Designs for Learning: Studying Science Museum Exhibits That Do More Than Entertain

SUE ALLEN
Exploratorium, San Francisco, CA 94123, USA

Received 2 October 2003; revised 6 April 2004; accepted 7 April 2004

DOI 10.1002/sce.20016
Published online in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: Science museum staff face a constructivist dilemma as they design their public spaces: the exhibits should facilitate science learning, yet they also need to support a diverse visiting public in making their own personal choices about where to attend, what to do, and how to interpret their interactions. To be effective as teaching tools, exhibits need to be highly intrinsically motivating at every step of an interaction in order to sustain involvement by an audience who views their visit primarily as a leisure activity. Given these challenges, it is vital to support the design process with a strong program of research and evaluation. I give a personal perspective on one institution's research and evaluation work over the last decade, focusing on four areas: immediate apprehendability, physical interactivity, conceptual coherence, and diversity of learners. © 2004 Wiley Periodicals, Inc. *Sci Ed* **88**(Suppl. 1)**:**S17–S33, 2004

INTRODUCTION: A PERSONAL PERSPECTIVE ON A LEARNING DILEMMA

In this article I present a personal perspective on a subset of museum learning environments, specifically the exhibition space of hands-on science museums. This choice is based on my own experience of studying and working in such environments over the last decade at the Exploratorium, a well-known museum of "science, art, and human perception" in San Francisco. Further, I focus my comments on the use of exhibits[1] in these environments by visitors in "self-guided" mode (i.e., without mediation and guidance by staff members), because this is the most common type of experience that casual visitors have at our museum.

At first glance, the exhibit space of a science museum seems an appealing educational alternative to a school science classroom: hands-on exhibits are novel, stimulating, evidence-rich, multisensory, and fun. The environment provides myriad personal choices, without any teachers forcing learners to do something unappealing, without curricular constraints,

Correspondence to: Sue Allen; e-mail: suea@exploratorium.edu

[1] In this article I use "exhibit" to mean an individual element (typically engaging visitors for about a minute at our museum), and "exhibition" or "exhibit collection" to refer to larger, themed groupings.

without testing or accountability. “No one ever flunked a museum,” as was often pointed out by Frank Oppenheimer, the Exploratorium’s founder (Semper, 1990). But science museums are actually very difficult environments to engineer for learning, precisely because of these same attributes. On the exhibit floor there is no accountability, no curriculum, no teachers to enforce concentration, no experienced guide to interpret and give significance to the vast amounts of stimulus and information presented. Without restrictions, visitors have complete freedom to follow their interests and impulses as they move through a public space packed with exhibits all vying for attention. This quality of totally unrestricted choice in what to attend to has huge implications for learning in the museum setting. As a newcomer to the field, I did not appreciate the magnitude of this effect, but after a decade of research and evaluation I consider it the single greatest constraint underlying exhibit design.

In a school setting, a teacher can use a variety of strategies to regulate her students’ progress, ensuring that they all arrive at the rewarding or significant climax of a lesson. By contrast, if an exhibit has a boring or effortful or confusing component, visitors have no way of knowing whether the reward for persisting will be worth the effort; and in an environment full of interesting alternatives, they are very likely to simply leave the exhibit and move on. Because of this, it is not enough that an exhibit has a culminating point or experience that is rewarding to visitors; *every intermediate step* in the visitors’ experience must be sufficiently motivating that they make the choice of continuing to invest time and attention there. This dependence on personal motivation at every step of an exhibit interaction makes it incredibly difficult to get visitors to follow a narrowly constrained learning agenda, particularly one involving sequential steps and moderate or high levels of intellectual effort.

Such challenges form part of the constructivist dilemma as it applies to museums: We expect these institutions to provide a hugely diverse visiting public with entertainment, the freedom to choose their own path, follow their personal interests, do their own inquiry, and create their own meanings. Yet at the same time, we want our museums to be respected educational institutions where people can spend an hour and come away having learned some canonical science. This dilemma plays out at every grain-size, from the largest organizational tensions between market and mission to the smallest design challenges of a single exhibit element. Over the last decade I have come to believe that it is indeed possible to create exhibit environments where visitors are simultaneously in a constant state of free choice and in the process of learning some form of science. But it is difficult, and calls for a program of research that focuses on the detailed features of the physical environment in which such learning is deeply situated. In the following sections I highlight some of the Exploratorium’s attempts to meet this challenge, and I summarize several key considerations for designing such environments effectively.

EARLY SUCCESS: THE ORIGINAL EXPLORATORIUM MODEL

The Exploratorium was started in 1969, brainchild of physicist Frank Oppenheimer. It was one of the first institutions to design and build hands-on exhibits based on what has since become the “standard model” of learning in science museums. The emphasis was on giving visitors direct experience with natural, physical, and technological phenomena, on the assumption that this would allow them to build the confidence and skills to understand the world around them. Although Oppenheimer’s educational model was complex and nonprescriptive, the basis of his approach to exhibits was compatible with that of Dewey, giving central importance to the role of direct experience of phenomena, and trying to present the learner with a problematic experience from which he/she could conduct genuine inquiry. It is also compatible with the Piagetian notion of disequilibration as a driver for learning through change of existing knowledge schemas.

A typical exhibit of this type, *Water Standing on Air,* is shown in Figure 1. This exhibit leads visitors through the following inquiry cycle. (1) Surprising phenomenon: visitors are invited to turn the large plexiglass cylinder over so that the water rushes down through a metal mesh barrier across the middle. Unexpectedly, the water does not all go through the holes in the mesh, but several inches of it come to rest above the mesh. (2) Exploration: visitors try rotating the cylinder repeatedly and in different ways to see what makes the effect work or fail. This leads to discovery of such things as speed of rotation being helpful, and the fact that the suspended water is unstable above the mesh. (3) Explanation: The label explains the cause of the phenomenon in terms of physics principles, in this case a combination of surface tension at the holes and the slight pressure differential between the air above and below the mesh. (4) Relevance: The label makes a connection to the everyday experience of capturing liquid in a straw by dipping it into the liquid and capping it with one's thumb.

