of experience with the matter at hand? Does the collaborative activity take place only once, or over an extended series of meetings? Do the learners know one another, or are they strangers? Collaborative learning *can* take place in all permutations of these contexts, but the measures taken to support collaborative learning under the different contexts will likely differ wildly. That said, there are two preconditions for collaborative learning that underlie nearly all contexts, the first being:

1. **Learners should have an awareness of joint goals**

Nearly every conception of collaborative learning assumes that it is best undertaken when the would-be learners understand the joint goals of the collaborative activity (Dillenbourg, 1999), with some using the degree of shared understanding,, i.e., “convergent conceptual change,” as *the* marker for the success of a collaborative learning activity (Roschelle, 1992). In a context like a museum exhibit, though, the length of use is unlikely to promote immediate and dramatic conceptual change in visitors – it has been demonstrated that visitors often take weeks, or months, to integrate museum experiences into their conceptions (Falk & Dierking, 2000; Falk et al., 2004). Besides, establishing shared understanding (e.g., accepting a common explanation for the role of CO2 in ozone depletion) is not the only possible joint goal for collaboration – sometimes it can be acquiring skills (e.g., learning to critique writing, as in O'Donnell & Dansereau, 1992), and other times it can be the joint accomplishment of a task (like writing a report, analyzing data, or solving problem). In the case of the exhibit used in this research, the joint goal is the cooperative accomplishment of a task: eliminating cancer cells from the simulated human tissue.

One advantage of computer-based museum exhibits is that visitors’ each and every interaction with the exhibit as they attempt to accomplish a task can be recorded. A fair amount can be surmised about a learner’s understanding of the shared task goal by looking at his or her actions, especially when those actions have direct, score-able impact on the eventual success or failure of an endeavor, as when visitors eliminate more cancer cells than healthy cells with their surgery actions. For that reason, this experiment is using the following measure[31] as an indicator of *individual*-level understanding of the joint goal:

M1: High *Individual Scores*

In a similar vein, if the groups are able to accomplish the aim of the activity (eliminating cancerous cells from the simulation), this betrays a certain amount of understanding of the joint goal of the activity, so the following measure can be used as an indicator of the *group's* level of understanding of the goal:

M2: Better group *Outcomes*

Finally, because this is a collaborative activity, a measure of whether or not the members of a group understood the collaborative goals of the activity can be measured by how successfully they were able to divide the task. In this activity, the optimal task division strategy was a very simple one: divide-and-conquer. The degree to which a group divided the task can be measured by the amount of the Shared Display that was occupied solely by a given individual, i.e., that individual's degree of territorial *Ownership*. Therefore, a measure that speaks to the group's understanding of the collaborative aspects of the shared task is this:

M3: High degree of *Ownership*

The second precondition common to nearly all forms of collaborative learning is that the learners actually interact with one another in meaningful ways:

2. Learners should interact with one another

In the case of this exhibit, the would-be learners interact mainly via *Conversation* (there are no chat programs or written artifacts to mediate interactions). The quantity of conversation is one issue (collaborative learning cannot take place if there is no interaction) but the *quality* of those interactions is perhaps more important. Collaborative learning cannot take place if the interactions are not *on task*, and so many educational researchers studying collaborative learning in classrooms will code conversational utterances as being on- or off-task to establish whether or not collaboration is likely to be occurring (Hertz-Lazarowitz, 1992). Researchers working in museums will often go beyond just measuring the amount and degree of on-task-ness of conversations and code for more specific conversational content, tuned to the exhibit(s) under study, that speak directly to the learning goals designers had in mind for the exhibit(s) (e.g., Allen, 2002; K. Crowley & Callanan, 1998; Leinhardt & Knutson, 2004; Scott. G. Paris & Hapgood, 2002). Because a major learning goal of this exhibit is for visitors to understand the underlying rules and processes of the simulation, examining the amount of tactical or strategic content in learner conversation

[31] Unlike the previous research question concerning the heads-down phenomenon, there is less clear evidence that the "Complex" condition will lead to poorer performance on measures of collaborative learning, so each of these items will be listed as measures, not as hypotheses, and tested using two-tailed statistical tests.

can indicate something about the degree to which visitors are trying to come to understand those rules. For these reasons, the following two measures are used to speak towards the quality of learner interactions:

M4: High proportion of *On-task Utterances*

M5: High proportion of *Tactical/Strategic Utterances*

Studying the content of learner conversations tells us something about the quality of collaborative learning likely to be going on (or not going on), but studying the *structure* of communication patterns also speaks to the likely success or failure of collaboration. In detailed ethnographic studies of collaborating groups in a classroom, the degree of *reciprocity* of communication was found to be a very strong predictor of whether or not the collaborative endeavor would succeed or fail (Barron, 2000, 2003). It makes a certain amount of intuitive sense that groups which respond to one another's remarks, whether or not they were building on those prior remarks or attempting to refute them, end up doing better on accomplishing joint goals as compared to groups that continually put forth stand-alone utterances. For this reason, the proportion of utterances made in response to other utterances, the proportion of *Interactional* remarks, is also deemed to be an indicator of likely success of collaborative learning:

M6: High proportion of *Interactional Utterances*

Finally, we come to one remaining category that is accepted by many collaborative learning researchers as a precondition for "good" collaborative learning: *equity*. Educational researchers have found that when some learners do not participate as much as others in a joint learning activity, a phenomena variously dubbed at "social loafing" (Slavin, 1992) or "the bystander effect" (Hudson & Bruckman, 2004), they miss out on more than just the collaborative learning benefits of the activity. A learner's level of participation in a joint learning activity directly correlates with individual learning gains (Cohen, 1994; Kapur & Kinzer, 2007; O'Donnell & Dansereau, 1992; Schellens et al., 2005). Perhaps for this reason, several CSCL researchers use measures of participation equity to judge whether a collaborative system is successful or not (Kapur & Kinzer, 2007; Schwartz, 1999), although the value of measuring participation equity also comes up in CSCW literature (DiMicco et al., 2007; Morris et al., 2004; Rodden et al., 2003). Given that museums have the additional responsibility to serve a wide variety of visitors, supporting equity is especially important. The exhibits that only appeal to or work for a small minority of visitors are the exhibits that tend to be rotated off the floor in favor of more equitably accessible exhibits. For all of these reasons, the third category we will turn to, to judge the quality of collaborative support provided by the different experimental conditions, is:

3. **Learners should participate equitably in the activity**

Specifically, equity in participation (which shows that no visitors are getting left out, from interacting with the activity or from interacting with one another) and equity in performance (which shows that the exhibit's activity is equally accessible to all group members) will be used as measures. A more successful collaborative exhibit should show:

M7: Lower *Participation Inequity*

M8: Lower *Conversation Inequity*

M9: Lower *Performance Inequity*

7.2.2.2 Evidence for Awareness of Joint Goals

Individual Scores: M1

There were not many learning goals for this experiment, given that the software was presented to participants without any of the usual supporting text and imagery typically found with a museum exhibit, and visitors were all forced to play the same role (Surgeon). That said, by constructing the simulation to be as true-to-life as possible (given its degree of abstraction), and the impacts of user actions on that simulation to be as true-to-life as possible (given their abstraction), concepts can be acquired merely by learning to interact with the simulation to "play the game" of eliminating cancer. One goal that remained relevant in this limited set-up was for participants to come to understand that in order to eliminate cancer from a patient, doctors must be diligent about removing all of the cancerous cells they can, even if it means risking some collateral damage to do so.

Section 6.2.2.3 Individual Task Performance describes two scoring mechanisms used to evaluate **M1**, *Individual Scores*. The *Adjusted Weighted Efficacy* method is preferred, as it most closely parallels what is valued in the simulation (and in real life): it is more valuable to kill cancer cells than to damage them, and more valuable to kill or damage cancer cells than to kill or damage normal cells. Essentially, it rewards players for pursuing aggressive treatments (surgeries, in this case). To use a real-world example, under this scoring scheme, full mastectomies would be "worth" more than lumpectomies. The *Unweighted Efficacy* method, on the other hand, rewards precision – damaged and killed cells are worth the same amount of points, but cancer cells have a positive valence, and normal cells are negative. The *Unweighted Efficacy* score is provided just to give another perspective on how visitors interact with the simulation. Given the fact that the simulation was tuned to have aggressively-spreading cancer, however, the aggressive treatment approach is superior.

Although these *Individual Score* measures give only a myopic view, at best, of what a user might be able to learn by using the software, they do have the potential to discriminate between those users who understood the shared task and how to go about it, and those that did not. One of the perennial concerns for educational software is that the added complication of attending to a dynamic display (or, in this case, a "Complex" O-UI interface) has the potential to add enough extraneous cognitive load so as to disrupt the user's ability to use the software appropriately (Lowe, 2003; Najjar, 1998). If the "Complex" O-UI experimental condition shows a higher propensity for poor task performance, as measured by the *Individual Scores*, it indicates that there may be an issue with excessive cognitive load arising from the "Complex" O-UIs.

Table 29. Results of the analysis of *Weighted* and *Unweighted* measures of *Independent Task Performance. (See* Figure 16 *for a visualization).*

	Conditions		
Dependent Variables	**Simple O-UI** *n* = 31	**Complex O-UI** *n* = 31	**Significance** (2-tail, paired t test) *n* = 62, *df* = 30
Adjusted Weighted Efficacy	M = - 0.29 SD = 0.24	M = - 0.16 SD = 0.11	$t(30) = 2.89$ **$p < 0.004$**
Unweighted Efficacy	M = 0.17 SD = 0.29	M = 0.74 SD = 0.17	$t(30) = 10.10$ **$p <$ 0.3.7 E-111**

The results in Table 29 indicate that "Complex" participants significantly outperform "Simple" participants on both the *Adjusted Weighted Score*, nearly two-to-one, and the *Unweighted Score*, by over four times. These results indicate that designers needn't be concerned that the extra cognitive load "Complex" O-UIs induce will inherently prevent visitors from successfully engaging with a software-based learning activity. It seems that the extra cognitive load, which in this case was carefully designed to be germane, not extraneous (see Section 6.2.1 Independent Variable: Opportunistic User Interface Complexity), may even help users to be more "mindful" in their task execution – the data here reflects qualitative observations that, during the study, "Simple" players tended to act with less planning and forethought.

Ownership: M2

The guiding idea behind the *Ownership* measure, **M2**, is that in order to effectively collaborate on the joint task, players to avoid overlapping their partners and duplicating their work. Because "Simple" players were shown to visually attend more to the Shared Display in Section 7.2.1.1 The Heads-Down Effect and Visual Attention Management, one might expect "Simple" participants to also demonstrate higher levels of *Ownership*. A check of the results, however, shows that the result is quite the opposite (see Table 30). From this evidence, it seems that "Complex" players are better at engaging in task division. This echoes informal observations made during the experiments: it seemed as though participants in the "Simple" condition, rather than using the Shared Display to better coordinate their task division efforts, would instead rush to move their incision rectangles to the same places as their partners'. The "Simple" players seemed to be unaware of, or uninterested in, the collaborative aspects of the joint goal. This seemingly emergent competition in the "Simple" condition will be returned to later, in Section 8.2 The "Simple" Condition and Competition.

Table 30. Result of the analysis of the *Individual Task Performance* measure relevant to prediction H7. (See Figure 17 for a visualization).

	Conditions		
Dependent Variables	**"Simple" O-UI Participants** *n* = 31	**"Complex" O-UI Participants** *n* = 31	**Significance** (2-tail paired t test) *n* = 62, *df* = 30
Ownership (degree of sole occupation of territory)	*M* =0.55 *SD* = 0.13	*M* = 0.71 *SD* = 0.15	$t(30) = 5.59$ $\boldsymbol{p <}$ **0.000004**

Outcomes: M3

The other prediction related to goal awareness is the group-level *Outcome* measure, **M3**. In this context, the group can succeed by eliminating the cancer, or fail by causing the patient to die as a result of too much collateral damage or allowing the cancer cells to take over. The *Outcome* is the percentage of episodes that a group participated in that ended in success. While "Simple" participants do see slightly more episodes ending in success (around 88%) than "Complex" users (around 81%) this difference is not significant (see Table 31). This may be the result of a ceiling effect: 73% of "Simple" players and 71% of "Complex" players never lost a single session.

Table 31. Result of the analysis of the *Group Task Performance* measure relevant to prediction H9. (See Figure 18 for a visualization)

	Conditions		
Dependent Variables	**"Simple" O-UI Groups** *n* = 9	**"Complex" O-UI Groups** *n* = 9	**Significance** (1-tail, paired t test) *n* = 18, *df* = 8
Outcome (percentage of episodes ending in success)	*M* =0.88 *SD* = 0.25	*M* = 0.81 *SD* = 0.38	(not significant)

Summary: Awareness of Joint Goals

Overall, "Complex" participants demonstrated higher awareness of joint goals, both by scoring significantly higher on individual measures of task performance (**M1**) and by demonstrating, if not more awareness of, at least adherence to, the collaborative aspects of the shared task (**M2**). The measure of group goal awareness, Outcome (**M3**), was essentially the same for both conditions, although a ceiling effect may have been involved.

7.2.2.3 Evidence for Interaction between Visitors

The assumption of many educational researchers who frame their work from a socio-cultural perspective is that talk is a necessary component of small-group learning: without using talk to externalize individual mental models, it is impossible for members of a group to learn from one another (Sawyer, 2006). Thus, one simple measure of interaction in small-group learning (and one initially applied by this research) is the quantity of conversation. Recall from Section 7.2.1.2 The Heads-Down Effect and Awareness of Shared Context that the difference in utterance frequency between the "Simple" (M = 20.96, SD = 10.35) and "Complex" O-UI (M = 15.55, SD = 8.34) conditions was found to be statistically significant, $t(8) = 2.60$, $p < 0.032$, in favor of the "Simple" groups. To further explore how group conversations may be affected by differences in O-UI "complexity," an analysis with more nuance needs to be conducted. For this purposes, utterances were coded into *On-task/Off-task*, *Functional/Nonfunctional*, and *Interactional/Non-interactional* categories.

Utterance Coding Procedures

Remarks that were out-loud readings of on-screen messages (e.g., "'You killed the cancer!'", or "'Use your stylus to draw lines'") were excluded from analysis, since the presence or absence of such utterances is dependent on the presence or absence of on-screen messages. The *On-task/Off-task* designation (**M4**) was made on the basis of whether the remark added new information about the task at hand, either in terms of observational comments ("There's a lot of cancer cells"), tactical statements (e.g., "I'm going over here"), or more abstract or strategic ideas about the simulation ("I think if you leave anything [meaning, any cancer cells] there, does it automatically grow back?"). Examples of utterances coded as *Off-task* included remarks on companions who were not present (e.g., "[Name redacted] is missing out!"), remarks on general competencies unrelated to current task execution (e.g., "My mother would be so proud to know I was [using a handheld computer]"), "trash talk" (e.g., "I rock! Take that!"), or "mutterings" not really intended to communicate ideas to others (e.g., "Oops!"). Also excluded from *On-task* designation were acknowledgement remarks (e.g., "Yeah," "I see," "OK," "Oh") when such remarks were used only to acknowledge another player's speaking turn, as in this exchange:

North: "We're doing better, now"
East: "Yeah."

In this example, the presence or absence of the "yeah" would not have changed the groups' mutual understanding of the situation. By way of contrast, acknowledgement utterances were coded as being *On-task* when they were made in response to action proposals, as in this exchange:

West: "All right, [South], stay over with me."
South: "OK."

In this situation, the presence or absence of the "OK" may very well alter the groups' mutual understanding of the situation, as West (and the other players) would not necessarily know if South was planning on going along with West's proposal without hearing the acknowledgement.

Determining whether or not an utterance had a *Functional* component (the Tactical/Strategic measure **M5**) to it required deciding of the utterance had either a tactical or strategic element to them – in other words, if it contained content that would likely improve the group's ability to cooperatively fight the cancer. Many of the remarks were more observational in nature – noting the general pace of the cancer growth, the difficulty level, or the skill of players – and did not necessarily promote any specific cooperative actions.

Remarks were considered to be *Interactional* (**M6**) if they were coded as either *Responses* to another group member's utterances, or *Continuations* of a prior remark made by the current speaker. The idea was to get a measure of conversational elaboration, which is often used in museum visitor studies as a marker for learning (Scott. G. Paris & Hapgood, 2002). Thus, utterances are coded as *New Statements*, and go uncounted for the *Interactional* measure, even if they directly address another player or make an insightful point, if they are not made in reply to another player or do not reference content from earlier in the dialogue.

Table 32. Demonstration of how the utterances of a conversational exchange were coded for *On-task*, *Functional*, and *Interactional* categories.

		On-task?	Functional?	Interactional?
North:	"Come on, go away."	No	No	No
East:	"Oh, wow, you got it most of the way clean."	Yes	No	No
North:	"I've gotta go down the edge here."	Yes	Yes	No
East:	"You do that."	Yes	Yes	Yes
North:	"Wow."	No	No	No
East:	"Whoops."	No	No	No

Table 32 provides an example of how utterances are coded for the three categories studied here. The first statement ("Come on, go away.") does not qualify for any of the three categories because it was a muttered remark addressed at the cancer cell being eliminated, and not to any particular group member, and adds no new content to the shared context. The last two remarks ("Wow" and "Whoops") are excluded for similar reasons. The second utterance ("Oh, wow, you got it most of the way clean."), while considered to be *On-task*, does not contain any tactical or strategic content – it is primarily observational in nature – and so is not considered to be *Functional*. Although it is directly addressed to the other player (East), it is not a reply and doesn't refer back to any prior content, so it is not considered to be *Interactional* either. The third utterance ("I've gotta go down the edge here.") is considered to be both *On-task* (adding to the group's overall informational content) and *Functional* (because it illuminates the approach she will use in the next few moments of play). Likewise, East's reply ("You do that") replies and indicates his approval of North's proposed strategy, and so it qualifies for all three categories.

