Comparing the Visitor Experience at Immersive and Tabletop Exhibits

TITLE

Comparing the Visitor Experience at Immersive and Tabletop Exhibits

AUTHORS

Dr. Toni Dancstep, Senior Researcher
Visitor Research and Evaluation
Exploratorium
tdancu@exploratorium.edu
Toni Dancstep (formerly Dancu) leads learning research in the museum field and exhibit evaluation studies at the Exploratorium.

Dr. Joshua P. Gutwill, Director of Visitor Research
Visitor Research and Evaluation
Exploratorium
jgutwill@exploratorium.edu
Josh Gutwill oversees all learning research and evaluation studies at Exploratorium exhibits and floor programs.

Lisa Sindorf, Research Associate
Visitor Research and Evaluation
Exploratorium
lsindorf@exploratorium.edu
Lisa Sindorf conducts learning research and evaluation studies at Exploratorium exhibits and floor programs.

ABSTRACT

Many museum professionals believe that immersive exhibits—those that surround visitors—provide more attractive, engaging and effective learning experiences than tabletop exhibits. We investigated this claim by comparing visitors’ experiences of the two exhibit types, using pairs of exhibits that differed in scale (immersive vs. tabletop), but shared the same content and similar visitor activity. We randomly selected, videotaped, interviewed, and sent follow-up surveys to sixty families who experienced
immersive exhibits and sixty families who experienced tabletop exhibits. We found that each design type had strengths. Learners at immersive exhibits returned to the exhibit more often, mentioned the exhibits’ positive aspects, and saw themselves as part of the exhibit. Conversely, learners spent longer periods of time at tabletop exhibits, and engaged in more content-related reasoning. Study results partially support the view that immersive exhibits may be more fun and engaging than table tops. However, results also counter the expectations that being immersed in exhibit experiences will lead to greater physical and intellectual engagement.

**KEYWORDS**

Immersive exhibits, tabletop exhibits, museum exhibit, interactive, STEM learning, mathematical reasoning, spatial reasoning, spatial language, geometry, applied research
Introduction

What is the educational power of immersive experiences? In describing the diorama-like worlds in Wes Anderson’s films, contemporary author Michael Chabon suggests that immersivity may not be our most potent educational tool: there is a paradoxical power of the scale model; a child holding a globe has a more direct, more intuitive grasp of the earth’s scope and variety, of its local vastness and its cosmic tininess, than a man who spends a year in circumnavigation” (2013).

Yet, when we turn to the world of museums, it is unclear whether this paradox applies. According to prior research, many museum professionals believe that immersive exhibits, compared with hands-on tabletop exhibits, attract a broader range of visitors, engage visitors for longer periods of time, and are more memorable—leading to enhanced visitor learning (Gilbert 2000, 2002; see also Coombes 1995). With limited research supporting either standpoint, we designed a study to assess the comparative impacts of the two types of museum experiences. Specifically we endeavored to answer the research question (depicted in Figure 1): How does the museum visitor experience differ at exhibits that visitors can manipulate with their hands (tabletops) compared to exhibits that they can get inside (immersives)?

[Insert Figure 1 about here.]

[Figure 1 Caption] Figure 1. Pictorial Representation of Research Question.
Immersive exhibits that surround visitors are becoming more and more common, especially in science and natural history centers (Gilbert 2000, 2002). In 2000, Hallie Gilbert conducted a seminal study to learn why museums were creating more immersive exhibits. She surveyed and interviewed museum and exhibit development directors around the nation. Gilbert’s explorations revealed that museum professionals perceived immersives to attract a broader range of visitors and engage those visitors for longer periods of time. The practitioners also believed the immersive exhibits to be more memorable, and therefore more capable of leading to enhanced visitor learning (Gilbert 2000, 2002; see also Coombes 1995).

Surprisingly, little research has been conducted to determine the extent and manner in which immersive exhibits provide a better visitor experience compared to more typical tabletop exhibits. A few research studies and summative evaluations have found that immersive exhibits can have strong emotional impacts and leave memorable impressions, often achieving “iconic” status (Anderson and Roe 1993; Bitgood, Ellingsen and Patterson 1990; Borun 2004; Korn and Jones 2000; Phillips 1999). Other evaluations indicated that certain immersive environments encouraged long holding times (the time a person spends using or gazing at an exhibit) (Gilbert 2000, 2002). In one instance, researchers found that changes to make an envirorama (a diorama that visitors can step into) more user-friendly and inclusive of immersive features resulted in an increase to the number of visitor stops, the amount of reading done, and the amount of time spent by visitors (Harvey, Loomis, Bell and Marino 1998).