Returning to the constructivist dilemma for museums, how did this model help people to learn science while in a state of free choice? I believe it used a combination of strategies:

a. It emphasized the aspects of science most easily and pleasurably learned in a physically complex and chaotic environment, namely those involving exploration, physical manipulation, and experimentation.
b. It de-emphasized many other aspects of science, including anything requiring memorization (e.g., detailed vocabulary, quantitative relationships), or anything requiring long chains of inference or effortful thinking (e.g., designing experiments to discriminate among competing models, arguing the relative merits of two explanations).

Figure 1. A typical exhibit: *Water Standing on Air*.

c. It articulated (through its label) and supported (through its physical design) a simple cycle of inquiry. Curiosity was used as the driving force that nudged visitors throughout this cycle a step at a time, from novel phenomenon to exploration to explanation and finally to relevance. The questions "what's going on?" and "so what?" were used as label headings to help raise visitors' curiosity at these intermediate steps, as well as to scaffold them through the cycle. The fundamental role played by curiosity gave this inquiry cycle a natural and effortless quality, particularly for visitors with little or no background in science.

From a constructivist perspective, this model was ahead of its time. While many museums were still dominated by transmission-based theories of learning, science (and children's) museums embraced this model because it put the visitor in a very active role as learner: experimenting, hypothesizing, interpreting, and drawing conclusions. At the same time, these efforts were not coupled with a program of research on learning, and some of its limitations have become apparent over the last decade, as I will describe shortly.

AN INSITUTIONAL RESEARCH AGENDA

Since the arrival of Kathleen McLean a decade ago, the Exploratorium's exhibit staff has been creating experimental exhibits that embody a broader range of learning models. Integral to this effort, we have established a department of visitor research and evaluation within the museum, with a research agenda at the intersection of the academic and practitioner communities. For us, research and practice are deeply linked. I wish to highlight this integrated approach by discussing four aspects of the learning environment that have emerged as important in our research and that of other institutions: immediate apprehendability, physical interactivity, conceptual coherence, and diversity of learning modes.

Immediate Apprehendability: Creating Effortless Backdrops for Challenge

Cognitive overload is a huge problem in museums of all kinds (e.g., Evans, 1995; Hedge, 1995), but perhaps especially in hands-on science museums. Consider the challenge: visitors are faced with a gyrating landscape of hundreds of exhibits, none of which they have probably seen before, and none of which has standardized controls, mechanisms, or explanations. Adults wanting to support their children must make sense of each novel device, decipher the instructions, guide their children toward the key experience, interpret this experience for themselves, translate the significance of it for their children, assess the result, and make on-the-fly adjustments as needed to optimize their children's learning. Over and over, every few minutes, adults coach their children in technical and cognitive skill-building without previous training. The effort it takes to negotiate a museum is apparent through the common phenomenon of "museum fatigue," in which visitors can only engage deeply with exhibits for a limited period (typically about 30 min) before they lose their focused attention and begin to "cruise," looking for anything particularly compelling before moving on (Falk et al., 1985). Museum fatigue is an important factor that limits the degree to which visitors can effectively learn any form of science.

In considering properties of museum environments that promote both intrinsic motivation and science learning, the prevalence of cognitive overload points to the central importance of something I call "immediate apprehendability." By this, I mean the quality of a stimulus or larger environment such that people introduced to it for the first time will understand its purpose, scope, and properties almost immediately and without conscious effort. Immediate apprehendability is similar to the idea of affordance (Gibson, 1977; Norman, 1988) but generalized to include things with no direct physical use, such as labels, as well as complex possibilities for action such as the playing of a specific game. Clearly apprehendability depends on the prior knowledge of the person introduced to the environment (the museum visitor), but it is possible to consider it as a property of the environment to the extent that the visitors share perceptual and conceptual schemata.

User-Centered Design. One way to achieve immediate apprehendability and reduce cognitive overload in exhibits is through "user-centered design" (also called "end-user" or "natural" design). This approach promotes the creation of objects that, by virtue of their physical forms and locations, invite certain kinds of use and not others. Such design often goes unseen and unappreciated because, ironically, masterful design results in objects that seem obvious and simple to use.

Started by industrial and human factors designers (e.g., Dreyfus, 1967; Loewy, 1951; Papanek, 1971), user-centered design came into mainstream psychology through the work of cognitive scientist Donald Norman. He used the term "affordances" to refer to the directly perceivable properties of objects that determine how they could possibly be used. "Plates are for pushing. Knobs are for turning. Slots are for inserting things into. . . . When affordances are taken advantage of, the user knows what to do just by looking: no picture, label, or instruction is required. Complex things may require explanation, but simple things should not" (Norman, 1988, p. 9). The principles of user-centered design were particularly compatible with the exhibit development philosophy at the Exploratorium, and after many iterations of design, research, and reflection, the idea of affordances now permeates the institution.

Beyond the concept of affordances, other relevant principles from this field that we have tested and found to be robust include:

- Using natural mappings to take advantage of physical analogies and cultural standards. For example, turning a knob clockwise increases the amount of whatever is being adjusted.
- Limiting the available controls to one small set at a time. While this can be done most easily with electronic interfaces, it is also possible with physical exhibits. One way is by making controls differentially salient. For example, Figure 2 shows *Touch the Spring*, an exhibit that displays an image of a spring created by a concave mirror. The primary surprising phenomenon of this exhibit is that the image looks like a real spring, but when visitors try to touch it, their hands pass straight through. Another way to interact with the exhibit is by pointing a flashlight at the image; surprisingly, it casts a shadow. But this interaction is a secondary one, and the flashlight is deliberately positioned to the side of the exhibit where it dangles inconspicuously, most often discovered after the primary visitor interaction. Another way to "chunk" the controls into manageable sets is to separate a complex exhibit into adjacent stations, each of which emphasizes a different aspect of the phenomenon to be explored (e.g., different electric circuit constructions to make a fan turn, light a bulb, turn a motor, etc.).
- Standardizing for consistency to reduce cognitive load for subsequent interactions. At the Boston Museum of Science, the Hearphones (small earpieces that provide audio versions of label text) are always positioned in the same place on the exhibit label so that blind or low-vision visitors don't have to search for them each time. At the Exploratorium, all of the magnets in the Electricity & Magnetism section are painted red and white to aid in conceptual clarity.