A check for inter-coder reliability on one complete transcript was initially at 87.1%, 82.4%, and 99.97% agreement for *On-task*, *Functional*, and *Interactional* coding, respectively, after a first pass. Owing to the strong agreement between the primary coder and the reliability-check coder on *Interactional* coding, no further work was done to verify the accuracy of the *Interactional* coding[32]. After a further clarification of definitions for the other two coding categories, *On-task* and *Functional*, the inter-coder agreement was brought up to 97.2% and 92.6%, respectively. A cross-check on a new, full transcript was found to be at 94.4% for *On-task* coding and 91.9% for *Functional* coding. The primary coder reviewed and amended the *On-task* and *Functional* codings of the remaining transcripts to reflect the negotiated definition changes.

Table 33. Results of the analyses of Proportions of conversational utterances coded as belonging to each of the following categories, for each condition. (See Figure 19 for a visualization).

	Conditions		
Dependent Variables	**"Simple" O-UI Participants** *n* = 31	**"Complex" O-UI Participants** *n* = 31	**Significance** (2-tail, paired t test) *n* = 62, *df* = 30
On Task Proportion	*M* = 0.36 *SD* = 0.18	*M* = 0.49 *SD* = 0.17	*t*(30) = 4.04 ***p* < 0.0004**
Functional Proportion	*M* = 0.21 *SD* = 0.14	*M* = 0.30 *SD* = 0.19	*t*(30) = 2.79 ***p* < 0.0092**
Interactional Proportion	*M* = 0.47 *SD* = 0.20	*M* = 0.49 *SD* = 0.19	(none)

Coded Utterance Results: M4, M5, and M6

The analysis of coded utterances is presented in Table 33. Although no predictions were made for the impact of O-UI "complexity" on the different coded utterance proportions, the results show that "Complex" participants demonstrate significantly higher *Proportions* of *On-task* (**M4**) and *Functional* (**M5**) utterances, while the proportion of *Interactional* (**M6**) utterances remains virtually the same. Recall that the comparison being made is a paired t-test, meaning that the utterances made by one participant in one condition are compared against the utterances made by that same participant in the other condition, so the O-UI clearly has a strong impact on a person's *On-task* and *Functional* remarks. Although "Complex" users on the whole spoke less than "Simple" users, as established in Section 7.2.1.2 The Heads-Down

[32] Broken down by category (by dividing the number of utterances assigned to a particular category by *both* coders by the number assigned to that category by *either* coder), there was agreement on the *New Statements* 97.2% of the time, *Responses* 98% of the time, and *Continuations* 88.9% of the time. The slightly lower agreement on *Continuations* was a result of their much lower frequency in group dialogues – only 8 or 9 occurred in the transcript tested.

Effect and Awareness of Shared Context, nearly half (M = 0.49, SD = 0.17) of "Complex" participants' utterances were considered to be *On-task*, compared to a little better than one third of the remarks made by "Simple" participants (M = 0.36, SD = 0.18).

Utterances can come in different sizes, though. Taking a look at the number of words used in the *On-task* utterances eliminates the possibility that the "Complex" participants' *On-task* utterances were overly simplistic in nature (word count can be used as a very rough measure of an utterance's complexity). Consistent with the finding that "Complex" participants are more likely to make *On-task* utterances when they speak, the "Complex" participants had a larger proportion of their total word count devoted to *On-task* utterances, with about 70% of the words used during a session belonging to *On-task* remarks (M = 0.70, SD = 0.15). This can be contrasted against "Simple" users' lower *On-task* word count proportion (M = 0.52, SD = 0.15), where roughly half of all words uttered are *On-task*, which is significantly lower than the Complex users' proportion, $t(30) = 3.70$, $p < 0.0004$, paired 2-tailed. A similar pattern holds for the word counts devoted to *Functional* utterances, with "Complex" users devoting above 40% of their words to *Functional* remarks (M = 0.44, SD = 0.24), a significantly higher proportion, $t(30) = 2.59$, $p < 0.01$, paired 2-tailed, than "Simple" users (M = 0.30, SD = 0.21). See Figure 82 for a plot.

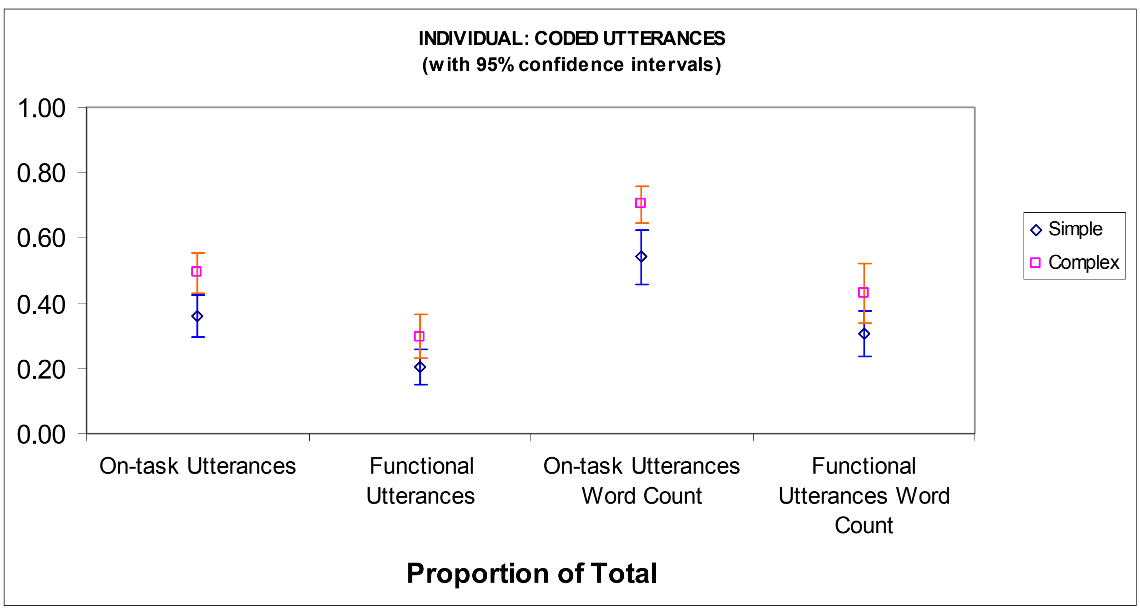


Figure 82. Plot of the proportion of utterances classified as either On-task or Functional, and the proportion of words used by each participant that went into On-task or Functional utterances, with 95% confidence intervals.

The lack of a difference between conditions on the *Interactional* proportion is also important to note. The "complexity" of the O-UI, while affecting *overall* levels of conversation, as described in Section 7.2.1.2 The Heads-Down Effect and Awareness of Shared Context, doesn't seem to affect the *proportion* of conversation that is devoted to building on prior discussions amongst the group members. To illustrate why this is an important finding: it might be reasonable to assume that, since "Complex" participants must devote more of their working memory to their O-UIs, they may also be less likely to monitor utterances from other players[33], or they may be less able to maintain an internal representation of the conversational history, or both. The results here show that while it is true that there are fewer remarks overall for "Complex" participants to reply to or reference in their conversations, participants in the "Complex" condition are no less likely to respond to other players' remarks, nor are they less likely to recall elements of prior conversation, than "Simple" participants. A cross-check of the word counts used in the utterances coded as *Interactional* (namely, *Responses* or *Continuations*) shows that the proportion of words devoted to "Complex" participants' *Interactional* utterances (M = 0.40, SD = 0.21) is nearly identical to that for "Simple" participants (M = 0.42, SD = 0.22). See Figure 83 for a visual representation.

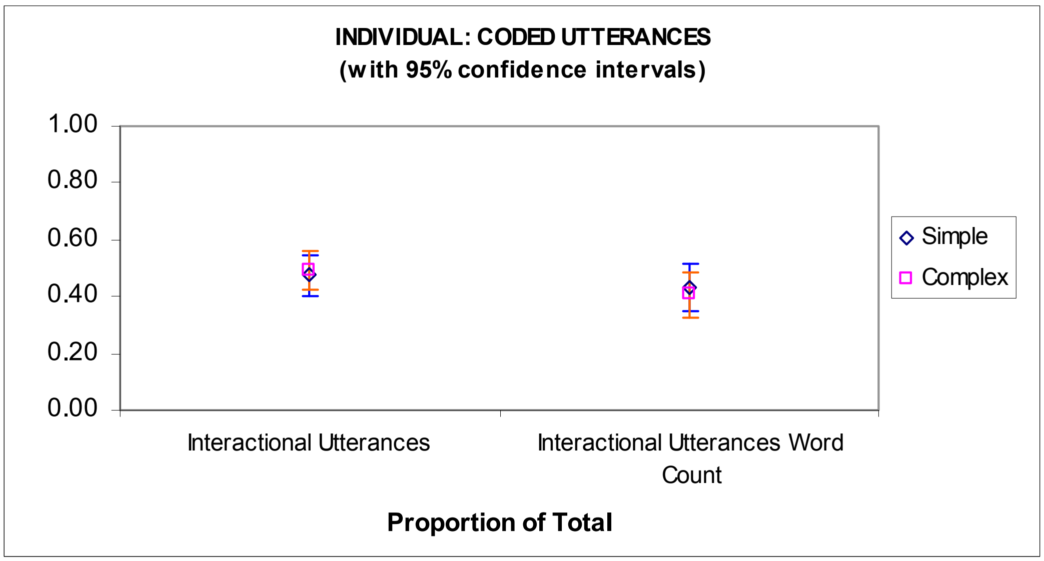


Figure 83. Depiction of the proportion of utterances classified as either *On-task* or *Functional*, and the proportion of words used by each participant that went into *On-task* or *Functional* utterances, with 95% confidence intervals.

Summary: Interaction between Visitors

Overall then, it seems that even though the total frequency of Interactions is lower for "Complex" participants, they still engage in higher-quality *Interactions* than "Simple" participants, measured in terms of On-task (**M4**) and Functional (i.e., Tactical/Strategic, **M5**) utterances. There was no difference between "Simple" and "Complex" participants on the Interactional measure (**M6**) of utterances, so it seems that while "Complex" participants speak less overall, they are not neglecting their companions any more than "Simple" participants may be. This ties into the observations made in Section 7.2.1.3 Demonstrating Relationship of Heads-Down Visual Attention Behaviors and Awareness, which showed that there was no correlation for "Complex" participants' conversation frequencies and their visual attention behaviors.

[33] It should be mentioned that there is some evidence from carefully constructed cognitive psychology experiments that verbal and visual stimuli do not operate additively on a person's working memory resources (Friedman & Miyake, 2000; Shah & Miyake, 1996). The way the brain structures the processing of the differing modalities allows people to process more incoming stimuli when it is divided between visual and verbal modalities than when it is provided via one modality alone.

7.2.2.4 Evidence for Equity

Other studies looking at the suitability of group activities for supporting collaborative learning have relied on measures of *Participation Inequity* to determine if a given activity is likely to result in consistent learning gains for each participant. This section will review the results of both the *Participation Inequity* (**M7**) calculations as well as the *Conversation Inequity* (**M8**) and *Performance Inequity* (**M9**) calculations for groups in both conditions.

Table 34. Results of the *Inequity* analysis of the *Participation* (**M7**) and *Conversation* (**M8**) measures. For each measure, the population standard deviation was computed across the members in each group, providing an estimate of how dissimilar group members were in their engagement behaviors. (See Figure 20 and Figure 21 for visualizations).

		Conditions		
Dependent Variables		**"Simple" O-UI Groups** *n* = 9	**"Complex" O-UI Groups** *n* = 9	**Significance** (2-tail, paired t test) *n* = 18, *df* = 8
Participation Inequity (STDEVP of inputs per minute)	**Moves**	***M* = 16.27** *SD* = 15.78	***M* = 5.68** *SD* = 2.92	$t(8) = 2.24$ **$p < 0.056$**
	Damage Events	***M* = 6.26** *SD* = 5.78	***M* = 10.05** *SD* = 3.29	**none**
	Total Participation	***M* = 21.31** *SD* = 21.26	***M* = 11.20** *SD* = 4.64	**none**
Conversation Inequity (STDEVP of utterances per minute)		***M* = 1.30** *SD* = 0.92	***M* = 1.28** *SD* = 0.77	**none**

An initial reading of the data in Table 34 suggests that with the exception of *Moves*, participants in the "Simple" and "Complex" conditions do not exhibit significantly different degrees of within-group dissimilarity on *Participation* (**M7**) and *Conversation* (**M8**) measures. So, for the most part, there is no strong reason to prefer one degree of O-UI "complexity" over another, if one is concerned about the *Conversation Inequity* the UI might engender. The participants in the "Simple" condition did exhibit a significantly higher level of inequity in the frequency of *Moves* made, however. To determine whether or not this difference on inequity of moves is reason for concern, a return to the original data is warranted. A clue for what to look for in the analysis can be found in the following conversational exchange. It takes place between two older female participants (East and South) after switching from the "Complex" condition to the "Simple" condition:

West: "More like a, more like a classic video game"

East: "Yeah, this is more like a teenage boy thing cause you gotta have that eye - uh"

South: "Yeah, pow pow pow" <Laughs>

East:	"But you gotta move around really fast, you know what I mean? They're good at that."
South	"Yeah I know."
East:	"I'm not as good at it."
East:	[addressed to West] "Ok, you got em! You guys could do this game by yourself, you could do…" …
East:	"You guys have that dexterity thing."
South:	<laughs>

The exchange above motivated another look at the *Engagement* measures, this time taking gender and age into account, to see if they were interacting with the *Engagement Inequity* measures. Since this qualifies as a *post hoc* analysis, given that *Gender* and *Age* were not originally considered to be independent variables under study, any statistical significance associated with inferential results should be treated carefully. With respect to *Age*, there were slight negative correlations with most of the *Engagement* measures. This might be expected – younger people are popularly acknowledged to have a greater affinity for, and facility with, computer-based activities. Only one of these correlations reached statistical significance: the negative correlation between *Age* and *Total Participation*, *Pearson's* $r(70) = -0.202$, $p < 0.045$. The magnitude of the correlation is rather low, however, so it would seem that *Age* is not providing any major interaction effects.

Gender, on the other hand, does seem to have a modifying effect on the *Move Frequency* measure (see Figure 84). While females and males are very similar in their *Move Frequencies* in the "Complex" condition (M=19.83, SD = 7.96, and M = 21.70, SD = 10.28, respectively), they show very different *Move Frequencies* in the "Simple" condition (M=42.80, SD = 21.42, and M = 63.95, SD = 25.82, respectively). Merely looking at these descriptive statistics is enough to show that while participants in the "Complex" condition show relatively similar *Engagement* in terms of *Move Frequency*, there are stark gender-related differences for the "Simple" participants.

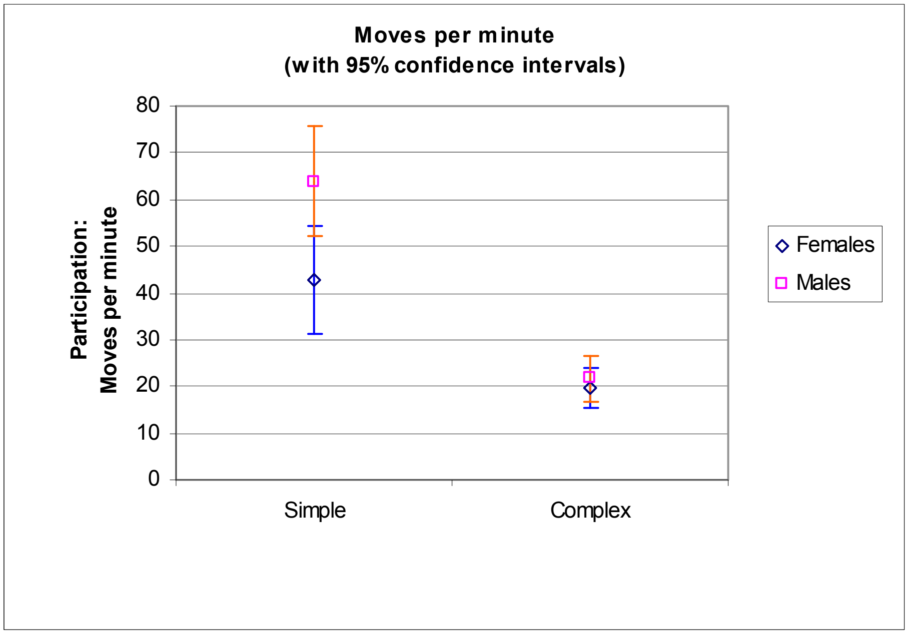


Figure 84. The modifying effect of *Gender* on the main effect of O-UI "complexity" on *Move Frequency* is depicted in this chart, with 95% confidence intervals.

It may not be appropriate to apply an ANOVA (Analysis of Variance) in a post-hoc fashion, but the results do show that while the main effect of O-UI complexity is still the predominant factor, gender does indeed seem to play a role in generating some of the variance seen in *Move Frequency* (see Table 35). Taken together with the descriptive statistics and the comments made by participants, it seems prudent for educational designers to weigh this factor when selecting a level of O-UI "complexity." It seems that while "Simple" groups suffer less from many symptoms of the heads-down effect (as shown by the results of Section 7.2.1 Establishing the Existence of the Heads-Down Effect), more "Complex" O-UIs promote better within-group equity of *Participation* in input activities that rely on the Shared Display (i.e., *Moves*). Moreover, it seems that some of this difference in equity may be related to the fact that "Simple" O-UIs seem to promote inequities along gender. True, this inequity only seems to apply to *Participation* in input activities that rely on visually attending to the Shared Display, hardly a broad-brush damning of the use of "Simple" O-UIs. Still, if a goal of museum exhibits is to support, as best as possible, equitable *Participation* amongst a wide variety of potential learners, inequities rooted in demographic variables (like gender, age, race, or creed) should be avoided or minimized if possible.