Our expectations were aligned with these museum directors’ beliefs, and with the prior, limited, research findings: We hypothesized that immersive exhibits would be more
engaging, enjoyable, social, and memorable than tabletops. We also anticipated that along with enhanced engagement and enjoyment would come greater intellectual involvement.

Understanding the comparative effects of immersive and tabletop exhibits on learners is important because immersives often cost more money to build, take more time to create, are more difficult to prototype, and require more physical space. While some researchers and evaluators have studied immersive exhibits, none have yet explored the differences between visitors’ learning experiences at immersive and tabletop exhibits. The present study aims to contribute to the museum field’s knowledge base on immersive exhibits by comparing visitors’ physical, affective, and cognitive engagement with each of these types of exhibits.

[A Head] Defining Immersive Exhibits

For the purpose of this research, we defined immersive exhibits as exhibits in which the exhibited phenomenon partially surrounds the visitor by extending from in front of their body (including virtual aspects of a visitor’s body, such as a shadow or image in a mirror) to behind their back, over their head or under their feet (Dancu and Gutwill 2009). We further narrowed our focus to immersive exhibits that involved the visitor’s whole body—their torso as well as their hands or feet—and those that were interactive—where visitors’ inputs directly affected the exhibits’ outputs (Allen and Gutwill 2004; Mclean 1993). This definition emerged from a review of the small body of academic literature on immersion, along with a broad survey of exhibits that had been deemed immersive by museum professionals.
In the museum literature, the term “immersive” has been used to describe many different types of informal learning experiences. For example, “immersive” describes certain replicated environments that recreate realistic, life-sized settings that place visitors in a certain time, location, or situation (Bitgood, et al. 1990; Gilbert 2000, 2002; Gyllenhaal 2002). These exhibits, such as the recreated underground Coal Mine exhibit at the Museum of Science and Industry, are often multi-sensory and multi-modal (see also Birney 1990; Jones and Wageman 2000; Larson 2002). “Immersive” may also describe virtual experiences in simulated or online worlds, such as Second Life or Minecraft (see also Dede, Salzman, Loftin and Ash 2000; Gyllenhaal 2002; Tecchia, Ruffaldi, Frisoli, Bergamasco and Carrozzino 2007; Twiss-Garrity and Fisher 2007; Psotka 1995). And Falk et al. (2004) used the term “whole body” to describe larger-than-life immersive replicas of objects or phenomena. These whole-body exhibits enable visitors to physically enter the replica, creating a multi-sensory and kinesthetic experience. Such exhibits are epitomized by the iconic walk-through Giant Heart exhibit at the Franklin Institute.

Without a unified definition of “immersive” in the literature, and with so many kinds of experiences falling into the category, we decided to create a taxonomy that would help narrow our undertaking. We began exploring the museum fields’ existing immersive exhibits by first identifying as many immersives as possible via a literature review, a short Google search, and a query on the listserv of the Association of Science and Technology Centers. This produced a list of fifty immersive exhibits. From this set of immersives and the supporting literature, we developed a list of exhibit attributes that could be employed in a cluster analysis, which in turn would reveal the taxonomy, or categories, of immersive exhibits. Once we had a large number of attributes (around
twenty five), we scored each exhibit on the presence or absence of each. We then narrowed the set of attributes to those that were most distinguishing: those that were present at more than 75% of our exhibits, or absent from more than 75% of our exhibits. These distinguishing attributes were:

- Authenticity—contains authentic objects
- Geographic—location is an important aspect of the exhibit
- Involves Sound—sound is used to set the scene
- Visitor Interactivity—visitor inputs affect exhibit outputs
- Whole Body—requires visitor to use their torso beyond just walking through or sitting in a ride
- Themed Exhibition—part of a larger themed exhibition

Next, we used those distinguishing attributes to conduct a cluster analysis, which statistically groups sets of objects, such as exhibits, into clusters that are more similar within the cluster than they are between different clusters. The scores on the distinguishing attributes helped us identify the following five classes of immersive exhibits that emerged from the analysis:

- Recreated with a Focus on Geography. These exhibits recreate an experience where the geographic location, regardless of authenticity, is highly important to that experience (e.g., *Dynamic Earth Cave* at the Anniston Museum of Natural History in AL).
- Recreated with a Focus on Sound. These exhibits recreate an experience
where sound is highly important to the immersive aspect of that experience (e.g., Coal Mine at the Museum of Science and Industry in Chicago, IL).