Figure 2. *Touch the Spring*, an exhibit that displays an image of a spring created by a concave mirror.

Familiar Activities as Schemas. Findings from our research and evaluation studies suggest that immediate apprehendability may also be increased through the use of familiar activities as over-arching schemas. Within a few seconds, visitors understand the broad nature of the activity, and can attend to interpretation of the specific rules and objects involved. Several common examples follow:

- Making a complex machine work. *Bike Cycle* is a robotic simulation of a bicycle rider, with hydraulic pistons in place of major leg muscles. Four buttons control the contraction of these pistons. To make the robot pedal requires pressing these buttons in the right sequence and with exact timing, a significant challenge usually requiring multiple attempts. Yet we found that 73% of visitors persisted until they had succeeded, and 70% said they thought the level of challenge was about right. This exhibit offered visitors immediate apprehendability at the large scale (the robot is immediately recognizable as a simulated person on a bike that is not moving), as well as the small scale (the only controls are four large buttons, centrally located, easy to push, and color-coded to match the relevant pistons).
- A competition, especially a race. *Downhill Race* is an exhibit in which visitors race wheels of different shape and weight down an inclined plane. The presence of two lanes side by side, the exhibit title, and the set of six disks with slightly varying features, all contribute to visitors' immediate apprehension that the exhibit involves racing the disks down the plane. Most people quickly know what to do. Within the race schema, visitors have time and attention to think about the characteristics of the disks that affect their speed, and several studies have shown that this exhibit provokes predictions, explanations, and new understandings that can be articulated (Gutwill, 2002; Perry & Tisdal, 2004). Rowe (2002) previously analyzed a similar

exhibit and noted the power of a competition or race as an idea that structured visitor activity.

- Watching and waiting. A successful exhibit schema need not involve energetic physical activity. As an example, we were surprised to discover the high levels of learning elicited by an "empty" tank of frogs in one of the Exploratorium's temporary exhibitions. The circular, galvanized iron tank contained a mini-ecosystem of pebbles, water, and plants, and was in fact full of frogs, but because they were nocturnal, most were hidden and sleeping during the museum's open hours. While the staff at first considered this a problem, a study of visitors' conversations revealed that this exhibit was the most likely of any in the exhibition to elicit complex inferences from visitors, such as theories about camouflage and froggy lifestyle (Allen, 2002). It was also the second most likely element to evoke stories and personal associations, and the second most likely to elicit declarations of previous knowledge about frogs. Apparently visitors recognized the exhibit as an example of something where one had to wait and watch in order to be rewarded by a frog sighting, and within this framework they accepted the challenge and used it as an opportunity to share stories and knowledge. None of the 30 pairs of visitors expressed frustration or annoyance at having to look for the frogs, even those who failed to find any.

Familiar activities can also be used on the scale of an entire exhibition. For example, the detective puzzle used to frame the *Whodunit* exhibition about forensic science allowed visitors to immediately apprehend the purpose and relationship of distinct elements such as the crime scene, DNA lab, and simulated autopsy laboratory (Walter, in press).

Immediate Apprehendability on All Scales. Csikszentmihalyi and Hermanson (1995) characterize ideal learning at exhibits as initially driven by curiosity and interest, and then sustained via a "flow" state, in which visitors become fully involved with mind and body in an intrinsically motivated activity. To create a flow state, activities generally need a level of challenge that closely matches the person's skills, as well as a clear set of goals and rules. However, I believe we need to optimize not just the level of challenge but also its timing and context: research has suggested that visitors will only engage in a challenge if they are comfortable and oriented (e.g., Hayward & Brydon-Miller, 1984). Thus immediate apprehendability may be a particularly important quality for the early or framing stages of an experience, because it lessens distracting stimuli and helps to put visitors within a comfortable framework from which to be curious.

In fact, one might view the whole museum visit as an alternating series of challenges and apprehensions on different scales: A family on holiday chooses to visit the Exploratorium because they are curious to see something novel and interesting. They drive to the museum, expecting its location to be immediately apprehendable, but due to the poor road signs they become lost and spend time and energy finding the building, and perhaps even finding the entrance. Once inside, the admissions desk, bathrooms, and coffee cart are immediately apprehendable, allowing them to move effortlessly into the main exhibition space. They become curious about the movement of a nearby exhibit, a large compound pendulum with arms that flail chaotically. Fortunately the control of the exhibit, a large brass knob, is immediately apprehendable and they move quickly into a phase of exploration, turning the knob to see what different kinds of behavior they can provoke from the exhibit. Now curious about the purpose of such a strange device, they scan the label and easily identify "What's going on" as the part they want to read. They invest time and effort in reading and making sense of this scientific explanation.

In general, neither staff nor visitors would tell this story by including the immediately apprehendable aspects of the environment. By definition, they are the things we don't notice. However, they are just as important from the design perspective, because they reduce the ever-present cognitive load on visitors, freeing them to focus on those aspects of the environment that are rewarding to them and worthy of their attention. As visitor researcher Serrell emphasizes, "people appreciate being given information that will help them make intelligent choices" (1996, p. 72).

At the Exploratorium, we have tried to employ such principles in the redesign of our public space, including more lighting and seating, some acoustic baffling to reduce ambient noise (Fry, 2002), and orientational devices such as maps, 3-dimensional models, and way-finding signs. The importance of such "comfort" factors on learning should never be underestimated: in one key study, Falk and Balling (1977) showed that children on school field trips showed higher levels of learning if they were given a pre-visit orientation based on the layout of the museum and the opportunities to eat and make purchases from its store, than if they were given an orientation based on relevant biological concepts. More recently, Maxwell and Evans (2002) link the physical environment to learning through psychological processes such as cognitive fatigue, distraction, motivation, and anxiety, and they provide evidence that learning is enhanced in quieter, smaller, better-differentiated spaces.