Table 35. An analysis of variance of the *Move Frequency*, considering O-UI "complexity" and Gender as the two independent variables.

Analysis of Variance of Move Frequency					
Source	***df***	**Sum of Squares**	**Mean Square**	***F***	***P***
O-UI Complexity	1	19,026.52	19026.52	57.53	$p < 3.75E{-10}$
Gender	1	3,647.61	3647.61	11.03	$p < 0.0016$
O-UI Complexity x Gender	1	1,772.93	1772.93	5.36	$p < 0.024$
Within-groups error	56	18,521.21	330.74		
Total	59	42,968.27			

Finally, as with measures of *Participation* and *Conversation*, *Task Performance* (**M9**) measures should be examined for the presence of any inequities (see Table 36). With respect to *Ownership*, it seems that the members within each group behave very similar to one another, and there is no effect related to O-UI "complexity." The "Simple" participants do show a greater amount of *Inequity* than "Complex" participants on the two *Efficacy* measures, and this difference is significant for the *Adjusted Weighted Efficacy*. The interpretation, then, is that the "Complex" O-UI condition encourages participants to "operate" somewhat more similarly to one another. Combined with the superior *Task Performance* exhibited by "Complex" participants, as described in Section 7.2.2.2 Evidence for Awareness of Joint Goals, one might conclude that the "Complex" O-UI supports both better and more equitable task execution.

Table 36. Results of the analysis of the *Inequity* of *Individual Task Performance* measures. (See Figure 22 for a visualization).

	Conditions		
Dependent Variables	**Simple O-UI** $n = 9$	**Complex O-UI** $n = 9$	**Significance** (2-tail, paired t test) $n = 18$, $df = 8$
Score Inequity: Adjusted Weighted Efficacy	$M = 0.14$ $SD = 0.09$	$M = 0.08$ $SD = 0.06$	$t(8) = 2.38$ $\mathbf{p < 0.04}$
Score Inequity: Unweighted Efficacy	$M = 0.17$ $SD = 0.08$	$M = 0.11$ $SD = 0.10$	(not significant)
Ownership Inequity	$M = 0.08$ $SD = 0.05$	$M = 0.08$ $SD = 0.04$	(not significant)

CHAPTER 8

Discussion and Future Work

The concept of an Opportunistic User Interface (O-UI) is applicable to a wide array of contexts, especially as more and more people purchase mobile devices with sophisticated display and communication capabilities. Users with mobile devices could quite literally stumble across a physical activity site – a museum exhibit, a travel-planning kiosk, a pizza parlor – an be able to use their device as an O-UI to interact with the activity hosted at that physical site (engaging with an educational simulation, planning a trip with their partner, or selecting mutually-agreeable pizza toppings with a group of friends). This work speaks to any scenario where an O-UI is used to interface with a shared display – effectively creating a Multi-Machine User Interface (MMUI) – in the service of collaborative tasks. Increasingly, the focus of both learning and work is on collaborative group tasks, and increasingly technology is being employed to aid and abet these tasks (classrooms and workplaces were the very first places Multi-Machine User Interfaces were explored). The principles gleaned from this work should apply to any scenario wherein equitable, cooperative participation in a joint task is the goal.

8.1 The "Verdict" on O-UI "Complexity"

The evidence from the preceding chapter shows that the prediction of the first research question, that O-UI "Complexity" exacerbates the heads-down phenomenon, holds true. Of the ten specific hypotheses made, **H1** – **H10**, seven were significantly confirmed, two were inconclusive, and only one was countermanded (although it was based on a faulty assumption, and so was not measuring what it was intended to measure).

It is clear from the evidence that "Complex" O-Us promote externally-visible visual attention management behaviors in line with the described heads-down phenomenon. The predictions of **H1** (gaze *Proportion*), **H2** (gaze *Duration*), and **H4** (gaze *Synchronicity Degree*) all held, demonstrating that "Complex" participants did indeed gaze longer, and for longer unbroken durations, at their O-UIs, and that "Simple" participants were better able to establish joint attention. (**H3** was dismissed from consideration because it relied on a misleading measure, gaze *Frequency*).

The next task was to look for secondary symptoms that would demonstrate whether or not the observed poor visual attention behaviors were in fact resulting in poorer awareness of the shared context. The evidence showed that while "Complex" participants reported about the same levels of *Awareness* (**H5**) as "Simple" participants, correlations with visual attention behaviors revealed that self-reported awareness was improved when participants engaged in more "heads-up" visual attention behaviors (**H8a**), and impaired when they engaged in more "heads-down" behaviors (**H8b**). More objective measures than self-reports showed that regardless of participant perceptions, "Complex" participants were significantly less engaged with elements of the shared context, *Conversing* (**H6**) less with their companions, and interacting less with the Shared Display (as evidenced by *Move* frequency, **H7**). Correlations with visual attention behaviors showed that while more "heads-up" behaviors were associated with more interaction with the Shared Display (i.e., Moves, **H10a**), and "heads-down" behaviors corresponded with less interaction with the Shared Display, (**H10b**), no such pattern existed for *Conversation* for "Complex" players (**H9a** and **H9b**) – they were equally likely to speak t their partners regardless of their visual attention behaviors.

With respect to the second research question, support for collaborative learning, even though the "Complex" participants suffered more from the heads-down effect, they went on to perform equivalently or better on each of the nine measures chosen to speak to the potential for the activity to support collaborative learning. The "Complex" participants performed significantly better on **M1** (Individual Scores) , **M2** (Ownership), **M4** (On-task Utterances), **M5** (Tactical/Strategic Utterances), **M7** (on Moves component of Participation Inequity), **M9** (on Weighted Efficacy measure of Performance Inequity). "Complex" participants performed at about the same level on **M3** (Group Outcomes), M6 (Interactional Utterances), **M7** (on Damages component of Participation Inequity), **M8** (Conversation Inequity), and **M9** (on the Unweighted Efficacy and Ownership measures of Performance Inequity). Overall, then, the "Complex" condition was judged superior for the purpose of supporting collaborative learning, especially since Section 7.2.2.4 Evidence for Equity demonstrated that the "Simple" condition promoted gender inequity in performance.

At this point, one must legitimately ask – were those who flagged the heads-down phenomenon as being problematic just *wrong* about the deleterious impact of poor visual attention management behaviors? A quick look at how the groups' gaze behaviors correlate with the *Outcomes* experienced by the groups quickly dismisses that question (see Table 37). Both "Simple" and "Complex" groups showed significant negative correlations between the average *Proportion* of time the members in each group spent looking at the O-UI and the degree of positive *Outcomes*. Likewise, "Complex" groups showed a significant negative correlation between the average *Duration* of time the members in the group gazed at the O-UI, and the *Outcomes*. Inversely, "Complex" groups showed a significant positive correlation between the average *Proportion* of time the members in each group spent looking at the Shared Display and the degree of positive *Outcomes* ("Simple" groups had a similar, but non-significant, trend). Likewise, both the "Simple" and "Complex" groups showed significant positive correlations between the *Gaze Synchronicity Degree* and the degree of positive *Outcomes*. The implication seems to be that groups, and especially "Complex" groups, need to attend to the Shared Display, and not spend too much contiguous time gazing at the O-UIs, to perform well in terms of ultimate *Outcomes*. A corroborating piece of evidence is that, for "Complex" players, higher frequencies of gaze shifts also correlate significantly and positively with game *Outcomes*.

Table 37. Correlations between *Individual Division of Attention* measures, averaged across each group, and the *Group Task Performance Outcome* measure for each group.

	Gaze Proportion O-UI	**Gaze Proportion Shared Display**	**Gaze Duration O-UI**	**Gaze Duration Shared Display**	**Gaze Shift Frequency**	**Gaze Synchronicity Degree**
"Simple" Outcomes (n = 11, df = 9)	**- 0.696** $t(9) = 2.91$, $p < 0.018$	0.493 (not significant)	- 0.507 (not significant)	0.250 (not significant)	- 0.100 (not significant)	**0.690** $t(9) = 2.86$ $p < 0.019$
"Complex" Outcomes (n = 11, df = 9)	**- 0.718** $t(9) = 3.09$, $p < 0.013$	**0.662** $t(9) = 2.65$, $p < 0.027$	**- 0.732** $t(9) = 3.22$, $p < 0.011$	0.261 (not significant)	**0.724** $t(9) = 3.15$, $p < 0.012$	**0.741** $t(9) = 3.31$ $p < 0.00907$

The preceding makes a case that, despite the poor attention management exhibited by "Complex" users, the "Complex" O-UI may actually be superior for supporting collaborative learning, although it can still be improved. To the extent that they desire to support collaborative museum exhibits, designers can, and probably should, consider making use of the displays of O-UIs. This work demonstrated that even while using an O-UI with a near worst-case version of output "complexity," the users still better the performance of participants using "Simple" O-UIs on a large host of measures of collaborative quality. There seems to be little reason *not* to fully make use of the graphical display capacities of mobile devices, as long as the displays are germane to the task at hand. From the preceding it does seem to be the case that "Complex" groups improved their performance on the activity when they engaged in more "heads-up," and less "heads-down," visual attention behaviors, however, so designers may wish to explore how to help O-UI users manage their attention. This brings us to the first Design Recommendation to emerge from this work.

8.1.1 Design Recommendation 1: O-UI-s for Museum Exhibits

8.1.1.1 Context

When designing for:

- A synchronous, co-located museum exhibit
- Ongoing dynamic event-driven activity
- Cooperative activity with homogeneous roles
- No externally-imposed task division
- Loosely-coupled input
- Loosely-coupled output provided via shared display and O-UIs
- Many dynamic (but germane) visual elements to attend to on O-UI display screen

8.1.1.2 Recommendation

To improve performance on joint task, additional mechanisms are needed to ensure:

- Users do not gaze at O-UI for long unbroken *Durations*
- Users do not gaze *Proportionally* more at the O-UI than the Shared Display
- Users shift their gaze *Frequently* between displays, and at regular intervals
- Users gaze at the Shared Display *Synchronously* from time to time

8.2 The "Simple" Condition and Competition

Although O-UI "complexity" does lead to attention management behaviors consistent with the predictions of museum professionals, it also seems better suited for supporting collaborative learning – an endeavor which, by all rights, should be reliant on good attention management behaviors. How can this be?

8.2.1 Demonstrating Emergent Competition in the "Simple" Condition

The answer may lie, paradoxically, with the visual attention paid to the Shared Display. Participants in the "Simple" condition were observed to "herd" around the same locations on the Shared Display, which is why they scored significantly lower than "Complex" users on the *Ownership* measure of task division. One possibility is that this could be due to "blind copying," a phenomenon observed in one other study of SDG usage, where a user relies on a partner to "show the way" and merely replicates what he or she has done (Kerawalla, Pearce, Yuill, Luckin, & Harris, 2008). And yet, "blind copying," which seems to imply a certain aimlessness, doesn't fit with what was anecdotally observed during the experiment: participants in the "Simple" condition seemed not just to be following each other, but to be *racing* each other to get to certain regions of the Shared Display; essentially, they appeared to be actively competing against one another.

Recall from Section 7.2.1.3 Demonstrating Relationship of Heads-Down Visual Attention Behaviors and Awareness that when the correlations between self-reported measures of awareness and "heads-up" attention behaviors were computed, the correlations were strongest for "Simple" participants on the question that references awareness to their partners (Q7: "I was aware of how well my partners were doing at all times"). By way of contrast, the correlations were strongest for "Complex" participants on the question that referenced awareness of the group (Q9: "I was aware of how well our group was doing at all times"), hinting that the participants in the "Complex" condition were more group-aligned, whereas the participants in the "Simple" condition were more inclined towards perceiving their status as being separate from the status of the other group members. It seemed like the constant monitoring of the Shared Display, and their partners' actions displayed on it, prompted "Simple" participants to engage in more self-versus-other comparisons.

Excessive monitoring, and awareness of being monitored, may also be what underlies the gender inequities seen in the "Simple" condition. Recall from Section 7.2.2.4 Evidence for Equity, that female players were observed setting down their handhelds at times during the "Simple" condition, as after this exchange:

East:	"But you gotta move around really fast, you know what I mean? They're good at that."
South	"Yeah I know."
East:	"I'm not as good at it."
East:	[addressed to West] "Ok, you got em! You guys could do this game by yourself, you could do…"

East was clearly very aware of how her performance compared to that of the younger male participants, and was demonstrating a self-consciousness about her ability to contribute to the shared task, which seemed to prompt her to withdraw from the activity. Could excessive monitoring be the culprit? Not enough instances of female withdrawal exists to confirm monitoring as the cause, but there is enough evidence to demonstrate that emergent competition is the result of monitoring.

Gaze Synchronicity Degree is effectively a measure of the degree of monitoring taking place during a session, although it is a composite measure, combining the proportion of time dyads viewed the Shared Display with the proportion of time triads viewed it synchronously and so on. By separating out these component degrees, and calculating their correlations with certain other measures, we can see how these measures change as the degree of Gaze Synchronicity increases. For example, Figure 85 shows the correlations of *Ownership* with different degrees of *Gaze Synchronicity*. The strong positive correlation between *Ownership* and *Dyad Gaze Synchronicity* for "Simple" players shows that when only two "Simple" partners are gazing synchronously at the shared display, high *Ownership* is promoted. That changes, though, as more partners begin monitoring the shared display – the correlation changes to a slightly negative one for Triads and a moderately negative one for Quads. This trend shows that with more monitoring, "Simple" players are overlapping more – although a glance at the correlation pattern for "Complex:" participants shows that increased monitoring does not have the same effect on them.

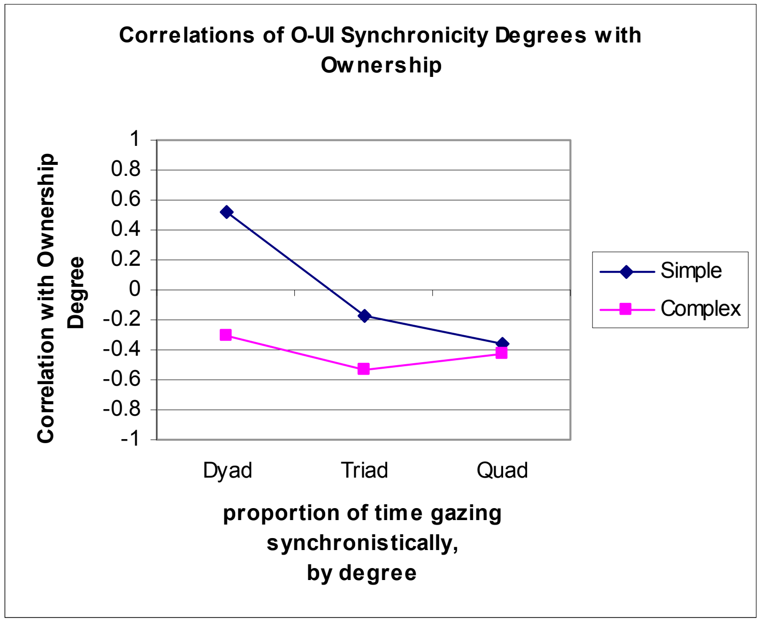


Figure 85. Plot of the correlations between the degree of *Ownership* and different degrees of *Gaze Synchronicity*. Notice that the correlations remain relatively the same for “Complex” participants, while for “Simple” participants, the correlation changes from a strong positive one to a strong negative one as the *Synchronicity Degree* increases. The implication is that more comparative monitoring results in more crowding/herding positioning behaviors.

The preceding plot supports the informal observations that increased monitoring results in increased “herding” behaviors in “Simple” participants, and the next plot supports the informal observations that the herding becomes more “frantic” for “Simple” participants as monitoring increases (see Figure 86). “Simple” participants start out with a slightly negative correlation between the number of *Moves* and *Gaze Synchronicity*, but as the *Synchronicity* degree increases, the correlation becomes quite positive. Once again, “Complex” participants’ *Move* behaviors do not seem to be overly affected by monitoring.

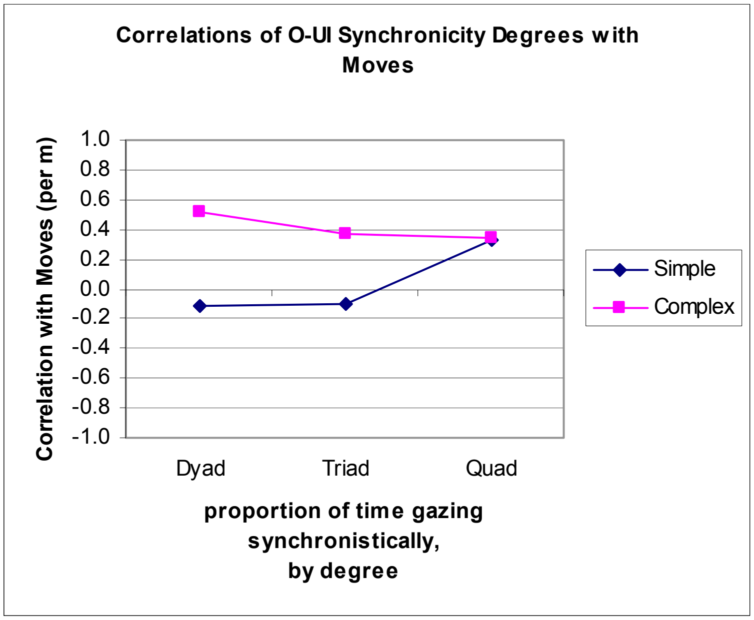


Figure 86. Plot of the correlations between the number of *Moves* and different degrees of *Gaze Synchronicity*. Notice that the correlations remain relatively the same for "Complex" participants, while for "Simple" participants, the correlation changes from a slight negative one to a strong positive one as the *Synchronicity Degree* increases. The implication is that more comparative monitoring results in more *Move* making.