- Recreated with a Focus on Whole-body. These exhibits recreate an experience, and enable visitor to use their whole body (typically torso) to investigate and explore that experience (e.g., Giant Heart at the Franklin Institute in Philadelphia, PA).

- Whole-body Interactive. These exhibits enable visitor to use their whole body (typically torso) to investigate and explore a phenomenon, and visitors directly affect the outcome of the exhibit/activity (e.g., Sky Bike, a high-wire bicycle at the Franklin Institute in Philadelphia, PA or Colored Shadows at the Exploratorium in San Francisco, CA).

- Whole-body Non-interactive. These exhibits enable visitor to use their whole body (typically torso) to investigate and explore a phenomenon, but visitors cannot directly affect the outcome of the exhibit/activity (e.g., Earthquake House at the Oregon Museum of Science and Industry, or Tactile Dome at the Exploratorium in San Francisco, CA).

The research and development team (see project description below) chose to create and study Whole-body Interactive immersive exhibits, and compare those exhibits to tabletop exhibits. For the purpose of this comparative research, we defined tabletop exhibits as hands-on interactive exhibits (typical of science museums), where the exhibit exists only in front of the visitor (i.e., does not surround the visitor).
Our study was situated within a larger National Science Foundation-funded exhibition development project, Geometry Playground. The Exploratorium, San Francisco’s museum of science recently developed Geometry Playground, a traveling exhibition where visitors explore and experiment with geometric shapes—turning, tiling, and designing with handheld manipulables, and also entering larger-scale immersive structures that might allow for alternate experiential impacts. Geometry Playground aimed to integrate “body-in” with “hands-on” learning.

There were two reasons this project was ideal for comparing immersive and tabletop designs. First, geometry exhibits allowed us to study a thinking skill that is important to multiple aspects of science and mathematics: spatial reasoning (e.g., Battista and Borrow 1997; Ben-Chaim, Lappan and Houang 1989; Pallrand and Seeber 1984; Tartre 1990; Tracy 1987). For example, researchers have found that the use of physical interactives in a classroom setting can lead to enhanced spatial reasoning scores (Battista and Borrow 1997; Ben-Chaim, Lappan and Houang 1985; Clements 1999; Clements and Battista 1990), and even some children’s toys have been related to spatial reasoning abilities (e.g., Baenninger and Newcombe 1995; Cherney and London 2006; Tracy 1987; Voyer, Nolan and Voyer 2000). Further, research has linked large-scale interactions that require gross-motor exploration to enhanced spatial reasoning skills (Acredolo 1977; Hazen 1983). Spatial reasoning is also necessary when interacting with geometry concepts, regardless of scale.
The second (and more important) reason we chose to do this study as part of Geometry Playground is that geometry is scale invariant; geometric shapes and phenomena are the same regardless of size. That is, a one-inch square cube that you hold in your hands is mathematically equivalent to a ten-foot square cube that you explore from the inside out. Similarly, scaling up or down does not alter geometric rules, such as transformation and rotation. This enabled us to create comparable content pairs: immersive and tabletop exhibits that explored the exact same content via similar visitor activity.

Situating the research directly within the museum development process allowed us to intentionally create study-worthy geometry exhibits that featured the tensions between immersive and tabletop experiences. Content pairs meant that the three exhibits in the Immersive set were matched according to content with the three exhibits in the Tabletop set. Each pair was also experientially matched as much as possible, while still respecting and incorporating the particular benefits that each type of exhibit affords. In the immersive exhibit, *Personal Space*,\(^1\) visitors walk around on a large floor pad while the exhibit draws a line exactly halfway between every two people, according to the Voronoi algorithm. At the tabletop exhibit, *Dividing Space*, visitors move red pegs to different locations on the board and roll blue pucks through the pegs. The exhibit draws a line exactly halfway through every two objects, according to the Voronoi algorithm. Figure 3 shows each of the Voronoi exhibits used for the study.

[Insert Figure 3 about here.]
Figure 3. Voronoi Diagram exhibits explored a mathematical rule for dividing space.