Immediate apprehendability, with its emphasis on comfortable and effortless understanding by users, might seem antithetical to a museum whose learning model rests on the motivating properties of surprise, cognitive dissonance, and curiosity. However, based on studies at our museum and many others, I believe that it forms the essential backdrop which makes such exhilarating foreground experiences possible.

Physical Interactivity: A More Critical Look

Physical interactivity, the ability of an exhibit to respond to visitor actions, is considered a cardinal feature of science (and children's) museums. From the constructivist perspective, this is the part of science learning that involves giving the learner access to the key phenomena of the natural world. Research on visitor learning in museums suggests that interactivity promotes engagement, understanding, and recall of exhibits (for a recent review, see Schneider & Cheslock, 2003). For example, Maxwell and Evans cite evidence that both children and adults recall actions they themselves perform better than those they observe. Richards and Menninger (2000) evaluated specially-designed interactive galleries at the J. Paul Getty Museum and found that holding time was greater in those galleries. Borun and Dritsas (1997) identified interactive design features such as multioutcome and multimodal as key ingredients in exhibits that will foster family learning at science centers.

At the Exploratorium we conducted a set of rigorous studies on the role of physical interactivity in exhibits displaying microscopic organisms on video screens (of which we have numerous examples). There was considerable interest in this topic among exhibit developers, because keeping animals alive is expensive and time-consuming, and several visiting scientists have suggested that we use pre-recorded video of "best-case" animal behavior rather than having visitors struggle with the interactional complexities of finding and focusing on the live specimens. We therefore created three different versions of an exhibit called *Glowing Worms*: (1) highly interactive (with changeable lighting, focus, dish location) and live specimens; (2) less interactive (with changeable lighting, focus) and live specimens, and (3) noninteractive (with pre-recorded video) and no live specimens. To the satisfaction of our exhibit developers, the results showed that visitors who saw one of the live, interactive versions of the exhibit stayed longer, rated the exhibit as more enjoyable, and were able to reconstruct more relevant details of their experience, than the noninteractive version (Allen, S., & Feinstein, N., manuscript in preparation, The effect of physical interactivity on visitor behavior and learning).

However, we have also discovered that "more is not necessarily better" when it comes to interactive features. In the same study of *Glowing Worms*, we found no significant differences between the experiences of visitors who used the more interactive version and those who used the less interactive version. This suggests that, at least for this exhibit, having some form of interactivity with live animals was key, but adding more did not improve the experience or enhance learning. In addition, studies of several other exhibits from the Exploratorium's collection have led us to identify five common pitfalls of designing exhibits with high levels of interactivity or multiple interactive features: (1) multiple options with equal salience can overwhelm visitors, (2) interactivity by multiple simultaneous users can lead to disruption, (3) interactivity, even by a single visitor, can disrupt the phenomenon being displayed, (4) interactive features can make a critical phenomenon difficult to find, and (5) secondary features can displace visitors' attention from the primary one (Allen & Gutwill, 2004). We interpret such findings as suggesting that exhibits may have an optimal degree of interactivity, and that formative evaluation is essential for ensuring that the interactive features work together harmoniously.

While museum researchers do not always have the resources to conduct rigorously controlled experimental studies, we can also learn from the summative evaluation findings of our exhibit collections. Interestingly, some of the Exploratorium's most attractive and sustaining exhibits in recent years have used little or no physical interactivity at all (for example, *Energy from Death*, a glass case containing animals in various stages of decomposition and consumption by smaller creatures; or a display of seven optical illusions at which more visitors stopped at all seven than stopped at any one). Even more interesting were the results of an analysis of visitors' conversations in a temporary exhibition about *Frogs* (Allen, 2002). For these microphoned visitors, the exhibits containing live animals in glass terrariums were more attractive and evoked significantly more frequent and more diverse "learning-talk" (conversations showing evidence of learning) than the physically interactive exhibits. The single exhibit in the exhibition that evoked the greatest diversity of learning-talk (as well as the longest holding times) was a videotape of frogs catching and eating their prey. This suggests that, while recognizing the power of interactive experiences, we should be skeptical about sweeping claims that interactivity is essential to learning, or even that it necessarily creates the most powerful, memorable, or attractive experiences in our museums.

In parallel with this work, the Exploratorium has been one of a number of institutions to push harder on the original model of an interactive exhibit, investigating alternative ways to create interactive exhibits that are "minds-on" as well as "hands-on." For example, we have created and studied exhibits that elicit "active prolonged engagement" (APE) among visitors, combining access to phenomena with opportunities for deeper cognitive experiences. The interim summative evaluation (Perry & Tisdal, 2004) showed that APE exhibits evoked a visitor experience that differed from more traditional exhibits (such as *Water Standing on Air,* Figure 1) in several ways. First, APE exhibits engaged visitors for more than twice as long as traditional exhibits. Second, they elicited a broader variety of "driving questions," defined as the initial thought that visitors had as they approached the exhibit. While the traditional exhibits shared a similar driving question for visitors (specifically, "What's going on here?"), APE exhibits elicited a range of questions, including "Can I do this?" (at an electrical circuit construction) and "What fascinating aspect of this can I see and share?" (at a sheet of freezing ice that visitors can melt with their hands or view through polarizing filters). Finally, visitors' physical interactions with APE exhibits varied in both pattern and sequence more than with the traditional exhibits, and labels were often referred to later in the sequence of exhibit use, suggesting that APE exhibits were eliciting more self-generated explorations by visitors than the more traditional exhibits.