A final piece of evidence supporting the assertion that "Simple" players became more "frantic" and "competitive" as they engaged in more monitoring of one another comes from an analysis of the correlations between different *Damage* events and *Gaze Synchronicity Degrees*. For "Simple" participants, there is a clear reversal: while *Gaze Synchronicity* correlates positively at the beginning with killing and damaging cancer cells, and negatively with killing and damaging normal cells, as monitoring increases, this pattern reverses (see Figure 87). This seems to support the observed behaviors – that players got more careless as they tried to compete against one another for the same "kills." This finding may be related to the fact that gender seemed to have an effect on *Participation* in the "Simple" condition. One researcher studying children's use of an educational SDG activity under different reward structures (competitive versus cooperative versus shared mouse versus individual) found that "the competitive [condition] hampered thoughtful decision-making, skewing instead towards impulsive clicking" for the male subjects, but not the females, who showed consistently high performance across all of the shared-use conditions (Pawar et al., 2007). By way of comparison, "Complex" players show relatively consistent correlations for each of the damage types, regardless of monitoring degree.

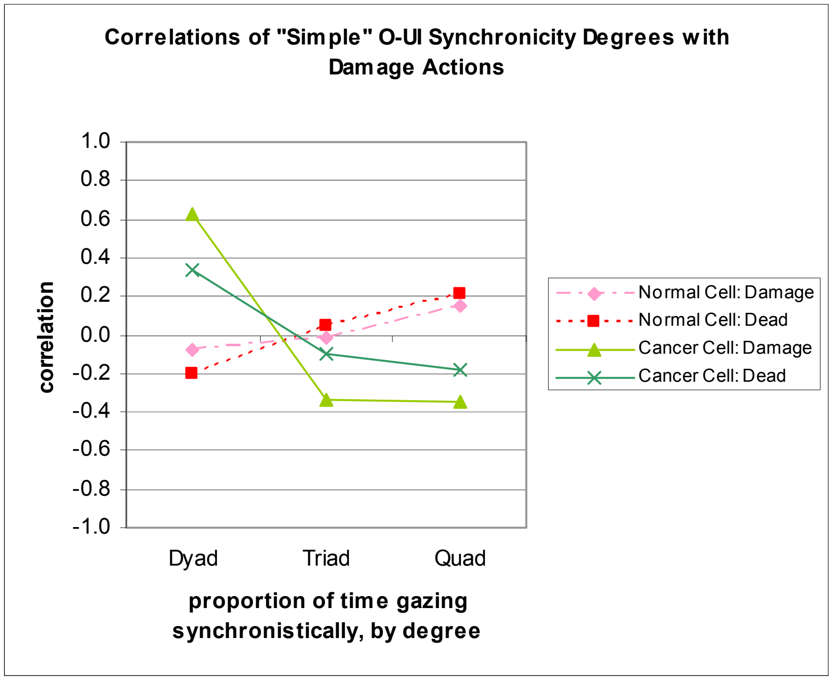


Figure 87. Plot of the correlations between different types of *Damage* actions and different degrees of *Gaze Synchronicity* for "Simple" participants. Notice that the correlations for dead and damaged cancer cells change from strong positive to moderately negative, whereas the correlations for dead and damaged normal cells change from mildly negative to mildly positive, indicating that "Simple" participants become more careless with their choice of targets as monitoring increases.

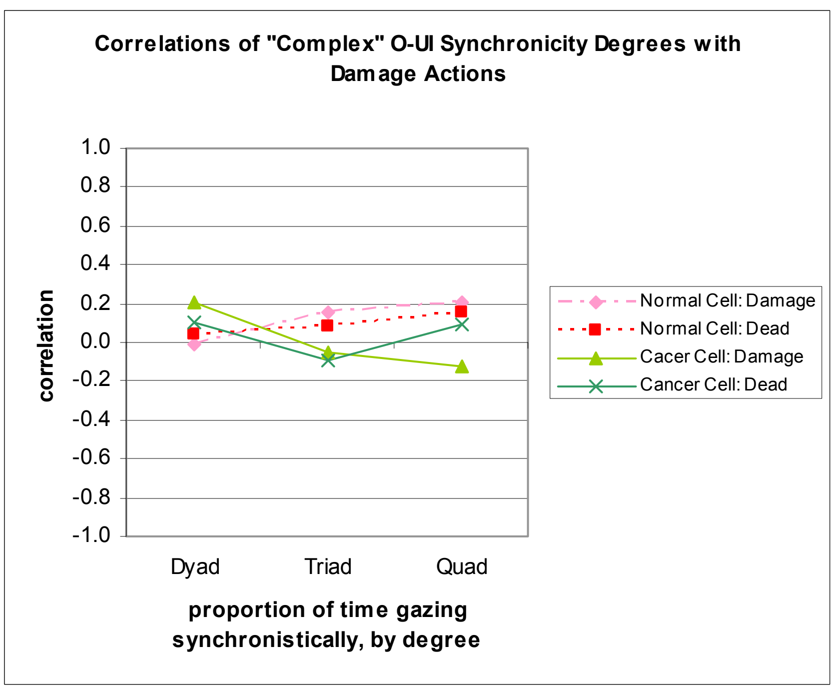


Figure 88. Plot of the correlations between different types of *Damage* actions and different degrees of *Gaze Synchronicity* for "Complex" participants. Notice that the correlations remain relatively the same for each type of damage action, regardless of the degree of monitoring.

8.2.1 Competition and Coupling

Increased monitoring clearly is associated with emergent competition in the "Simple" condition, but nor the "Complex" condition. Relating this finding back to larger themes found in Computer-Supported Collaborative Work (CSCW), and Human-Computer Interaction (HCI), the effect of supplying a "simpler" O-UI is the same as creating a tighter *coupling* between the users' outputs. (See Section 4.1.2.2 Coupling for a description of coupling in the context of collaborative computer activities). As the O-UI is made "simpler" – here, by replacing dynamic visual elements with static elements – the users devote more and more of their visual attention to the shared display. Thus, as O-UI designs become "simpler," the overall experience begins to resemble that of a WYSIWIS (What You See Is What I See) system. WYSIWIS systems exhibit the tightest possible output coupling, wherein users are privy to the exact same output.

Researchers working in CSCW have long posited that tighter coupling of input and output encourages – enforces, even – collaboration. This research, though, shows that output coupling is not enough to ensure, or even promote, collaboration – despite experiencing a more WYSIWIS arrangement of output, "Simple" users actually performed significantly more poorly on measures of collaboration like task division, and even displayed non-collaborative, competitive behaviors. Something else beyond output coupling is clearly needed to promote (or enforce) collaboration.

The observation that tight output coupling does not necessarily lead to collaborative behaviors (and can even lead to competition, which is what seemed to happen in this research) calls into question some of the fundamental assumptions underlying the Single-Display Groupware (SDG) paradigm. SDG systems provide one large shared display, augmented with multiple means of providing input (e.g., multiple mice) so that all members of a co-located group can synchronously use the system[34]. The earliest studies of SDG systems did not really examine whether or not SDG systems promoted effective collaboration; they just pitted SDGs against more traditional, single input desktops and found (not surprisingly) that the users were more equitably engaged and reported a preference for the SDG system (Inkpen et al., 1995; Stacey D. Scott et al., 2002; Stewart et al., 1999; Stewart et al., 1998). The underlying assumption is that being privy to the tightly-coupled shared output would encourage users to collaborate (the degree of input coupling was never explicitly discussed, but was loose in all cases, allowing users to engage in independent actions within the shared space).

This research is not the first to find than an SDG system with tight output coupling but loose input coupling fails to promote collaboration. At least two other studies noticed similar outcomes, one in the context of a shared drawing program (Benford et al., 2000), and another in the context of a document layout task (Birnholtz et al., 2007). Both noticed that SDG users often used the system in a parallel rather than collaborative fashion, and in the document layout task, seemed to be more motivated by personal gain than by group gain, behaving in an implicitly competitive fashion, much as the "Simple" users in this research behaved. Both of these research groups tried to improve collaboration by experimenting with

[34] MMUIs can be thought of as being a special case within the class of SDG systems, where the means of providing input are themselves computational devices with output displays.

coupling the user input more tightly, although they did not phrase it as such. Birnholtz et al. tried the extreme of input coupling, replacing the multiple mice that were providing input to the SDG with a single mouse to be shared by all members of the group in an attempt to enforce collaboration. Their approach took a metaphorical two steps forward and one step backwards, for they found that while it increased group discussions, it also increased frustration of the group members and encouraged lopsided participation. Benford et al. took a more relaxed approach to input coupling, allowing user inputs to modify each other to produce effects in the drawing program that would be impossible to attain solo. They reported observing improved collaboration, but no formal evaluation was performed, and it is hard to project how their input coupling approach could be adapted for use in non-generative tasks.

This research done here suggests a third approach – rather than try to manipulate the degree of *input* coupling to forestall emergent competition among users of a Single-Display Groupware system, loosen the *output* coupling. The "Complex" condition was identical to the "Simple" condition in all ways aside from the degree of output coupling, but the "Complex" users showed none of the sensitivity to monitoring that seemed to be driving the emergent competitive behaviors of the "Simple" users. This leads to the second Design Recommendation to emerge from this work:

8.2.2 Design Recommendation: Coupling, Shared Displays, and Competition

8.2.2.1 Context

When designing for:

- Synchronous, co-located activity with a shared display
- Cooperative activity with homogeneous roles
- No externally-imposed task division
- Loosely-coupled input

8.2.2.2 Recommendation

To prevent emergent competition, loosely-couple the output.

8.2.3 Coupling and Privacy

The results of this work show that loosening output coupling works to forestall emergent competition, but it was not structured so that explanations could be offered as to why it worked to forestall competition. Perhaps in addition to the degree of monitoring, *constancy* of monitoring is a prerequisite for competition – Simple" O-UI users essentially perform *all* of their actions while viewing the Shared Display. "Complex" O-UI users, on the other hand, provide at least some of their input actions while viewing the O-UI, and not the Shared Display. Not only does this break up the monitoring of the Shared Display, it also places the user's attention in the context of a "private" workspace, despite the fact that the image on the O-UI is really just a small region of the larger shared space.

Several groupware researchers have touted the need for separate private and public spaces to support group work (Elwart-Keys, Halonen, Horton, Kass, & Scott, 1990; Mandviwalla & Olfman, 1994; Stacey D. Scott, Grant et al., 2003; Shoemaker & Inkpen, 2001), although the rationale seldom gets much deeper than “that is how people work together in the real world,” or “people don’t always want to be on display” for reasons of embarrassment. This may tie in with the gender effect on participation that was observed with “Simple” groups, but not with “Complex” groups. Female students have long been observed to participate less in public venues like classrooms than their male counterparts (Wilkinson & Marrett, 1985). Several female users who were observed to only tentatively engage with the simulation while using the “Simple” O-UI seemingly had no reticence to engage with the simulation while using the “Complex” O-UI. This may have been due to the perceived level of privacy offered by the “Complex” O-UI.

One group of CSCL researchers articulated a theory-based explanation for the possible educational utility of dual private and public spaces: they associate private spaces with individual, cognitivist theories of learning, whereas public spaces associated with group, sociocultural theories of learning (Vahey et al., 2007). Moving between the two, then, allows users to engage with multiple styles of learning, a best-of-both-worlds approach. While interacting with the private interfaces, users get a chance to try things out, and they then can take their individually-developed ideas or actions and share them with companions when they are fully-formed. This may have been a contributing reason for the larger proportion of *on-task* and *functional* remarks made by “Complex” users – the “private” time allowed them to better focus their thoughts.

8.3 Future Work

8.3.1 Visual Attention Management Mechanisms

The first design recommendation to emerge from this work acknowledges the need for a mechanism to help O-UI users better manage their visual attention behaviors, but acknowledges that how to do so is still very much an open question. One group of researchers working with MMUI systems found that when they displayed passive highlighted indicators on the private display to cue users to direct their attention back to the shared display, the indicators were misunderstood or ignored (Sugimoto et al., 2004). In some of the formative work performed for this research, pop-ups on the O-UIs were found irritating. Attempting to force user attention to the shared display, however, by requiring them to press buttons on the shared display were found to be highly disruptive and annoying (Sugimoto et al., 2004). Another research group employed an animated avatar used to help museum visitors move their attention from a movie playing on a PDA to a movie playing on a wall-mounted display (Kruppa & Aslan, 2005), but it may be too similar to an animated paperclip for users to fully embrace.

Rather than using a “stick” approach, and attempting to force users to shift their visual attention at moments that suit the designers, it might be fruitful to pursue a “carrot” approach, by “fissioning” task-critical information between the displays. Akin to “output fissioning” (Kray et al., 2004), “information fissioning” would attempt to divide up information that needs to be consulted on a fairly regular schedule, so that the user would be motivated to look at the shared display more frequently (and perhaps gain some awareness of his or her partners' actions as a side effect).

8.3.2 Gender and Participation Equity

The "private" workspace offered by "Complex" O-UIs may have encouraged participation equity by increasing participation among group members who would otherwise shy away from "performing" in a public space. This can be confirmed by manipulating the degree of "privacy" in a MMUI setup, from keeping participants' contributions totally anonymous on one end of the spectrum, to making the actions (and the actor behind them) very prominently and publicly indicated on the public display.

8.3.3 Tight Strategic Coupling of "Simple" O-UI

This work demonstrated that, contrary to the theory underlying the concept of coupling, the degree of collaboration clearly does not increase linearly with the degree of output coupling. The research in this study did not vary the input coupling, however, although other work has suggested that tightening input coupling can have the effect of making users feel restricted (Birnholtz et al., 2007). This may be due to the way input coupling is conceived of in CSCW and CSCL literature. When the term "tight input coupling" is used, the assumption is that it is a low-level, operational coupling of input, as when two users are needed to press a key simultaneously. This sort of input coupling takes place at the *tactical* level, as it is a "mode of procedure" as the Merriam-Webster dictionary defines "tactics." For obvious reasons, tight input coupling of input at the tactical level can be cumbersome, and is generally not recommended – frustration is the usual result, as exemplified by (Birnholtz et al., 2007). "Strategy" operates at a larger level of granularity than tactics, however, a "careful plan or method" (Merriam-Webster again) which employs smaller-granularity tactics. Many CSCL applications relied on activity designs, like the jigsawing method of task division, to encourage participation in collaborative activities (see Section 4.3 Computer-Supported Collaborative Learning (CSCL)). Such activity structures are, essentially, a tight coupling of input at the *strategic* level, although it was never labeled as such in the CSCL literature. A review of prior work using the "strategic input coupling" lens, suggests that tight input coupling at the *strategic* level may be a better way to encourage collaboration amongst SDG users than tight coupling at the tactical level.

The constant joint monitoring of the Shared Display when using a "Simple" O-UI seems to encourage competition among group members. Can this be alleviated by adding tight *strategic* input coupling? The original design for *MUSHI-Lignancy* called for visitor to take on different, complementary roles, and prior to leaving the Exploratorium, over 60 visitors participated in a formative Phase III of this study, which employed the multi-role version of *MUSHI-Lignancy*. This time, the exhibit was placed on the floor of the Exploratorium in the midst of the life science region of the museum. The data analysis for Phase III, which involved audio transcripts and video coding similar to that of Phase II, is ongoing, but should shed some light on the concept of strategic input coupling.

APPENDICES

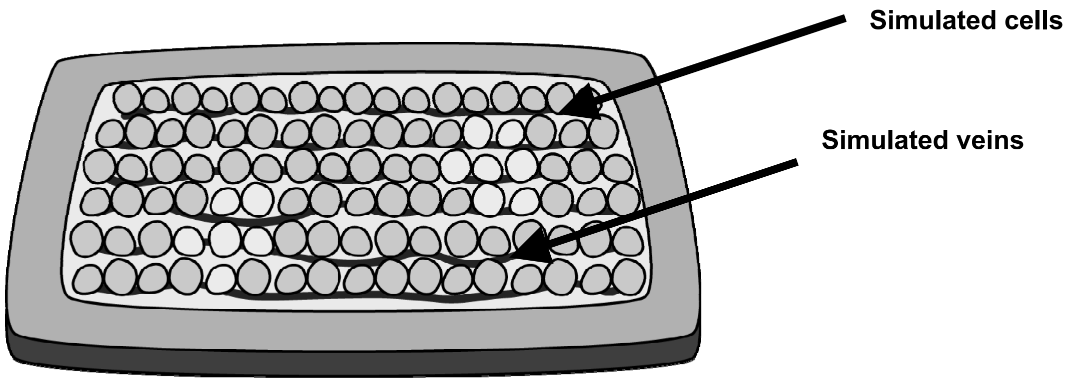
APPENDIX A
Protocol script used in Phase II of this study

Hello. My name is Leilah Lyons. I am a graduate student at the University of Michigan. I study how to make games for museums and schools. I do not work for the Exploratorium itself, but they have given me permission to conduct a study here. My study is on an exhibit that allows visitors to play an educational computer game together. It is similar to many of the exhibits here at the Exploratorium because it is very interactive, and it deals with scientific medical topics.