Perry and Tisdal suggest that one factor in accounting for the long holding times at APE exhibits is that, unlike the traditional exhibits studied, the APE exhibits were designed to support the use of exhibits by more than one member of a social group. This is compatible with findings by Borun and Dritsas (1997) that exhibits that allow for multiple simultaneous users facilitate family learning, at least in the absence of the kinds of interference problems that Allen and Gutwill (2004) describe. In addition to the social aspect of design, Gutwill (2002) has explored some other characteristics of the exhibits that account for their impact on visitor behavior. In a detailed study of the APE exhibit *Downhill Race,* he assessed two types of outcome: the degree to which visitors articulated the correct principle underlying the exhibit (in this case, that disks with more mass at their center roll faster down the slope) and the degree to which they were engaged in inquiry with the exhibit (coded as a combination of physical and intellectual engagement). He found two interesting trade-offs in terms of exhibit design: (1) When the exhibit was used with wooden disks specially designed to heighten the salience of the key variable of mass distribution, more visitors succeeded in figuring out the correct principle behind the disks' behavior. However, the overall holding time decreased, and fewer visitors who could read the principle in the label engaged deeply with the exhibit. (2) A more traditional exhibit label (containing an invitation to interact and then an explanation of the underlying principle) resulted in more visitors articulating the correct principle, but led them to be less deeply engaged with the exhibit. This research highlights a tension that can arise between designing exhibits to support scientific content versus scientific inquiry processes.

Conceptual Coherence: An Ongoing Challenge

In defining constructivism in the classroom context a decade ago, Driver et al. (1994) argued that "Learners need to be given access not only to physical experiences but also to the concepts and models of conventional science" (p. 7). However, museum researchers know that visitors to science museums tend to be highly literal and concrete in their interpretations of exhibits (e.g., Falk & Dierking, 1992; Gammon, 1999). In such an environment, how can we hope to create exhibit collections that successfully communicate the abstract concepts, themes, and models of science?

The original Exploratorium model was one of untested curricular idealism. As Oppenheimer said, "we view a science museum as a collection of props that constitute an interlocking web of mini-curricula" (1986, p. 6). Such mini-curricula were known to staff, and the hope was that visitors would also become aware of the intended connections underlying different groups of exhibits, perhaps with a little aid from loose clusterings, area titles, and label-text mentioning related exhibits. However, evaluation showed that most visitors did not fully recognize the intended themes, especially if they were abstract (such as feedback) rather than phenomenological (such as sound). Visitors did not easily infer such concepts from multiple, individually designed exhibits.

In the last decade we have explored various techniques for making abstract concepts and themes more apparent to visitors, including heightened selectivity about which exhibits to include in a collection, linear sequencing of exhibits, unified design aesthetics among exhibits in a cluster, advanced organizers (conceptual and/or spatial overviews at the entry to thematic collections), and labels that echo and reinforce the abstract theme rather than focusing on the individual exhibit. None of these are new to the museum field, but their effectiveness depends on their precise design, and their use in combination brings in new untested possibilities for visitor learning. We conducted one controlled study in which we discovered that the addition of simple partitions around a group of related exhibits increased the percentage of visitors who correctly identified their common theme from 30 to 51%, highlighting the impact of environmental factors on learning (Allen, 2003).

We have had some large-scale success with these methods. For example, our recent *Seeing* Collection, funded by the National Science Foundation, partly replaced our original *Light*, *Vision*, *Color*, and *Optics* collection, and was designed to provide insights into how the eye and brain function together as well as how culture affects our understanding of what we see. Comparing the experiences of visitors to the original *Optics* collection and the new *Seeing* collection, external evaluator Whitney (2003) found that significantly more visitors to *Seeing* talked about seeing as interpretation, mentioned specific parts of the visual system, said they had a new appreciation of the importance of attention in seeing, said they hadn't realized that people see things differently, and found that the exhibits related to their daily lives. These findings corresponded with four of the five original goals for the collection, and demonstrated that subtle and complex scientific ideas could be communicated successfully through groups of exhibits.

On the other hand, our *Traits of Life* collection was less successful in communicating its abstract theme—four characteristics common to all living things—to visitors. External evaluator Hein (2003) noted that our exhibition strategy of displaying a range of living creatures to illustrate their commonalities could also be interpreted as displaying their diversity, which many visitors noticed instead. Interestingly, Hein hypothesized about the likely cause of the two "alternative themes" most commonly cited by visitors (viz., cycles of life, and environmental interdependence), as follows: "[These concepts] have been extensively emphasized in both formal education and in popular literature and life science museum displays . . . Conceptual change is hard, and humans prefer to add new experiences, new data and new information to conceptual schemes they already have, rather than to accept new overarching ideas" (p. 36). It seems that our attempts at conceptual coherence, let alone conceptual change, face many of the same challenges as have been well documented in science classrooms over the last 40 years.

Achieving high levels of thematic clarity for exhibitions may be particularly difficult in an open environment where most exhibits are, by design, loosely related via a web of alternative possible connections. Visitors may simply not expect to look for strong inter-exhibit links; in a study of conversations in a temporary exhibition, Allen (2002) found that visitors made explicit connections among exhibits at only 5% of the exhibits at which they stopped.

Diversity of Learners: Helping Visitors Optimize Their Own Learning

If learning is fundamentally visitor-driven (the basic constructivist position), how can a single, static, unmediated exhibit collection support the huge diversity of learners who visit a science museum?

In 1992, the American Association of Museums published "Excellence and Equity," a call to museums to recognize education as central to their public service, and to be more inclusive of diverse people in all aspects of their operations and offerings. At the same time, the last decade brought to the fore several theoretical approaches to the idea of learner diversity. One was the theory of Multiple Intelligences (Gardner, 1991), which proposed that there are different cognitive styles for understanding the world, not limited to the verbal and logical-mathematical. A second was the "learning styles" systems proposed by Kolb (1984), McCarthy (1987) and others, that classified learners according to their preferred modes of perception and processing of information. A third was the recognition of different sensory modes of experience: visual, auditory, tactile, smell and even taste. These approaches validated what many museum educators already knew: visitors vary in their preferences, styles, and motivations for learning.

Museums have increasingly attempted to embrace such diversity by designing a broader range of learning experiences, often led by educators and supported by researchers. In one influential study, Borun and Dritsas (1997) determined that one of the seven characteristics of exhibits that facilitated learning by family groups was "multimodal," meaning that they appealed to different learning styles and levels of knowledge. Another key study validated the effectiveness of Universal Design, "the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design." (Center for Universal Design, 2002). Specifically, staff at the Boston Museum of Science modified a hall of dioramas based on universal design principles, with the goals of making the subject matter intellectually and physically accessible, and finding new methods for allowing participation by visitors regardless of their special needs (Davidson, 1991).They added a number of elements, including objects that afforded touching, listening, and smelling, as well as activity stations that invited visitors to make comparisons between animal body parts and human tools, or to examine different types of fur under a microscope and consider the implications for the animals' lifestyles. Davidson, Heald, and Hein (1991) showed that the changes improved the experience for visitors in general: dramatic increases in holding time, label use, and understanding of the area's major themes. A recent survey by Tokar (2003) suggests that designing for physical access has finally become common practice among institutions with hands-on science exhibits, partly in response to the Americans with Disabilities Act. However, universal design with its larger goals, including intellectual access and participation for all visitors, has not yet been integrated into mainstream design.

The Exploratorium has always supported multisensory learning modes, in part because of our content-focus on human perception in its various forms. However, we have been trying to design exhibits that appeal to a broader range of visitor learning styles, since our original exhibits emphasize concrete experience rather than abstract conceptualization, and active experimentation rather than reflective observation (to use McCarthy's framework). In particular, we have experimented with exhibit collections that are quieter, carpeted, with selective lighting and partitioning of the space, and exhibits that support observation and reflection rather than hands-on activity. Evaluation studies have shown that our visitors notice and value the diversity of such spaces within the museum (Hein, 2003; Gutwill-Wise et al., 2000).

Another area we have explored in the last decade is the use of narrative in exhibits, as an attempt to engage a more diverse audience and to explore the different kinds of learning that result. Narrative, loosely defined as personal storytelling and distinguished from scientific discourse by such qualities as personal viewpoint, event sequences, and relevance to human concerns, has long been recognized as a powerful exhibit technique in historical and cultural museums (e.g., Bedford, 2001; Rounds, 2002), and is even regarded as a fundamental mode of human thought by some psychologists (Bruner, 1986; Shank, 1990). In what has become an influential argument for constructivism, researchers such as Roberts (1997) and Silverman (1995) have argued in the last decade that narrative, and particularly narrative with multiple voices, should replace authoritative knowledge-dissemination as the iconic mode for museums to conduct their educational mission. However, narratives and personal stories have had a much less prominent role in science museums, where the dominant mode is still hands-on inquiry with a single-voiced authoritative explanation.

In our experimental exhibits, we have created opportunities for visitors to write, tell, or hear personal stories related to exhibits or exhibition topics. Our evaluations have suggested that, while visitors readily engage in telling or listening to stories on emotional topics such as love, AIDS, or the memories associated with precious objects from one's home country (Pearce, 2003; McLean, 2003), they are much less engaged by the stories we created to make existing phenomenon-based exhibits more personally meaningful (Gutwill-Wise & Allen, 2002). This latter endeavor has been part of a research project called *Finding Significance,* in which we created video-based additions to existing exhibits, and compared the effects of "narrative" video clips (in which four different people tell a personal story about their connection to the exhibit) with "inquiry" video clips (in which the same four people ask the audience short questions that invite further exploration or thinking about the exhibit) and with a control (the baseline exhibit without any video addition). Preliminary results (Allen & Gutwill, 2003) suggest that the inquiry videos were preferred by visitors, and helped them integrate the exhibit experience into their prior knowledge, while narratives showed little or no impact compared with the control condition. The exception was a narrative that took the form of an unfolding drama with a cast of young children: this showed some evidence of evoking personal connections in more visitors' lives than the control or inquiry conditions. Apart from the modest differences we have seen in terms of learning impact, we have come to appreciate the particular difficulties of creating effective narratives as enhancements to exhibits in our environment. These include: a noisy and chaotic environment that is not conducive to sitting and listening, an upper limit of a few minutes on the length of a narrative that will hold visitors' attention, visitors' preference for something concretely related to the exhibit rather than loosely associated or metaphorical, and attentional conflicts as visitors try to watch and listen to the videos and have their own group conversations at the same time. Also, we found it extremely challenging to find authentic narratives that were compelling to visitors and yet fundamentally about the exhibit, while at the same time avoiding the kinds of scientific explanations that science museums routinely provide to their usual audiences. Others (e.g., Martin, 1996) have explored the use of narrative on the scale of an exhibition or even a whole visit, but the most compelling such narratives remain those that tell the history of a life-endangering enterprise, and the field has yet to understand how we can harness narrative in the service of helping visitors understand exhibits about scientific phenomena and principles.

CONCLUSION

I have argued that designing science museum exhibits that engage visitors moment-by-moment and yet also support their science learning is a difficult but not impossible task. Given the challenges, I believe it is critical to support the design process with a strong program of research and evaluation, a commitment that was made by the Exploratorium a decade ago and continues to grow. I have highlighted our integrated approach and discussed four areas that have emerged as important for us: immediate apprehendability, physical interactivity, conceptual coherence, and diversity of learning modes.

Many of our studies suggest that more attention needs to be paid to user-centered design at all scales, from the precise affordances of exhibit controls to the layout and orientation of the surrounding environment. I have proposed that a key concept for framing the research on museum learning is that of immediate apprehendability, which reduces cognitive overload and frees visitors to focus on those aspects of the environment that are worthy of their attention. While we know some standard methods for achieving immediate apprehendability, such as using familiar activities as schemas for exhibit interactions, or placing highly readable and consistent directional signs throughout a museum to help visitors orient themselves, there is still much to learn. For instance, we still need to know more about the impacts of apprehendability versus challenge, and the way in which the tension between them plays out at different scales in the museum visit, at different times, and with different audiences. This will require studying the detailed prior knowledge of our visitors in situ, with respect to the exhibits and environments we create.