If you decide to participate, you will be asked to play the game and fill a short questionnaire. You might also be asked to answer a few short questions after the game to help us understand what happened while you played. The entire experience should take between 10 and 20 minutes. You are free to stop participating at any time, and you do not have to answer any written or verbal questions that you do not wish to answer. Do you understand?

[If the visitor says yes, proceed. Otherwise, address any misunderstandings.]

The subject of this game is how doctors treat cancer. There are no pictures or images from real life – everything in the exhibit looks sort of like a cartoon, like this.



In the game you will be asked to engage in simulated tasks, like performing surgery. Sometimes, the simulated patient may die. If this makes you uncomfortable, and you do not want to participate, please tell me now. Do you still wish to participate?

[If participant says no, end interview.]

Please give me your signed consent and assent forms now.

[If the visitor has already provided signed consent and assent forms, proceed, otherwise, ask the minor and his/her guardian to sign the forms.]

[Assign unique, randomly-assigned ID number to participant]

The large screen on the table is showing us a simulation of cells in a patient's body. Some of these cells are healthy, the pink cells that you see here…

[Points at pink cells on simulation screen]

…but some are cancerous, like the grayish-green cells you can see here...

[Points at green cells on simulation screen]

The cancer cells will try to spread. If the cancer cells spread too much, the patient will die. Your job is to help fight the cancer, but to do so without hurting the patient's healthy cells too much. You will each play the role of a doctor, and will be able to perform surgery on the patient. Work together to keep the patient alive. Does anyone have any questions about what I just explained?

[Wait for questions]

Remember, during the game, if any one of you has a question, or just wants to stop playing, please raise your hand at any time. I am now going to give each one of you a nametag with a letter on it, either 'N', 'W', 'E', or 'S', just so we can keep track of who is who without knowing your names.

[Researcher doles out sticky nametags labeled with 'N', 'W', 'E', or 'S' written on them, to each of the participants]

[Researcher should take note of the consent/assent forms regarding audiotaping and videotaping before making one of these statements]

You have all assented to being videotaped, and your parents have all given you permission, so I will start taping now. Remember, you can ask me to stop taping at any time. Experiment number ______ begins.

[Hold up paper with experiment # written on it in front of the camera]

OR

You have all assented to being audiotaped, and your parents have all given you permission, so I will start taping now. Remember, you can ask me to stop taping at any time. Experiment number ______ begins.

OR

Not everyone has consented to being recorded, so we will not video or audiotape you today.

All right, we're ready to begin! Each one of you pick up a handheld and press the buttons onscreen to join the simulation. The screen will give you instructions for what to do next.

APPENDIX B
Questionnaire used in Phase II of this study

Center for Highly-Interactive Computing in Education
College of Engineering, Dept. of EECS
University of Michigan
2260 Hayward St.
Ann Arbor, MI 48109-2121
Contact: Leilah Lyons, 734-274-1412, ltoth@umich.edu

Studying Cooperation in a Multi-User Computer-based Museum Exhibit

Post-Experiment Questionnaire

If any of the questions make you uncomfortable, or if you do not wish to answer them for any reason, you do not have to do so.

For each row, please circle the answer closest to your opinion. We use the term 'partner' to describe the other participants who played the game with you.

1.	**My** actions **helped** the group fight cancer.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
2.	Sometimes, **my** actions **hurt** how the group fought cancer.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
3.	My **partners'** actions **helped** the group fight cancer.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
4.	Sometimes, my **partners'** actions **hurt** how the group fought cancer.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
5.	I was aware of **how well my partners were** doing at all times.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
6.	There were times during the game where I needed to coordinate one of **my actions** with one of **my partner's actions**.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
7.	I was aware of **how well our group was doing** at all times.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
8.	I wish some of my **partners** had **behaved differently** during the game.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
9.	**I** was able to **contribute as much or more** as my partners did.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know

Postquestionnaire 1 of 4
Experiment #: ________________
Participant #: ________________
Date: ____/____/____
Time: ____:____AM / PM

In your own words, describe **what you did** during the game.

For example, what actions could you perform within the game? Did you use a strategy?

10.

Did any of your partners seem to use a different strategy from the one you used? If so, identify the player by their letter (**N**, **S**, **E** or **W**) and briefly describe how their strategy differed from yours.

11.

Please rate yourself and your partners on a scale from **1 (poor)** to **5 (best)**:

			Poor …….. Best
12.	How **skilled** was each player at playing the game?	N	1 2 3 4 5
		S	1 2 3 4 5
		E	1 2 3 4 5
		W	1 2 3 4 5
13.	How well did each player try to **organize the actions of the group**?	N	1 2 3 4 5
		S	1 2 3 4 5
		E	1 2 3 4 5
		W	1 2 3 4 5

14. Which player **helped you the most** during the game?

a. **N**

a. **S**

a. **E**

a. **W**

e. **(none of the above)**

Postquestionnaire 2 of 4
Experiment #: ________________
Participant #: ________________
Date: ____/____/____
Time: ____:____ AM / PM

For each row, please circle the answer closest to your opinion.

15.	I would have had more fun if I could have played the game **by myself**.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
16.	I would have had more playing the game fun if I was **competing** against the other participants.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
17.	I would play the game again.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
18.	I felt the game took a long time to learn.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
19.	I feel I learned something about science from the game.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
20.	I wish the game was more challenging.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
21.	The game should have lasted longer.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know

22.	I feel that this kind of game is appropriate for a science museum like the Exploratorium.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know

23. Please explain your answer:

24.	I feel that computer games are appropriate for a science museum like the Exploratorium	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know

25.Please explain your answer:

Postquestionnaire
Experiment #: _______________
Participant #: ________________
Date: ____/____/____
Time: ____:____ AM / PM

Now we will ask you a few questions about what you know or may have learned while playing the game.

26.	Cancer cells grow in only one area of the body.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
27.	The more blood a cancer cell has access to, the more it will grow.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
28.	It is easier to perform surgery on scattered cancer cells than clustered cancer cells.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
29.	It is possible to cure cancer without hurting healthy cells.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
30.	All cancer is hard to kill.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
31.	Doctors need to coordinate their actions to fight cancer well.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
32.	Healthy cells and cancer cells reproduce at the same rate.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know
33.	It's useful to kill healthy cells if they are near cancer cells.	Strongly Disagree	Disagree a little	Agree a little	Strongly Agree	Don't know

34. How did the cancer cells behave in the game?

35. What did you find most surprising, if anything, about the game?

35. If you are in school, what grade are you in?
(If not, list highest level of education completed)

What age are you?

____________________________________ ________________

Postquestionnaire 4 of 4
Experiment #: _______________
Participant #: _______________
Date: ____/____/____
Time: ____:____ AM / PM

BIBLIOGRAPHY

Abowd, G. D., Atkeson, C. G., Hong, J., Long, S., Kooper, R., & Pinkerton, M. (1997). Cyberguide: A mobile context-aware tour guide. *Wireless Networks, 3*(5), 421-433.
Acoustiguide. (n.d.). Acoustiguide: History of Audio Tours. Retrieved 5/12/2008, from http://www.acoustiguide.com/about/history.html
Albertini, A., Brunelli, R., Stock, O., & Zancanaro, M. (2005). Communicating user's focus of attention by image processing as input for a mobile museum guide. In *Proceedings of the 10th international conference on Intelligent user interfaces (IUI '05)* (pp. 299--301), San Diego, California, USA.
Alisi, T. M., Del Bimbo, A., & Valli, A. (2005). Natural interfaces to enhance visitors' experiences. *Multimedia, 12*(3), 80-85.
Allen, S. (2002). Looking for Learning in Visitor Talk: A Methodological Exploration. In G. Leinhardt, K. Crowley & K. Knutson (Eds.), *Learning Conversations in Museums* (pp. 259-303). Mahwah, New Jersey: Lawrence Erlbaum.
Allen, S. (2004). *Finding Significance*. San Francisco: Exploratorium.
Allison, D. K., & Gwaltney, T. (1991). How People Use Electronic Interactives: 'Information Age - People, Information & Technology'. In *Proceedings of Hypermedia & Interactivity in Museums Archives and Museum Informatics* (pp. 7–16), Pittsburgh.
American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.
Aoki, P. M., Grinter, R. E., Hurst, A., Szymanski, M. H., Thornton, J. D., & Woodruff, A. (2002). Sotto voce: exploring the interplay of conversation and mobile audio spaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '02)* (pp. 431--438), Minneapolis, Minnesota, USA.
Argyle, M. (1975). *Bodily Communication*. London: Methuen.
Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The Jigsaw Classroom*. Beverly Hills: Sage.
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.
Atkins, L. (2006). Frames and Games in the Science Museum: A Lens for Understanding Visitor Behavior. In *Proceedings of ICLS 2006*, Bloomington, Indiana, USA.
Back, M., Gold, R., Balsamo, A., Chow, M., Gorbet, M., Harrison, S., et al. (2001). Designing innovative reading experiences for a museum exhibition. *Computer, 34*(1), 80-87.
Bain, R., & Ellenbogen, K. M. (2002). Placing Objects within Disciplinary Perspectives: Examples from History and Science. In S. G. Paris (Ed.), *Perspectives on object-centered learning in museums* (pp. 153-170). Mahwah, N.J. :: L. Erlbaum Associates.
Ballantyne, R., & Packer, J. (2005). Solitary vs. Shared: Exploring the Social Dimension of Museum Learning. *Curator, 48*(2), 177-192.
Barab, S. (2006). Design-Based Research: A Methodological Toolkit for the Learning Scientist. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. New York, NY, USA: Cambridge University Press.
Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology, 72*(5), 593-604.
Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences, 9*(4), 403-436.
Barron, B. (2003). When Smart Groups Fail. *The Journal of the Learning Sciences, 12*(3), 307–359.
Bederson, B. B., Hollan, J. D., Druin, A., Stewart, J., Rogers, D., & Proft, D. (1996). Local tools: an alternative to tool palettes. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology* (pp. 169--170), Seattle, Washington, United States.
Bell, B., Bareiss, R., & Beckwith, R. (1993-1994). 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.
Benford, S., Bederson, B. B., Åkesson, K.-P., Bayon, V., Druin, A., Hansson, P., et al. (2000). Designing storytelling technologies to encouraging collaboration between young children. In *CHI '00:*

Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 556--563), The Hague, The Netherlands.
Benford , S., Rowland, D., Flintham, M., Drozd, A., Hull, R., Reid, J., et al. (2005). Life on the Edge: Supporting Collaboration in Location-Based Experiences. In *Proceedings of CHI 2005* (pp. 231-240), Portland, Oregon, USA.
Benjamin, W. (1969). The Work of Art in the Age of Mechanical Reproduction (H. Zohn, Trans.). In H. Arendt (Ed.), *Illuminations* (pp. 217-251). New York: Schocken Books.
Benta, K.-I. (2005). Affective aware museum guide. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE '05)* (pp. 53-55).
Bier, E. A., & Freeman, S. (1991). MMM: a user interface architecture for shared editors on a single screen. In *Proceedings of the 4th annual ACM symposium on User interface software and technology* (pp. 79-86), Hilton Head, South Carolina, United States.
Birnholtz, J. P., Grossman, T., Mak, C., & Balakrishnan, R. (2007). An exploratory study of input configuration and group process in a negotiation task using a large display. In *Proceedings of CHI 2007* (pp. 91-100), San Jose, CA, United States.
Blandin, P. (2002). Case Study: The Grand Galerie De L'Evolution. In B. Lord & G. D. Lord (Eds.), *The Manual of Museum Exhibitions* (pp. 479-482). Walnut Creek, CA: AltaMira Press.
Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualizations. *Learning and Instruction, 14*(3), 325–341.
Boehm, B. (1988). A Spiral Model of Software Development and Enhancement. *Computer, 21*(5), 61-72.
Borun, M. (2002). Object-Based Learning and Family Groups. In S. G. Paris (Ed.), *Perspectives on object-centered learning in museums* (pp. 245-261). Mahwah, N.J. :: L. Erlbaum Associates.
Borun, M., Chambers, M., & Cleghorn, A. (1996). Families are learning in science museums. *Curator, 39*(2), 262-270.
Boyd, W. I. (1999). Museums as Centers of Controversy. *Daedalus, 128*(3), 185-228.
Bressler, D. M. (2005). *Science Now, Science Everywhere: Liberty Science Center's Mobile Learning Companion* (White paper): Liberty Science Center.
Bricker, L. J., Baker, M., Fujioka, E., & Tanimoto, S. (1998). *Colt: A System for Developing Software that Supports Synchronous Collaborative Activities*: University of Washington Technical Report.
Brightman, J. (2006, April 17). Study: Women Gamers Outnumber Men in 25-34 Age Group. *GameDaily*, from http://www.gamedaily.com/articles/features/study-women-gamers-outnumber-men-in-25-34-age-group/68821/?biz=1
Brøderbund. (1985). Where in the World is Carmen Sandiego? [Computer Software].
Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *The Journal of the Learning Sciences, 2*(2), 141-178.
Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in Learning: New Environments in Education* (pp. 289-325). Mahwah, NJ, USA: Erlbaum.
Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.
Bruner, J. S. (1961). The Act of Discovery. *Harvard Educational Review, 31*, 21–32.
Bruns, E., Brombach, B., Zeidler, T., & Bimber, O. (2007). Enabling Mobile Phones To Support Large-Scale Museum Guidance. *Multimedia, 14*(2), 16-25.
Callahan, P. (n.d.). What is the Game of Life? Retrieved 6/15/2008, from http://www.math.com/students/wonders/life/life.html
Callanan, M. A., Jipson, J. L., & Soennichsen, M. S. (2002). Maps, Globes, and Videos: Parent-Child Conversations about Representational Objects. In S. G. Paris (Ed.), *Perspectives on object-centered learning in museums* (pp. 261-284). Mahwah, N.J. :: L. Erlbaum Associates.
Cameron, F. (2003). Digital Futures I: Museum Collections, Digital Technologies, and the Cultural Construction of Knowledge. *Curator, 46*(3), 325-340.
Carlisle, R. W. (1985). What Do School Children Do at a Science Center? *Curator, 28*(1).
Chang, C. Y., 20 (1), 4-10. (2004). Could a laptop computer plus the liquid crystal display projector amount to improved multimedia geoscience instruction? *Journal of Computer Assisted Learning, 20*(1), 4-10.

Cheverst, K., Davies, N., Mitchell, K., Friday, A., & Efstratiou, C. (2000). Developing a context-aware electronic tourist guide: some issues and experiences. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 17--24), The Hague, The Netherlands.

Chou, L.-D., Lee, C.-C., Lee, M.-Y., & Chang, C.-Y. (2004). A Tour Guide System for Mobile Learning in Museums. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04)* (pp. 195-196), Taoyuan, Taiwan.

Chou, L.-D., Wu, C.-H., Ho, S.-P., Lee , C.-C., & Chen, J.-M. (2004). Requirement analysis and implementation of palm-based multimedia museum guide systems. In *Proceedings of 18th International Conference on Advanced Information Networking and Applications (AINA '04)* (pp. 352-357).

Chou, S.-C., Hsieh, W.-T., Gandon, F. L., & Sadeh, N. M. (2005). Semantic Web technologies for context-aware museum tour guide applications. In *Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA '05)* (pp. 709-714).

Ciolfi, L., Fernstrom, M., Bannon, L. J., Desbpande, P., Gallagher, P., McGettrick, C., et al. (2007). The Shannon Portal Installation: Interaction Design for Public Places. *Computer 40*(7), 64-71.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher, 32*(1), 9-13.

Cohen, E. G. (1994). *Designing groupwork: Strategies for heterogeneous classrooms*. New York: Teachers College Press.

Cole, H., & Stanton, D. (2003). Designing mobile technologies to support co-present collaboration. *Personal Ubiquitous Computing, 7*(6), 365-371.

Colella, V. (2000). Participatory Simulations: Building collaborative understanding through immersive dynamic modeling. *Journal of the Learning Sciences, 9*(4), 471-500.

Collins, A. (2006). Cognitive Apprenticeship. In K. Sawyer (Ed.), *Handbook of Learning Sciences.* New York: Cambridge University Press.

Conn, S. (1998). *Museums and american intellectual life, 1876-1926*. Chicago: The University of Chicago Press.

Cooper, J. (1993). Engaging the [Museum] Visitor: Relevance, Participation & Motivation in Hypermedia Design. In *Proceedings of International Conference on Hypermedia and Interactivity in Museums (ICHIM)*, Cambridge, England.

Cortez, C., Nussbaum, M., Santelices, R., Rodríguez, P., Zurita, G., Correa, M., et al. (2004). Teaching Science with Mobile Computer Supported Collaborative Learning (MCSCL). In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04)* (pp. 67-74), Taoyuan, Taiwan.

Crane, V., Nicholson, H., Chen, M., & Bitgood, S. (1994). *Informal science learning : what the research says about television, science museums, and community-based projects*. Dedham, MA :: Research Communications Ltd.

Crevier, D. (1993). *AI: The Tumultuous Search for Artificial Intelligence*. New York: Basic Books.

Crowley, K., & Callanan, M. A. (1998). Describing and supporting collaborative scientific thinking in parent-child interactions. *Journal of Museum Education, 23*(1), 12-17.

Crowley, K., & Eberbach, C. (2005). From Living to Virtual: Learning from Museum Objects. *Curator, 48*(3), 317-338.

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* (pp. 333-356). Mahwah, New Jersey: Lawrence Erlbaum.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM, 35*(6), 64--72.

Csikszentmihalyi, M. (1993). Why We Need Things. In S. Lubar & W. D. Kingery (Eds.), *History from Things: Essays on Material Culture* (pp. 20-29). Washington and London: Smithsonian Institution Press.

Csikszentmihalyi, M., & Hermanson, K. (1995). Intrinsic Motivation in Museums: Why Does One Want to Learn? In J. H. Falk & L. D. Dierking (Eds.), *Public Institutions for Personal Learning: Establishing a Research Agenda* (pp. 67-77). Washington. D.C.: American Association of Museums.

Curtis, M., Luchini, K., Bobrowsky, W., Quintana, C., & Soloway, E. (2002). Handheld use in K-12: a descriptive account. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02)* (pp. 23-30), Växjö, Sweden.

Cuthbert, A. J. (1999). Designs for collaborative learning environments: can specialization encourage knowledge integration? In *CSCL '99: Proceedings of the 1999 conference on Computer support for collaborative learning* (pp. 13), Palo Alto, California.

Danesh, A., Inkpen, K., Lau, F., Shu, K., & Booth, K. (2001). Geney™: Designing a Collaborative Activity for the Palm™ Handheld Computer. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 388-395), Seattle, Washington, United States.

Davis, E. A. (2003). Prompting Middle School Science Students for Productive Reflection: Generic and Directed Prompts. *Journal of the Learning Sciences, 12*(1), 91–142.

Davis, E. A. (2004). Creating Critique Projects. In M. C. Linn, E. A. Davis & P. Bell (Eds.), *Internet Environments for Science Education* (pp. 89 - 113). Mahwah, NJ: Lawrence Erlbaum Associates.

Deng, Y.-C., Do, M.-Z., Chang, L.-J., & Chan, T.-W. (2004). PuzzleView: Enhanced Workspace Displaying for Group Interaction with Tablet PCs. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04)* (pp. 177-181), Taoyuan, Taiwan.

Dewan, P., & Choudhard, R. (1991). Flexible user interface coupling in a collaborative system. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 41--48), New Orleans, Louisiana, United States.

Di Blas, N., Poggi, C., & Reeve, T. C. I. P. o. I., Bloomington, Indiana, USA. (2006). Collaborative Learning in a 3D Virtual Environment: Design Factors and Evaluation Results. In *Proceedings of the International Conference of the Learning Sciences (ICLS '06)* (pp. 127-133), Bloomington, IN, US.

Diamond, J. (1986). The Behavior of Family Groups in Science Museums. *Curator: The Museum Journal, 29*(2), 139-154.

Diamond, J. (1989). Desert Explorations - A Videodisk Exhibit Designed for Flexibility. *Curator, 32*(3), 161-173.

Diamond, J. (1996). Playing and learning. *ASTC Newsletter, 24*(4), 2-6.

Diamond, J. (1999). *Practical Evaluation Guide: Tools for Museums & Other Informal Educational Settings*. Walnut Creek, CA: AltaMira Press.

Diamond, J., Bond, A., Schenker, B., Meier, D., & Twersky, D. (1995). Collaborative multimedia. *Curator, 38*(3), 137-149.

Dietz, P., & Leigh, D. (2001). DiamondTouch: a multi-user touch technology. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology* (pp. 219--226), Orlando, Florida.

DiGiano, C., Yarnall, L., Patton, C., Roschelle, J., Tatar, D., & Manley, M. (2002). Collaboration Design Patterns: Conceptual Tools for Planning for The Wireless Classroom. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02)* (pp. 39-47), Växjö, Sweden.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative Learning: Cognitive and Computational Approaches* (pp. 1-15). Oxford: Elsevier.

DiMicco, J. M., Hollenbach, K. J., Pandolfo, A., & Bender, W. (2007). The Impact of Increased Awareness While Face-to-Face. *Human–Computer Interaction, 22*, 47-96.

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.

Edelson, D. C., & Reiser, B. J. (2006). Making authentic practices accessible to learners: design challenges and strategies. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. New York: Cambridge University Press.

Ellenbogen, K. M. (2002). Museums in Family Life: An Ethnographic Case Study. In G. Leinhardt, K. Crowley & K. Knutson (Eds.), *Learning Conversations in Museums* (pp. 81-101). Mahwah, New Jersey: Lawrence Erlbaum.

Elwart-Keys, M., Halonen, D., Horton, M., Kass, R., & Scott, P. (1990). User Interface Requirements for Face to Face Groupware. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '90)* (pp. 295-301).

Evans, E. M., Mull, M. S., & Poling, D. A. (2002). The authentic object? A child's-eye view. In S. G. Paris (Ed.), *Perspectives on object-centered learning in museums*. (pp. 55-77). Mahwah, NJ: Lawrence Erlbaum Associates.

Exploratorium (2001). Electronic Guidebook Forum Report [Online]. *Journal*, 92. Retrieved from http://www.exploratorium.edu/guidebook/

Exploratorium (2005). Electronic Guidebook Forum [Online]. *Journal*, 66. Retrieved from http://www.exploratorium.edu/guidebook/

Exploratorium. (2008). 2007-2008 Fact Sheet. Retrieved 5/12/2008, from http://www.exploratorium.edu/pr/presskits/General_Info/FY07_Fact_Sheet.pdf

Exploratorium. (n.d.-a). Floor Map of the Exploratorium. Retrieved 5/12/2008, from http://www.exploratorium.edu/visit/images/floormap.jpg

Exploratorium. (n.d.-b). The Fountain. Retrieved 5/12/2008, from http://www.exploratorium.edu/history/palace/fountain.html

Falk, J. H. (2006). An Identity-Centered Approach to Understanding Museum Learning. *Curator, 49*(2), 151-166.

Falk, J. H., Brooks, P., & Amin, R. (2001). Investigating the Role of Free-Choice Learning on Public Understanding of Science: The California Science Center LASER Project. In *Free-Choice Science Education: How We Learn Science Outside School*. Danvers, MA.

Falk, J. H., & Dierking, L. D. (1992). *The museum experience*. Washington, DC: Whalesback Books.

Falk, J. H., & Dierking, L. D. (2000). *Learning from museums : visitor experiences and the making of meaning*. Walnut Creek, CA :: AltaMira Press.

Falk, J. H., & Dierking, L. D. (Eds.). (1995). *Public Institutions for Personal Learning: Establishing a Research Agenda*. Washington. D.C.: American Association of Museums.

Falk, J. H., Dierking, L. D., & Holland, D. G. (1995). How Should We Investigate Learning in Museums? Research Questions and Project Designs. In J. H. Falk & L. D. Dierking (Eds.), *Public Institutions for Personal Learning: Establishing a Research Agenda* (pp. 23-30). Washington. D.C.: American Association of Museums.

Falk, J. H., Koran, J. J., Dierking, L., & Dreblow, L. (1985). Predicting visitor behavior. *Curator, 24*(8), 249-257.

Falk, J. H., Scott, C., Dierking, L., Rennie, L. J., & Cohen Jones, M. (2004). Interactives and Visitor Learning. *Curator, 47*(2), 171-198.

Farooq, U., Schafer, W., Rosson, M. B., & Carroll, J. M. (2002). M-Education: bridging the gap of mobile and desktop computing. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02)*, Växjö, Sweden.

Fehrer, E. (1990). Interactive museum exhibits as tools for learning: Exploration with light. *International Journal of Science Education, 12*, 35-39.

Fehrer, E., & Rice, K. (1985). Development of Scientific Concepts Through the Use of Interactive Exhibits in a Museum. *Curator, 28*(1).

Fleck, M., Frid, M., Kindberg, T., O'Brien-Strain, E., Rajani, R., & Spasojevic, M. (2002). From informing to remembering: ubiquitous systems in interactive museums. *Pervasive Computing, 1*(2), 13-22.

Forlines, C., Esenther, A., Shen, C., Wigdor, D., & Ryall, K. (2006). Multi-user, multi-display interaction with a single-user, single-display geospatial application. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology* (pp. 273--276), Montreux, Switzerland.

Friedman, N., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General, 129*, 1-23.

Frost, C. O. (2002). When the Object is Digital: Properties of Digital Surrogate Objects and Implications for Learning. In S. G. Paris (Ed.), *Perspectives on object-centered learning in museums* (pp. 79-94). Mahwah, N.J. :: L. Erlbaum Associates.

Fujimoto, T., & Matsuo, T. (2005). An assessment - navigation support system in social educational facilities. In *Proceedings of the 2005 International Conference on Active Media Technology (AMT '05)* (pp. 271), Takamatsu, Kagawa, Japan.

Fussell, S. R., Kraut, R. E., & Siegel, J. (2000). Coordination of communication: Effects of shared visual context on collaborative work. . In *Proceedings of CSCW 2000* (pp. 21-30).

Gammon, B. (1999). Visitors' use of computer exhibits: Findings from 5 grueling years of watching visitors getting it wrong. *Informal Learn. Rev. [Online], 38*.

Ganoe, C. (2002). Supporting the collaborative meeting place. In *CHI '02 extended abstracts on Human factors in computing systems*, Minneapolis, Minnesota, USA.

Gardner, M. (1970, October, 1970). Mathematical Games: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American, October, 1970*.

Gee, J. (2003). *What Video Games Have to Teach us About Learning and Literacy*. New York, NY: Palgrame Macmillan.

Gelman, R., Massey, C., & McManus, M. (1991). Characterizing supporting environments for cognitive development: Lessons from children in a museum. In L. B. Resnick, J. Levine & S. Teasley (Eds.), *Perspectives on socially-shared cognition* (pp. 226-256). Washington, D.C.: American Psychological Association.

Gergle, D., Kraut, R. E., & Fussell, S. R. (2006). The Impact of Delayed Visual Feedback on Collaborative Performance. In *Proceedings of the 2006 Conference on Human Factors in Computing Systems (CHI '06)* (pp. 1303-1312), Montreal, CA.

Goldman, S., Pea, R., & Maldonado, H. (2004). Emerging Social Engineering in the Wireless Classroom. In *Proceedings of the International Conference on the Learning Sciences (ICLS '04)*, Santa Monica, CA.

Goldstone, R. L. (2006). The Complex Systems See-Change in Education. *Journal of the Learning Sciences, 15*(1), 35-43.

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology, 46*, 414-466.

Greenberg, S., Boyle, M., & Laberge, J. (1999). PDAs and Shared Public Displays: Making Personal Information Public, and Public Information Personal. *Personal Technologies, 3*(1), 53-64.

Greeno, J. G. (2006). Learning in Activity. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. New York: Cambridge University Press.

Grinter, R. E., Aoki, P. M., Hurst, A., Szymanski, M. H., Thornton, J. D., & Woodruff, A. (2002). Revisiting the visit: understanding how technology can shape the museum visit. In *Proceedings of CSCW* (pp. 146-155), New Orleans, Louisiana, United States,.

Grudin, J. (1994). Computer-supported cooperative work: history and focus. *Computer, 27*(5), 19 - 26.

Gurian, E. H. (1999). What is the object of this exercise? A meandering exploration of the many meanings of objects in museums. *Daedalus, 128*(3), 163-183.

Haneef, A. M., & Ganz, A. (2002). Mobile agent based network access for mobile electronic guidebooks. In *International Mobility and Wireless Access Workshop (MobiWac '02)* (pp. 22-25).

Hardy, R., & Rukzio, E. (2008). Direct Touch-based Mobile Interaction with Dynamic Displays. In *CHI08 Workshop: Designing and Evaluating Mobile Phone-Based Interaction with Public Displays*, Florence, Italy.

Harrison, S., Minneman, S., & Balsamo, A. (2001). Methods & tools: how to XFR:"eXperiments in the future of reading". Genre as a Way of Design. *interactions, 8*(3), 31--41.

Heartney, E. (1997, August 28, 1997). Can Museums Survive Popularity? *Ann Arbor News*,

Hein, G. E. (1998). *Learning in the museum*. London ; New York :: Routledge.

Hertz-Lazarowitz, R. (1992). Understanding Interactive Behaviors: Looking at Six Mirrors of the Classroom. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups. The theoretical anatomy of group learning* (pp. 174–199). New York, NY: Cambridge University Press.

Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding Complex Systems: Some Core Challenges. *Journal of the Learning Sciences, 15*(1), 53-61.

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. . *Cognitive Science, 28*(1), 127-138.

Holland, J. (1998). *Emergence: From Chaos to Order*. New York, NY: Basic Books.

Hooper-Greenhill, E. (2000). Objects and Interpretive Processes. In E. Hooper-Greenhill (Ed.), *Museums and the Interpretation of Visual Culture* (pp. 103-123). London: Routledge.

Hornecker, E., & Stifter, M. (2006a). Digital backpacking in the museum with a SmartCard. In *CHINZ '06: Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction* (pp. 99--107), Christchurch, New Zealand.

Hornecker, E., & Stifter, M. (2006b). Learning from interactive museum installations about interaction design for public settings. In *Proceedings of the 20th conference of the computer-human*

interaction special interest group (CHISIG) of Australia on Computer-human interaction: design: activities, artefacts (pp. 135 - 142), Sydney, Australia.
Hourcade, J. P., Bederson, B. B., Druin, A., & Taxén, G. (2002). KidPad: collaborative storytelling for children. In *CHI '02: CHI '02 extended abstracts on Human factors in computing systems* (pp. 500--501), Minneapolis, Minnesota, USA.
Hsi, S. (2002). The Electronic Guidebook: A Study of User Experiences using Mobile Web Content in a Museum Setting. In *Proceedings of the International Workshop on Wireless and Mobile Technologies in Education (WMTE '02)* (pp. 48-54), Växjö, Sweden.
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.
Hsi, S. (2004). I-Guides in Progress: Two Prototype Applications for Museum Educators and Visitors Using Wireless Technologies to Support Informal Science Learning. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04)*, Taoyuan, Taiwan.
Hsi, S., Semper, R., Brunette, W., Rea, A., & Borriello, G. (2004). eXspot: A Wireless RFID Transceiver for Recording and Extending Museum Visits. In *Ubicomp 2004 Demonstration*, Nottingham, England
Huang, Y.-P., Chuang, W.-P., Chang, Y.-T., & Sandnes, F. E. (2007). Enhanced Interactivity in Learning-Guide Systems with RFID. In *First Annual RFID Eurasia* (pp. 1-6).
Hudson, J. M., & Bruckman, A. S. (2004). The Bystander Effect: A Lens for Understanding Patterns of Participation. *Journal of the Learning Sciences, 13*(2), 165-195.
Humphrey, T., Gutwill, J., & Team, E. A. (2005). *Fostering Active Prolonged Engagement: The Art of Creating APE Exhibits*. San Francisco: Exploratorium.
Impey, O., & MacGregor, A. (Eds.). (2001). *The Origins of Museums: The Cabinet of Curiosities in Sixteenth- and Seventeenth-century Europe*. Thirsk, North Yorkshire: House of Stratus.
Inkpen, K., Booth, K. S., Gribble, S. D., & Klawe, M. (1995). Give and take: children collaborating on one computer. In *CHI '95: Conference companion on Human factors in computing systems* (pp. 258--259), Denver, Colorado, United States.
Inkpen, K., McGrenere, J., Booth, K. S., & Klawe, M. (1997). The Effect of Turn-Taking Protocols on Children's Learning in Mouse-Driven Collaborative Environments. In *Proceedings of Graphics Interface '97*, Kelowna, BC, CA.
Jacobson, M. J. (2000). Butterflies, traffic jams, and cheetahs: Problem solving and complex systems. Paper presented at the *2000 annual meeting of the American Educational Research Association (AERA 2000)*, New Orleans, LA.
Jacobson, M. J. (2001). Problem solving, cognition, and complex systems: Differences between experts and novices. *Complexity, 6*(3), 41-49.
Jacobson, M. J., & Wilensky, U. (2006). Complex Systems in Education: Scientific and Educational Importance and Implications for the Learning Sciences. *Journal of the Learning Sciences, 15*(1), 11-34.
Jensen, J. J., & Skov, M. B. (2005). A review of research methods in children's technology design. In *IDC '05: Proceedings of the 2005 conference on Interaction design and children* (pp. 80--87), Boulder, Colorado.
Johansen, R. (1988). *Groupware. Computer support for Business Teams*. New York: The Free Press.
John, B. E., & Kieras, D. E. (1996). Using GOMS for User Interface Design and Evaluation: Which Technique? *ACM Transactions on Computer-Human Interaction (TOCHI), 3*(4), 287–319.
Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* Cambridge, MA: Cambridge University Press.
Kapur, M., & Kinzer, C. (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.
Kaynar, I., Pasek, Z., & Lyons, L. (2004). Creating an Informal Engineering Education Experience: Interactive Manufacturing Exhibit. In *Proceedings of International Conference on Engineering Education (ICEE '04)*, Gainesville, FL.
Kerawalla, L., Pearce, D., Yuill, N., Luckin, R., & Harris, A. (2008). ''I'm keeping those there, are you?'' The role of a new user interface paradigm – Separate Control of Shared Space (SCOSS) – in the collaborative decision-making process. *Computers & Education, 50*(1), 193–206.

Ketelhut, D. (2007). The Impact of Student Self-efficacy on Scientific Inquiry Skills: An Exploratory Investigation in River City, a Multi-user Virtual Environment. *Journal of Science Education and Technology, 16*(1), 99-111.

Klopfer, E. (2003). Technologies to Support the Creation of Complex Systems Models – Using StarLogo Software with Students. *Biosystems, 71*(1-2), 111-123.

Klopfer, E., Perry, J., Squire, K., & Jan, M.-F. (2005a). Collaborative Learning through Augmented Reality Role Playing. In *Proceedings of the 2005 conference on Computer Support for Collaborative Learning (CSCL05)* (pp. 311-315).