As museum staff, we also need to be explicit about our own intentions in terms of what aspects of our learning environment we regard as *appropriate* for visitors to struggle with and what we don't. We might all agree that visitors should not struggle to figure out how to open the front door (a problem we actually have, incidentally, due to handles that afford pushing but in fact pull to open). However, many aspects of learning are still controversial. For example: Should we make explanations of scientific phenomena easy to locate and understand, or do we want visitors to rise to the challenge of investigating phenomena in their own terms? Should we create more sequenced exhibits and linear paths to reduce the effort of navigation and connection-making among exhibits, or should we keep the floor-plan open because connection-making is exactly where we believe visitors should be spending their effort? These are ongoing questions requiring institutional prioritizing, as well as further research.

Our studies on physical interactivity have shown that it is not a simple and universal prescription for effective learning. While some of our research supports the field's general conclusions that interactivity enhances visitor engagement, understanding, and recall, we have also discovered that it is not always essential to a powerful and sustaining experience, and that too many interactive features may even hinder visitors' engagement and learning. We have also shown (along with other institutions) that interactive exhibits can be designed to elicit extended engagement and self-directed inquiry, particularly if they support engagement by multiple users and avoid a simple "one-line" explanation of a phenomenon. However, designing exhibits to maximize visitor-directed inquiry may sometimes decrease the chances that visitors will understand a specific scientific principle. Looking to the future, with interactivity on the increase in many museums, there is still a great need for further study of its impact on learning. In particular, the field badly needs more studies of nonverbal forms of learning because these may be the dominant forms for a three-dimensional physical interaction, especially for children. Nonverbal studies would also help to reveal the learning impacts of a controversial and increasingly common display technique we call "gratuitous interactivity," where the interactive features of an exhibit are unrelated to its central phenomenon.

Our efforts to investigate conceptual coherence across multiple exhibits have suggested that it is often difficult to make such curricular connections apparent to visitors (especially if the connecting theme is an abstract scientific concept or principle), in part because they tend to learn in a concrete, literal way at exhibits. On a related issue, we have found that visitors' perceptions of what an exhibit is fundamentally "about" are quite robust, and not easily changed by, say, the reworking of the exhibit label. Instead, it seems that conceptual coherence across exhibits depends on a number of factors, many of which are environmental, such as the presence or absence of partitions surrounding the exhibit group. Much useful research could still be done in this area, because such large-scale design issues are logistically difficult and expensive to test in real museum settings. How can we encourage visitors to make more connections among exhibits in a collection? What are realistic abstractions in relation to visitors' prior knowledge? And, if visitors are continually constructing personal meanings from our exhibits, to what extent can the abstract concepts and themes of science be framed and embodied to provide context and significance for an otherwise fragmented experience?

We have attempted to study visitors' diversity of intelligences and learning styles, mostly through the design of exhibit collections that incorporate a broader diversity of approaches and sensory modalities, and provide more ways to make everyday and personal connections to scientific material. Some of our studies have focused on the effects of incorporating multiple personal narratives into exhibits, a technique that has proven powerful in cultural and history museums or in exhibits on emotional topics, but which has so far been relatively ineffective at enhancing learning or personal meaning-making at phenomenon-based exhibits. More research is needed to determine whether our instantiations of narrative have simply been flawed, or whether perhaps the chaotic, high-energy environment of a science museum and the nonanecdotal nature of science combine to make storytelling an ineffective strategy for learning in these settings.

Finally, the last decade's work has highlighted for me the vital role of ongoing, focused research and evaluation. Effective design is highly nontrivial and can take many iterations and years of gradual evolution, even for simple devices such as telephones and radios. How much more challenging, then, is the design of a unique and novel exhibit that must be robust, easily usable by people of any age and background, and lead to the learning of some aspect of science or the world in a personally meaningful way? In the face of irreducible complexity of both physical systems and humans, we are unlikely to ever create generalizable enough design principles to obviate the need for research, prototyping, and evaluation. Much of this work will require careful and detailed study at many scales if we are to understand the myriad alternative ways in which visitors experience, interpret, and learn from our exhibition spaces.

I gratefully acknowledge feedback on an early draft of this article by Joshua Gutwill, Kathleen McLean, Beverly Serrell, and three anonymous reviewers. Much of the work reported in this article was funded by the National Science Foundation.

REFERENCES

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, NJ: Lawrence Erlbaum.

Allen, S. (2003, Oct.). To partition or not to partition: The impact of walls on visitor behavior at an exhibit cluster. Paper presented at the annual meeting of the Association of Science-Technology Centers, Minneapolis.

Allen, S., & Gutwill, J. (2003). Assessing techniques for facilitating meaning-making by science museum visitors. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Allen, S., & Gutwill, J. (2004). Designing science museum exhibits with multiple interactive features: Five common pitfalls. Curator, 47(2), 199–212.

American Association of Museums. (1992). Excellence and equity: Education and the public dimension of museums. Washington DC: American Association of Museums.

Bedford, L. (2001). Storytelling: The real work of museums. Curator, 44(1), 27–34.

Borun, M., & Dritsas, J. (1997). Developing family-friendly exhibits. Curator, 40(3), 178–196.

Bruner, J. (1986). Actual minds, possible worlds. Cambridge: Harvard University Press.

Center for Universal Design (2002). What is Universal Design? Definition, http://www.design.ncsu.edu/cud/univ_design/ud.htm.

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, DC: American Association of Museums.

Davidson, B. (1991). New dimensions for traditional dioramas: Multisensory additions for access, interest and learning. Boston, MA: Museum of Science.

Davidson, B., Heald, C. L., & Hein, G. E. (1991). Increased exhibit accessibility through multisensory interaction. Curator, 34(4), 273–290.

Dreyfus, H. (1967). Designing for people. New York: Paragraphic Books.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23(7), 5–11.
Evans, G. (1995). Learning and the physical environment. In J. H. Falk & L. D. Dierking (Eds.), Public institutions for personal learning: Establishing a research agenda (pp. 119–126). Washington DC: American Association of Museums.
Falk, J., & Balling, J. (1977). An investigation of the effect of field trips on science learning. Final report grant number SED77-18913. Washington, DC: National Science Foundation.
Falk, J. H., & Dierking, L. D. (1992). The museum experience. Washington, DC: Whalesback Books.
Falk, J. H., Koran, J. J., Dierking, L. D., & Dreblow, L. (1985). Predicting visitor behavior. Curator, 28, 249–257.
Fry, R. (2002). Delightful sound and distracting noise. Journal of Museum Education, 27(1), 14–17.
Gammon, B. (1999). Everything we currently know about making visitor-friendly mechanical interactive exhibits. Informal Learning Review, 39, 1–13.
Gardner, H. (1991). The unschooled mind: How children learn, how schools should teach. New York: Basic Books.
Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), Perceiving, acting, and knowing. Hillsdale, NJ: Erlbaum Associates.
Gutwill, J. (2002). Providing explanations to visitors affects their inquiry behavior: A study of the Downhill Race exhibit. (Available from the Exploratorium, 3601 Lyon St, San Francisco, CA 94123.)
Gutwill-Wise, J., & Allen, S. (2002). Finding significance: Testing methods for encouraging meaning-making in a science museum. Current Trends in Audience Analysis, 15, 5–11.
Gutwill-Wise, J., Soler, C., Allen, S., Wong, D., & Rezny, S. (2000). Revealing bodies summative evaluation. Available in summary form at www.exploratorium.edu/partner/evaluation.html.
Hayward, D. G., & Brydon-Miller, M. L. (1984). Spatial and conceptual aspects of orientation: Visitor experiences at an outdoor history museum. Journal of Environmental Systems, 13(4), 317–332.
Hedge, A. (1995). Human-factor considerations in the design of museums to optimize their impact on learning. In J. H. Falk & L. D. Dierking (Eds.), Public institutions for personal learning: Establishing a research agenda (pp. 105–117). Washington DC: American Association of Museums.
Hein, G. E. (2003). Final summary evaluation report for Traits of Life, a collection of life science exhibits. (Available from the Exploratorium, 3601 Lyon St, San Francisco, CA 94123.) Also available at www.exploratorium.edu/partner/evaluation.html.
Kolb, D. A. (1984). Experiential learning: Experience as a source of learning and development. Prentice Hall.
Loewy, R. (1951). Never leave well enough alone. New York: Simon & Schuster.
Martin, L. M. W. (1996). A Vygotskian approach to the design of a science center. Paper presented at the annual meeting of the American Educational Research Association, New York.
Maxwell, L. E., & Evans, G. W. (2002). Museums as learning settings: The importance of the physical environment. Journal of Museum Education, 27(1), 3–7.
McCarthy, B. (1987). The 4MAT System: Teaching to learning styles with right/left mode techniques. Barrington, IL: EXCEL, Inc.
McLean, K. (Ed.). (2003). Visitor voices [Special issue]. Journal of Museum Education, 28(3).
Norman, D. A. (1988). The design of everyday things. New York: Doubleday.
Oppenheimer, F. (1986). Working prototypes: Exhibit design at the Exploratorium. Exploratorium Publications, ISSN-0-943451-081-06.
Papanek, V. (1971). Design for the real world: Human ecology and social change. New York: Pantheon.
Pearce, M. (2003). "They said that the glass is full of friendship": Visitor stories in a memory exhibition. Journal of Museum Education, 28(3), 26–30.
Perry, D., & Tisdal, C. (2004). Going APE! at the Exploratorium: Interim Summative Evaluation Report. (Available from the Exploratorium, 3601 Lyon St., San Francisco, CA 94123.)
Richards, W. H., & Menninger, M. (2000). A discovery room for adults. In J. S. Hirsch, & L. H. Silverman (Eds.), Transforming practice: Selections from the Journal of Museum Education 1992–1999 (pp. 301–311). Washington, DC: Museum Education Roundtable. (Also published in 1993, Journal of Museum Education, 19(1), 6–11.)

Roberts, L. (1997). From knowledge to narrative: Educators and the changing museum. Washington, DC: Smithsonian Institution Press.

Rounds, J. (2002). Storytelling in science exhibits. Exhibitionist, 21(2), 40–43.

Rowe, S. (2002). The role of objects in active, distributed meaning-making. In S. Paris (Ed.), Perspectives on object-centered learning in museums (pp. 19–35). Mahwah, NJ: Lawrence Erlbaum.

Schneider, B., & Cheslock, N. (2003). Measuring results: Gaining insight on behavior change strategies and evaluation methods from environmental education, museum, health and social marketing programs. (Available from the Coevolution Institute, 423 Washington Street, San Francisco, CA 94111.)

Semper, R. (1990, Nov.). Science museums as environments for learning. Physics Today, 50–56.

Serrell, B. (1996). Exhibit labels: An interpretive approach. Walnut Creek: AltaMira Press.

Shank, R. C. (1990). Tell me a story. Evanston, IL: Northwestern University Press.

Silverman, L. H. (1995). Visitor meaning-making in museums for a new age. Curator, 38(3), 161–170.

Tokar, S. M. (2003). Universal design: An optimal approach to the development of hands-on science exhibits in museums. Unpublished master's thesis, Excelsior College, Albany, NY. Also available at http://www.library.excelsior.edu/theses/tokar.html.

Walter, C. (in press). Whodunit. In K. McLean & C. McEver (Eds.), Toward best practices: Conversations about exhibition development (working title). San Francisco: Exploratorium Publications.

Whitney, K. (2003). Seeing collection pre-post study final results. (Available from the Exploratorium, 3601 Lyon St., San Francisco, CA 94123). Also available at www.exploratorium.edu/partner/ evaluation.html.