Klopfer, E., Perry, J., Squire, K., & Jan, M. (2005b). Mystery at the Museum – A Collaborative Game for Museum Education. In *Proceedings of the conference on Computer Supported Collaborative Learning (CSCL '05)* (pp. 316-320), Taiwan.

Klopfer, E., & Yoon, S. (2005). Using Palm Technology in Participatory Simulations of Complex Systems: A New Take on Ubiquitous and Accessible Mobile Computing. *Journal of Science Education and Technology 14*(3), 287-295.

Klopfer, E., Yoon, S., & T., U. (2005). Young Adventurers- Modeling of Complex Dynamic Systems with Elementary & Middle School Students. *Journal of Computers in Math and Science Teaching, 24*(2), 157-178.

Knudsen, E. I. (2007). Fundamental Components of Attention. *Annual Review of Neuroscience, 30*, 57-78.

Korn, R. (1995). An analysis of differences between visitors at natural history museums and science centers. *Curator, 38*(3), 150-160.

Koshizuka, N., & Sakamura, K. (2000). Tokyo University Digital Museum. In *Proceedings of the International Conference on Digital Libraries: Research and Practice* (pp. 85-92), Kyoto, Japan.

Krasner, G., & Popo, S. (1988). A Cookbook for Using the Model-View-Controller User Interface Paradigm in Smalltalk-80. *Journal of Object-Oriented Programming, 1*(3).

Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human Computer Interactions, 18*, 13-49.

Kray, C., Wasinger, R., & Kortuem, G. (2004). Concepts and issues in interfaces for multiple users and multiple devices. In *Workshop on Multi-User and Ubiquitous User Interfaces 2004 (MU3I '04) at the International Conference on Intelligent User Interfaces / Computer-Aided Design of User Interfaces* (pp. 7-12).

Kruppa, M. (2004a). The better remote control – Multiuser interaction with public displays. In *Proceedings of the Workshop on Multi-User and Ubiquitous User Interfaces 2004 (MU3I '04) at the International Conference on Intelligent User Interfaces (IUI '04)*, Madeira, Portugal.

Kruppa, M. (2004b). *A User Study On Parallel Presentations For Inhomogeneous User Groups, Technical Report 563* (No. Technical Report 563): Saarland University.

Kruppa, M., & Aslan, I. (2005). Parallel Presentations for Heterogenous User Groups - An Initial User Study. In M. Maybury, et al. (Ed.), *Intelligent Technologies for Interactive Entertainment: Proceedings of the First International Conference, INTETAIN 2005, Madonna di Campiglio, Italy, November 30 – December 2, 2005, Lecture Notes in Computer Science* (pp. 54-63). Berlin / Heidelberg: Springer.

Kruppa, M., Lum, A., Niu, W., & Weinel, M. (2005). Towards Mobile Tour Guides Supporting Collaborative Learning In Small Groups. In *Workshop on New Technologies for Personalised Information Access at User Modeling* (pp. 54-63), Edinburgh, UK.

Kusunoki, F., Sugimoto, M., & Hashizume, H. (2002). Toward an Interactive Museum Guide System with Sensing and Wireless Network Technologies. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02)*, Växjö, Sweden.

Kyza, E. A., Golan, R., Reiser, B.J., & Edelson D.C. . (2002). Reflective inquiry: Enabling group self-regulation in inquiry-based science using the Progress Portfolio tool. In *Proceedings of CSCL 2002*, Boulder, CO.

Lai, C.-Y., Wu, C.-C., Kao, H.-C., & Chen, S.-M. (2005). Using a Handheld Concept Mapping Tool for Cooperative Learning. In *Proceedngs of 2005 National Educational Computing Conference (NECC05)*, Philadelphia, PA.

Lai, C. Y., & Wu, C. C. (2006). Using handhelds in a Jigsaw cooperative learning environment. *Journal of Computer Assisted Learning, 22*(4), 284 - 297.

Lave, J. (1988). *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. New York: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.
Lee, J., Lyons, L., Vath, R., Kawamura, M., Quintana, C., & Soloway, E. (2005). MUSHI: Demonstrating A Multi-User Simulation with Handheld Integration. Demonstration at *Fourth International Conference for Interaction Design and Children (IDC 2005)*, Boulder, CO, USA.
Leinhardt, G., & Knutson, K. (2004). *Listening in on museum conversations*. Walnut Creek, CA :: AltaMira Press.
Lesh, R. (2006). Modeling Students Modeling Abilities: The Teaching and Learning of Complex Systems in Education. *Journal of the Learning Sciences, 15*(1), 45-52.
Light, P., Foot, T., & Colboum, C. (1997). Collaborative interactions at the microcomputer keyboard. *Educational Psychology, 7*(1), 13-21.
Linn, M. C., Davis, E. A., & Bell, P. (2004). Inquiry and Technology. In M. C. Linn, E. A. Davis & P. Bell (Eds.), *Internet Environments for Science Education* (pp. 3 - 27). Mahwah, NJ: Lawrence Erlbaum Associates.
Liu, L., Hmelo-Silver, C. E., & Marathe, S. (2005). How Representation Matters: Comparing Collaborative Learning with Alternative Versions of Hypermedia. In *Proceedings of the 2005 Computer-Supported Collaborative Learning Conference (CSCL 2005)* (pp. 398 - 402), Taipei, Taiwan.
Liu, L., Marathe, S., & Hmelo-Silver, C. E. (2005). Function before form: An alternative approach to learning about complex systems. Paper presented at the *2005 Annual Meeting of the American Educational Research Association (AERA 2005)* Montreal, QC.
Lord, B., & Lord, G. D. (2002). *The Manual of Museum Exhibitions*. Walnut Creek, CA: AltaMira Press.
Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 157-176.
Luchini, K., Quintana, C., & Soloway, E. (2004). Design Guidelines for Learner-Centered Handheld Tools. In *Proceedings of CHI 2004* (pp. 135-142), Vienna, Austria.
Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279-281.
Lundby, K., Smørdal, O., A., L., & Fjuk, A. (2002). Networked PDAs in a Community of Learners. In *Proceedings of CSCL 2002* (pp. 548-549), Boulder, CO.
Lyons, L., Lee, J., Quintana, C., & Soloway, E. (2006a). MUSHI: A Multi-Device Framework for Collaborative Inquiry Learning. In *Proceedings of the International Conference of the Learning Sciences (ICLS2006)*, Bloomington, IN, USA.
Lyons, L., Lee, J., Quintana, C., & Soloway, E. (2006b). Preliminary Evaluation of a Synchronous Co-located Educational Simulation Framework. In *Extended Abstracts of the 2006 Conference on Human Factors in Computing Systems (CHI 2006)*, Montreal, CA.
Lyons, L., & Pasek, Z. (2005). Beyond Hits: Gauging Visitor Behavior at an On-line Manufacturing Exhibit. In *Proceedings of Museums and the Web 2005 (MOW '05)*, Vancouver, CA.
Lyons, L., & Pasek, Z. (2006). Gauging Visitor Behavior at an Interactive Engineering Exhibit. In *Proceedings of the 2006 Conference of the American Society for Engineering Education (ASEE '06)*, Chicago, IL.
MacDonald, G. F., & Alsford, S. (1995). Canadian Museums and the Representation of Culture in a Multicultural Nation. *Cultural Dynamics, 7*(1), 15-35.
Macedonia, M. (2003). Revitalizing museums with digital technology. *Computer 36*(2), 94-96.
Mandryk, R. L., Inkpen, K. M., Bilezikjian, M., Klemmer, S. R., & Landay, J. A. (2001). Supporting children's collaboration across handheld computers. In *CHI '01 extended abstracts on Human factors in computing systems* (pp. 255-256), Seattle, Washington.
Mandviwalla, M., & Olfman, L. (1994). What Do Groups Need? A Proposed Set of Generic Groupware Requirements. *ACM Transactions on Computer-Human Interaction (TOCHI), 1*(3), 245-268.
Mantyjarvi, J., Paternò, F., Salvador, Z., & Santoro, C. (2006). Scan and tilt: towards natural interaction for mobile museum guides. In *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 191--194), Helsinki, Finland.
Margolis, J., & Fisher, A. (2002). *Unlocking the clubhouse : women in computing*. Cambridge, Mass.: MIT Press.
Martin, L. M. W., & Toon, R. (2005). Narratives in a Science Center: Interpretation and Identity. *Curator, 48*(4), 407-425.

Maximea, H. (2002a). Exhibition Galleries. In B. Lord & G. D. Lord (Eds.), *The Manual of Museum Exhibitions* (pp. 197-207). Walnut Creek, CA: AltaMira Press.

Maximea, H. (2002b). Planning & Designing Exhibition Facilities. In B. Lord & G. D. Lord (Eds.), *The Manual of Museum Exhibitions* (pp. 69-110). Walnut Creek, CA: AltaMira Press.

Mayer, R. E. (2004). Should there be a three-strike rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*(1), 14-10.

McLean, K. (1999). Museum Exhibitions and the Dynamics of Dialogue. *Daedalus, 128*(3), 83-107.

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

Meisner, R., Craubner, A., Dech, S., & Eales, P. (1999). Art meets science-computer animation of remote sensing data for communicating information on the environmental complexity of the polar regions. In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS '99)* (pp. 2730-2732).

Melton, A. W. (1935). *Problems of Installation in Museums of Art*. Washington, D.C: American Association of Museums.

Mevarech, Z. R. (1994). The effectiveness of individualized versus cooperative computer-based integrated learning systems. *International Journal of Educational Research, 21*(1), 39-52.

Miles, R. (2007). A Natural History Museum in Transition: Reflections on Visitor Studies in Practice. *Visitor Studies, 10*(2), 129 - 135.

Milrad, M., Perez, J., & Hoppe, U. (2002). C-Notes: Designing a Mobile and Wireless Application to Support Collaborative Knowledge Building. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02)*.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). Visuospatial Working Memory, Central Executive Functioning, and Psychometric Visuospatial Abilities: How Are They Related? *Journal of Experimental Psychology: General, 130*, 621-640.

Miyoshi, T., & Murata, A. (2001). A method to evaluate properness of GUI design based on complexity indexes of size, local density, aliment, and grouping. In *IEEE International Conference on Systems, Man, and Cybernetics* (pp. 221-226), Tucson, AZ, USA.

Moher, T. (2006). Embedded phenomena: supporting science learning with classroom-sized distributed simulations. In *Proceedings of CHI 2006* (pp. 691-700).

Moher, T., Ding, X., Wiley, J., Hussain, S., Singh, P., Srinivasan, V., et al. (2003). Combining handhelds with a whole-class display to support the learning of scientific control. In *CHI '03 extended abstracts on Human factors in computing systems*, Ft. Lauderdale, Florida, USA.

Moritsch, O., Hornecker, E., & Stifter, M. (2004). Assessing visitor behavior and attitudes in the medien.welten exhibition. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems* (pp. 1536), Vienna, Austria.

Morris, M. R., Huang, A., Paepcke, A., & Winograd, T. (2006). Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proceedings of the 2006 SIGCHI conference on Human Factors in computing systems (CHI)* (pp. 1201 – 1210), Montréal, Québec, Canada.

Morris, M. R., Morris, D., & Winograd, T. (2004). Individual audio channels with single display groupware: effects on communication and task strategy. In *Proceedings of the 2004 ACM conference on Computer Supported Cooperative Work* (pp. 242 - 251), Chicago, Illinois, USA.

Motorola. (2006). Motorola Provides Unique Look at the Past with Technology of the Future Retrieved 4/27/2008, from http://www.motorola.com/mediacenter/news/detail.jsp?globalObjectId=6675_6633_23&page=archive

Muller, K. (2002). Museums and Virtuality. *Curator, 45*(1), 21-33.

Myers, B. A. (2000). The pebbles project: using PCs and hand-held computers together. In *CHI '00 extended abstracts on Human factors in computing systems* (pp. 14-15), The Hague, The Netherlands.

Myers, B. A. (2001). Using handhelds and PCs together. *Communications of the ACM, 44*(11), 34-41.

Myers, B. A., Stiel, H., & Gargiulo, R. (1998). Collaboration using multiple PDAs connected to a PC. In *Proceedings of the ACM conference on Computer Supported Cooperative Work (CSCW)* (pp. 285-294), Seattle, Washington, United States.

Najjar, L. J. (1998). Principles of Educational multimedia user interface design. *Human Factors, 40*, 311-323.

Nearlife. Nearlife Recent Projects Information. Retrieved 5/19/2008, from http://www.nearlife.com/project/recent.html
Newhouse, V. (1999, February 7, 1999). In Paris, a Pair of Visions for the Future of Museums. *New York Times*,
Nicks, J. (2002). Curatorship in the Exhibition Planning Process. In B. Lord & G. D. Lord (Eds.), *The Manual of Museum Exhibitions* (pp. 345-347, 353-372). Walnut Creek, CA: AltaMira Press.
Norman, D. A., & Draper, S. W. (1986). *User Centered System Design: New Perspectives on Human-Computer Interaction*: Lawrence Erlbaum Associates.
Nova, N., Wehrle, T., Goslin, J., Bourquin, Y., & Dillenbourg, P. (2007). Collaboration in a multi-user game: impacts of an awareness tool on mutual modeling. *Multimedia Tools and Applications, 32*(2), 161--183.
O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups. The theoretical anatomy of group learning* (pp. 120-141). New York, NY: Cambridge University Press.
O'Hara, K., Kindberg, T., Glancy, M., Baptista, L., Sukumaran, B., Kahana, G., et al. (2007). Collecting and Sharing Location-based Content on Mobile Phones in a Zoo Visitor Experience. *Computer Supported Cooperative Work, 6*(1-2), 11-44.
O'Hara, K., Perry, M., Churchill, E., & Russell, D. (Eds.). (2003). *Public and Situated Displays: Social and Interactional Aspects of Shared Display Technologies (Computer Supported Cooperative Work)* Dortrecht, The Netherlands.: Kluwer.
Okuma, T., Kourogi, M., Sakata, N., & Kurata, T. (2007). A pilot user study on 3-D museum guide with route recommendation using a sustainable positioning system. In *Proceedings of the International Conference on Control, Automation and Systems (ICCAS '07)* (pp. 749-753).
Olson, J. S., Olson, G. M., Mack, L. A., & Wellner, P. (1990). Concurrent editing: the group's interface. In *Proceedings of Interact '90* (pp. 835-840), Cambridge, UK.
Onda, A., Oku, T., Chiu, P.-Y., Yu, E., Yokoi, M., Choh, I., et al. (2004). Visual-dimension interact system (VIS): exhibiting creative process for museum visitor experience. In *Proceedings of the International Conference on Cyberworlds* (pp. 278-285).
Oppenheimer, F. (1968). Rationale For A Science Museum. *Curator, 11*(3).
Ou, J., Oh, L. M., Yang, J., & Fussell, S. R. (2005). Effects of Task Properties, Partner Actions, and Message Content on Eye Gaze Patterns in a Collaborative Task. In *Proceedings of CHI 2005* (pp. 231-240), Portland, Oregon, USA.
Paek, T., Agrawala, M., Basu, S., Drucker, S., Kristjansson, T., Logan, R., et al. (2004). Toward universal mobile interaction for shared displays. In *Proceedings of the ACM conference on Computer Supported Cooperative Work (CSCW)* (pp. 266-269), Chicago, Illinois, USA.
Palincsar, A. S., & Herrenkohl, L. R. (1999). Designing Collaborative Contexts: Lessons from Three Research Programs. In A. M. O'Donnell & A. King (Eds.), *Cognitive Perspectives on Peer Learning* (pp. 151-177). Mahwah, NJ: Lawrence Erlbaum.
Palincsar, A. S., & Ladewski, B. G. (2006). Literacy and the Learning Sciences. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. New York: Cambridge University Press.
Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.
Paris, S. G. (2002). *Perspectives on object-centered learning in museums*. Mahwah, N.J. :: L. Erlbaum Associates.
Paris, S. G., & Hapgood, S. E. (2002). Children Learning with Objects in Informal Learning Environments. In S. G. Paris (Ed.), *Perspectives on object-centered learning in museums* (pp. 37-54). Mahwah, N.J. :: L. Erlbaum Associates.
Parush, A., Nadir, R., & Shtub, A. (1998). Evaluating the Layout of Graphical User Interface Screens: Validation of a Numerical Computerized Model. *International Journal of Human-Computer Interaction, 10*(4), 343-360.
Patterson, J. F. (1991). Comparing the programming demands of single-user and multi-user applications. In *UIST '91: Proceedings of the 4th annual ACM symposium on User interface software and technology* (pp. 87--94), Hilton Head, South Carolina, United States.
Pawar, U. S., Pal, J., Gupta, R., & Toyama, K. (2007). Multiple mice for retention tasks in disadvantaged schools. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 1581--1590), San Jose, California, USA.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47-87). Cambridge, England: Cambridge University Press.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences, 13*(3), 423–451.

Peart, B. (1984). Impact of Exhibit Type on Knowledge Gain, Attitudes and Behavior. *Curator, 27*(3), 220-236.

Penner, D. E. (2000). Explaining systems: Investigating middle school students' understanding of emergent phenomena. *Journal of Research in Science Teaching, 37*(8), 784-806.

Piaget, J. (1970). *Science of Education and Psychology of the Child*. New York: Oxford University Press.

Piaget, J. (1976). *The child's conception of the world.* Totowa, NJ: Littlefield, Adams, & Co.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., et al. (2004). A Scaffolding Design Framework for Software to Support Science Inquiry. *Journal of the Learning Sciences, 13*(3), 337–386.

Quintana, C., Shin, N., Norris, C., & Soloway, E. (2006). Learner-Centered Design: Reflections on the past and directions for the future. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. New York: Cambridge University Press.

Ramsey, G. F. (1938). *Educational Work in Museums in the United States: Development, Methods, and Trends*. New York: H. W. Wilson Company.

Raptis, D., Tselios, N., Tselios, N., & Avouris, N. (2005). Context-based design of mobile applications for museums: a survey of existing practices. In *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services* (pp. 153--160), Salzburg, Austria.

Rasmusen, E. (2006). *Games and Information: An Introduction to Game Theory* (4 ed.). Oxford: Wiley-Blackwell.

Reeves, S., Benford, S., O'Malley, C., & Fraser, M. (2005). Designing the spectator experience. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 741-750), Portland, Oregon, USA.

Rekimoto, J. (1998). A multiple device approach for supporting whiteboard-based interactions. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 344-351), Los Angeles, California, United States.

Repenning, A., & Ioannidou, A. (2004). Agent-Based End-User Development. *Communications of the ACM, 47*(9), 43-46.

Repenning, A., & Ioannidou, A. (2005). Mr. Vetro: A Collective Simulation Framework. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 1612-1619), Chesapeake, VA, USA.

Resnick, M., & Wilensky, U. (1998). Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. *Journal of the Learning Sciences, 7*, 153 - 172.

Roberts, L. C. (1997). *From Knowledge to Narrative: Educators and the Changing Museum*. Washington and London: Smithsonian Institution Press.

Robinson, E. (1928). *The Behavior of the Museum Visitor*. Washington D.C.: American Association of Museums.

Robinson, E. (1930). Psychological Problems of the Science Museum. *Museum News, 8*(5).

Rocchi, C., Stock, O., Zancanaro, M., Kruppa, M., & Krüger, A. (2004). The museum visit: generating seamless personalized presentations on multiple devices. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces*, Funchal, Madeira, Portugal.

Rodden, T., Rogers, Y., Halloran, J., & Taylor, I. (2003). Designing novel interactional workspaces to support face to face consultations. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 57--64), Ft. Lauderdale, Florida, USA.

Rogoff, B., & Lave, J. (Eds.). (1984). *Everyday Cognition: Its Development in Social Context*. Cambridge, MA: Harvard University Press.

Roschelle, J. (1992). Learning by Collaborating: Convergent Conceptual Change. *The Journal of the Learning Sciences, 2*(3), 235-276.

Roschelle, J., & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology, 1*(1), 145-168.

Roschelle, J., Tatar, D., Chaudbury, S. R., Dimitriadis, Y., Patton, C., & DiGiano, C. (2007). Ink, Improvisation, and Interactive Engagement: Learning with Tablets. *Computer, 40*(9), 42-48.

Rounds, J. J. (2006). Doing Identity Work in Museums. *Curator, 49*(2), 133-150.

Rummel, N., & Spada, H. (2005). Learning to Collaborate: An Instructional Approach to Promoting Collaborative Problem Solving in Computer-Mediated Settings. *The Journal of the Learning Sciences, 14*(2), 201-241.

Sandifer, C. (2003). Technological novelty and open-endedness: Two characteristics of interactive exhibits that contribute to the holding of visitor attention in a science museum. *Journal of Research in Science Teaching, 40*(2), 121-137.

Sawyer, R. K. (2006). Analyzing Collaborative Discourse. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. New York, NY: Cambridge University Press.

Schauble, L., Gleason, M., Lehrer, R., Bartlett, K., Petrosino, A., Allen, A., et al. (2002). Supporting Science Learning in Museums. In G. Leinhardt, K. Crowley & K. Knutson (Eds.), *Learning Conversations in Museums* (pp. 333-356). Mahwah, New Jersey: Lawrence Erlbaum.

Schellens, T., Keer, H. V., Valcke, M., & Wever, B. D. (2005). The impact of role assignment as scripting tool on knowledge construction in asynchronous discussion groups. In *CSCL '05: Proceedings of the 2005 conference on Computer support for collaborative learning*, Taipei, Taiwan.

Schwartz, D. L. (1999). The productive agency that drives collaborative learning. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 197–218). New York: Elsevier Science.

Schwieren, J., & Vossen, G. (2007). Implementing Physical Hyperlinks for Mobile Applications Using RFID Tags. In *Proceedings of the 11th International Database Engineering and Applications Symposium (IDEAS '07)* (pp. 154-162).

Scott, S. D., Grant, K. D., & Mandryk, R. L. (2003). System guidelines for co-located, collaborative work on a tabletop display. In *ECSCW'03: Proceedings of the eighth conference on European Conference on Computer Supported Cooperative Work* (pp. 159--178), Helsinki, Finland.

Scott, S. D., Mandryk, R. L., & Inkpen, K. M. (2002). Understanding Children's Interactions in Synchronous Shared Environments. In *Proceedings of CSCL 2002*, Boulder, CO.

Scott, S. D., Mandryk, R. L., & Inkpen, K. M. (2003). Understanding children's collaborative interactions in shared environments. *Journal of Computer Assisted Learning, 19*, 220-228.

Serrell, B. (1997). Paying Attention: The Duration and Allocation of Visitors' Time in Museum Exhibitions. *Curator, 40*(2), 108-124.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*(1), 4-27.

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2008). *Research Methods in Psychology* (8 ed.). New York, NY: McGraw-Hill.

Shoemaker, G. B. D., & Inkpen, K. M. (2001). Single display privacyware: augmenting public displays with private information. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 522-529), Seattle, Washington, United States.

Silverman, L. H. (1995). Visitor meaning-making in museums for a new age. *Curator: The Museum Journal, 38*, 161-170.

Slavin, R. E. (1992). When and Why Does Cooperative Learning Increase Achievement? Theoretical and Empirical Perspectives. . In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups. The theoretical anatomy of group learning* (pp. 174–199). New York: Cambridge University Press.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*, 5-42.

Smith, H., Luckin, R., Fitzpatrick, G., Avramides, K., & Underwood, J. (2005). Technology at work to mediate collaborative scientific enquiry in the field. Paper presented at the *AIED 2005, Frontiers in Artificial Intelligence and Applications*.

Smith, R. B., Hixon, R., & Horan, B. (1998). Supporting flexible roles in a shared space. In *CSCW '98: Proceedings of the 1998 ACM conference on Computer supported cooperative work* (pp. 197--206), Seattle, Washington, United States.

Spencer, H. A. D. (2002). Exhibition Text. In B. Lord & G. D. Lord (Eds.), *The Manual of Museum Exhibitions* (pp. 393-402). Walnut Creek, CA: AltaMira Press.

Squire, K., & Jan, M. (2007). Mad City Mystery: Developing Scientific Argumentation Skills with a Place-based Augmented Reality Game on Handheld Computers. *Journal of Science Education and Technology, 16*(1), 5-29.

Squire, K., & Klopfer, E. (2007). Augmented Reality Simulations on Handheld Computers. *Journal of the Learning Sciences, 16*(3), 371-413.

Stahl, G. (2006a). Computer Supported Collaborative Learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 409-426). New York: Cambridge University Press.

Stahl, G. (2006b). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA: MIT Press.

Stanton, D., Neale, H., & Bayon, V. (2002). Interfaces to Support Children's Co-present Collaboration: Multiple Mice and Tangible Technologies. In *Proceedings of CSCL 2002*, Boulder, CO.

Stefik, M., Bobrow, D. G., Foster, G., Lanning, S., & Tatar, D. (1987). WYSIWIS revised: early experiences with multiuser interfaces. *ACM Transactions on Information Systems (TOIS), 5*(2), 147--167.

Stefik, M., Foster, G., Bobrow, D. G., Kahn, K., Lanning, S., & Suchman, L. (1987). Beyond the Chalkboard: Computer Support for Collaboration and problem Solving in Meeting. *Communications of the ACM, 30*(1), 32-47.

Steiner, K. E., & Moher, T. (2002). Encouraging task-related dialog in 2D and 3D shared narrative workspaces. In *CVE '02: Proceedings of the 4th international conference on Collaborative virtual environments* (pp. 39--46), Bonn, Germany.

Stewart, J., Bederson, B. B., & Druin, A. (1999). Single display groupware: a model for co-present collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit* (pp. 286-293), Pittsburgh, Pennsylvania, United States.

Stewart, J., Raybourn, E. M., Bederson, B., & Druin, A. (1998). When two hands are better than one: enhancing collaboration using single display groupware. In *Proceedings of the CHI 98 conference summary on Human factors in computing systems* (pp. 287-288), Los Angeles, California, United States.

Stock, O., & Zancanaro, M. (2007). *PEACH: Intelligent Interfaces for Museum Visits*. Berlin: Springer-Verlag.

Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Kruger, A., Kruppa, M., et al. (2007). Adaptive, intelligent presentation of information for the museum visitor in PEACH. *User Modeling and User-Adapted Interaction, 17*(3), 257 - 304.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, D.C., USA: Brookings Institution Press.

Streitz, N., Prante, T., Müller-Tomfelde, C., Tandler, P., & Magerkurth, C. (2002). Roomware®: the second generation. In *CHI '02: CHI '02 extended abstracts on Human factors in computing systems* (pp. 506--507), Minneapolis, Minnesota, USA.

Streitz, N. A., Geißler, J., Holmer, T., Konomi, S. i., Müller-Tomfelde, C., Reischl, W., et al. (1999). i-LAND: an interactive landscape for creativity and innovation. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 120--127), Pittsburgh, Pennsylvania, United States.

Suchman, L. (1987). *Plans and situated actions : the problem of human-machine communication*. New York: Cambridge University Press.

Sugimoto, M., Hosoi, K., & Hashizume, H. (2004). Caretta: a system for supporting face-to-face collaboration by integrating personal and shared spaces. In *Proceedings of the 2004 conference on Human factors in computing systems (CHI '04)* (pp. 41-48), Vienna, Austria.

Sweeney, T. (2001). Wired Worlds: Exploring the Digital Frontier. A pioneering gallery of digital media discovery. *Computer Graphics and Applications, 21*(1), 72-77.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General, 119*, 176-192.

Taborsky, E. (1990). The Discursive Object. In S. M. Pearce (Ed.), *Objects of Knowledge* (pp. 50-77). London: The Athlone Press.

Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science, 312*(5777), 1143 – 1144.

Tani, M., Horita, M., Yamaashi, K., Tanikoshi, K., & Futakawa, M. (1994). Courtyard: integrating shared overview on a large screen and per-user detail on individual screens. In *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence* (pp. 44-50), Boston, Massachusetts, United States.

Tanikawa, T., Ando, M., Wang, Y., Yoshida, K., Yamashita, J., Kuzuoka, H., et al. (2004). A case study of museum exhibition - historical learning in Copan ruins of Mayan civilization. In *Proceedings of Virtual Reality* (pp. 257-258), Chicago, Illinois, USA.

Taylor, B. (2003). Enhancing the Value of Museum Websites: Lessons from the Practical Engagement Front. *Journal of Academic Librarianship, 39*(2/3), 107-122.

Taylor, S. M. (1991). *Try It! Improving Exhibits through Formative Evaluation*. Washington, D.C.: Association of Science-Technology Centers.

The British Museum. (n.d.). The Eastern Zoological Gallery in the East Wing of The British Museum. Retrieved 5/16/2008, from http://www.britishmuseum.org/explore/highlights/highlight_image.aspx?image=ps329281.jpg&retpage=16290

The Walt Disney Company. (n.d.). Canada Pavilion. Retrieved 5/16/2008, from http://disneyworld.disney.go.com/wdw/parks/attractionDetail?id=CanadaPavilionAttractionPage

Thom-Santelli, J., Boehner, K., Gay, G., & Hembrooke, H. (2006). Beyond just the facts: transforming the museum learning experience. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems* (pp. 1433--1438), Montreal, Quebec, Canada.

Thompson, M., & Moher, T. (2006). HelioRoom: Problem-solving in a whole class visual simulation. In *Proceedings International Conference of the Learning Sciences (ICLS 2006)* Bloomington, IN, USA.

Tomlinson, B. (2005). A Heterogeneous Animated Platform for Educational Participatory Simulations. In *10th Computer Supported Collaborative Learning Conference (CSCL '05)* (pp. 677-681), Taipei, Taiwan.

Traub, J. (2004, Oct 3, 2004). The Stuff Of City Life. *New York Times Magazine,* 23-27.

Tse, E., Shen, C., Greenberg, S., & Forlines, C. (2006). Enabling interaction with single user applications through speech and gestures on a multi-user tabletop. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces* (pp. 336--343), Venezia, Italy.

Turkle, S. (1984). *The second self : computers and the human spirit*. New York: Simon and Schuster.

Vahey, P., Tatar, D., & Roschelle, J. (2004). Leveraging Handhelds to Increase Student Learning: Engaging middle school students with the mathematics of change. In *Proceedings of ICLS 2004* (pp. 553-560).

Vahey, P., Tatar, D., & Roschelle, J. (2007). Using Handheld Technology to Move Between the Private and Public in the Classroom. In M. A. van 't Hooft & K. Swan (Eds.), *Ubiquitous computing: Invisible technology, visible impact*. Mahwah, NJ: Erlbaum.

van Merriënboer, J. J. G., & Ayres, P. (2005). Research on Cognitive Load Theory and Its Design Implications for E-Learning. *Educational Technology Research and Development, 53*(3), 5-13.

Vath, R., Lyons, L., Lee, J., Kawamura, M., Quintana, C., & Soloway, E. (2005). Addressing Assessment Challenges for a Multi-User Simulation with Handheld Integration (MUSHI). Poster presented at *The 4th International Conference for Interaction Design and Children (IDC 2005)*, Boulder, CO, USA.

Viney, L. (2005). Fuelling the future. *BP Magazine*(1), 56-60.

vom Lehn, D., Heath, C., & Hindmarsh, J. (2001). Conduct and Collaboration in Museums and Galleries. *Symbolic Interaction, 24*(2), 189-216.

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

Wagner, D., Schmalstieg, D., & Billinghurst, M. (2006). Handheld AR for Collaborative Edutainment. In *International Conference on Artificial Reality and Telexistence (ICAT '06)* (pp. 85-96).

Walter, T. (1996). From museum to morgue? Electronic guides in Roman Bath. *Tourism Management, 17*(4), 241-245.

Wang, F., & Hannafin, M. J. (2005). Design-Based Research and Technology-Enhanced Learning Environments. *Educational Technology Research and Development, 53*(4), 5-23.
Wang, Y., Yang, C., Liu, S., Wang, R., & Meng, X. (2007). A RFID & Handheld Device-Based Museum Guide System. In *Proceedings of the 2nd International Conference on Pervasive Computing and Applications (ICPCA '07)* (pp. 308-313).
Webb, N. M. (1984). Stability of small group interaction and achievement over time. *Journal of Educational Psychology, 76*(2), 211-224.
Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13*, 21-39.
Weiss, R. S., & Boutourline, S. J. (1963). The Communication Value of Exhibits. *Museum News, 42*(3), 23-27.
Wessel, D., & Mayr, E. (2007). Potentials and Challenges of Mobile Media in Museums. *International Journal of Interactive Mobile Technologies, 1*(1), 1-8.
White, T. (2006). Code talk: Student discourse and participation with networked handhelds. *International Journal of Computer-Supported Collaborative Learning, 1*(3), 359-382.
Wilcove, D. S., & Eisner, T. (2000). The impending extinction of natural history. *The Chronicle for Higher Education, B24*.
Wilensky, U., & Stroup, W. M. (1999). Learning through Participatory Simulations: Network-based Design for Systems Learning in Classrooms. In *Proceedings of CSCL '99, Computer-Supported Collaborative Learning*, Palo Alto, CA, USA.
Wilensky, U., & Stroup, W. M. (2000). Networked Gridlock: Students Enacting Complex Dynamic Phenomena with the HubNet Architecture. In *Fourth International Conference of the Learning Sciences* (pp. 282-289), Ann Arbor, MI, USA.
Wilkinson, L. C., & Marrett, C. B. E. (1985). *Gender Influences in Classroom Interaction*. Orlando, FL: Academic Press.
Witcomb, A. (2006). Interactivity: Thinking Beyond. In S. Macdonald (Ed.), *A Companion to Museum Studies* (pp. 353-361). United Kingdom: Blackwell.
Wolf, M., Lee, E., & Borchers, J. (2007). Education, entertainment and authenticity: lessons learned from designing an interactive exhibit about medieval music. In *CHI '07: CHI '07 extended abstracts on Human factors in computing systems* (pp. 1887--1892), San Jose, CA, USA.
Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17*, 89–100.
Woodruff, A., Szymanski, M. H., Aoki, P. M., & Hurst, A. (2001). The Conversational Role of Electronic Guidebooks. In *Proceedings of the 3rd international conference on Ubiquitous Computing* (pp. 187 - 208), Atlanta, Georgia, USA.
Yamada, S., Hong, J.-K., & Sugita, S. (1995). Development and evaluation of hypermedia for museum education: validation of metrics. *ACM Transactions on Computer-Human Interaction (TOCHI), 2*(4), 284 - 307.
Yamaguchi, T., Kaji, H., & Kusunoki, F. (2005). Pit a pat: design for support learning exhibitions. In *Proceedings of the 2005 International Conference on Active Media Technology (AMT '05)* (pp. 276), Takamatsu, Kagawa, Japan.
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 IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04)* (pp. 109-113), Taoyuan, Taiwan.
Zucker, B. F. (1987). *Children's Museums, Zoos, and Discovery Rooms*. Westport, CT: Greenwood Press.
Zurita, G., & Nussbaum, M. (2004). Computer supported collaborative learning using wirelessly interconnected handheld computers. *Computers & Education, 42*(3), 289-314.