The second pair of exhibits in the study involved transformations between 2D and 3D shapes. At the immersive exhibit, \textit{Spin a Shape Around You}, visitors hold a solid line at different angles and rotate it around themselves by spinning their bodies. The exhibit detects the line and, using long exposure, displays the 3D rotational solid created by the spinning motion on a large screen in front of the visitors, along the visitors’ axis of symmetry. The tabletop version, \textit{Spin a Shape}, asks visitors to place predesigned cards into a slot that spins the cards. Visitors guess which of the shapes on the label each card will create. Persistence of vision (the retention of an image on your retina long enough to build up a complete image) turns the 2D lines into a 3D rotational solid along the axis of symmetry (Exploratorium 2015). Figure 4 shows these two exhibits featuring 2D to 3D transformations.

Figure 4. Rotational solids exhibits showcased the axis of symmetry for rotational shapes.

In the third pair of exhibits used in the study, visitors play with visual transformations in a cylindrical mirror (Figure 5). The immersive version, \textit{Distorted Chair}, offers visitors a mathematically stretched and curved bench that is reflected as a normal chair in a large cylindrical mirror. (The warping of the image in the mirror, which
makes this distorted bench look normal, is called “anamorphosis.”) Visitors sit on the bench and use their bodies to explore the horizontal compression created by the shape of the mirror. At the tabletop exhibit, *Distorted Drawing*, visitors draw on a mathematically stretched and curved chalkboard grid while viewing their drawings in a cylindrical mirror. The cylindrical mirror reflects a horizontally compressed image of any drawing, so straight lines appear curved and curved lines appear straight.

[Insert Figure 5 about here.]

[**Figure 5 caption**]  **Figure 5.** Transformation exhibits highlighted anamorphic conversions within cylindrical mirrors.

This research aimed to answer the question: In what ways and to what extent do visitor learning experiences (physical, affective, and cognitive engagement) differ at immersive and tabletop exhibits? Below, we describe the methods used to answer this question, followed by a discussion of the results, implications, and limitations of this study.

**Methods**

**Study Design and Data Collection**

We employed a quasi-experimental design to compare visitor experiences at the immersive and tabletop exhibits. Immersive and Tabletop exhibit sets were switched out weekly over the course of the summer, so that each exhibit grouping was used by visitors on different days of the week, weeks of the month, and months over the summer data
collection period. This enabled us to approximate random assignment of visitors to treatment conditions, and then average visitors’ experiences within each grouping. Data collectors (those recruiting, videotaping, and interviewing visitors) were blind to the purpose of the research and did not hear the term “immersive” nor the true names of the exhibits while collecting data for the project.

We recruited sixty families to each set of exhibits, 120 families total. Visitors were randomly selected from the museum floor and screened to avoid special groups (such as museum members and homeschoolers) who might use the exhibits differently than typical museum visitors. Visitor groups with at least one child between the ages of eight and twelve were asked to participate. Slightly under half (42%) agreed to participate. While the whole visitor group explored the exhibits together, we randomly selected an adult and eligible child from the group to wear microphones, be followed by ceiling-mounted cameras, and respond to a set of interview questions after using the exhibits. Figure 6 shows the experimental design.

[Insert Figure 6 about here.]

[Figure 6 caption] Figure 6. Experimental Design

Each group visited only one set of exhibits (immersives or tabletops); when tabletops were on the museum floor, the immersives were removed entirely (and vice versa). Exhibits were placed midway through the museum, allowing visitors to become familiar with the museum experience before joining the study. Visitors were asked to use all three exhibits for as long as they liked. It is important to note that although recording
visitors undoubtedly created reactivity by visitors at the exhibits, that reactivity would be present in both conditions, thus washing out the effect on our comparisons. After using the exhibits, the participating adult and child were interviewed separately. Six weeks later, a brief follow-up survey was emailed to all adult participants (asking that the adult and child respond together). Between 35% and 43% (tabletops and immersives, respectively) of the participants responded to the follow-up survey.

[B-head] Statistical Analyses

All interview responses and 25% of video data were coded by two research assistants separately. Inter-rater agreement for all codes was high (Multon 2010; Fleiss, Levin and Paik 2004). Planned comparisons measured differences between participants who experienced the Immersive exhibits and those who experienced the Tabletop exhibits.

Below, we describe each of the outcome variables we used to investigate differences in visitors’ learning experiences, followed by the statistical results and effect sizes. Effect sizes help us understand the magnitude of a difference, independent of the number of participants in the study. Results are summarized in Table 1.

[C-head] Demographic Variables

There were no significant differences by treatment condition in the following demographic variables:
• Participating adult’s gender / age
• Participating child’s gender / age
• Total group size

[C-head] Physical Engagement Variables

To determine whether immersives were more engaging than tabletops, we analyzed the video data for two distinct physical behaviors: time spent at the exhibits, and the number of times adults and children returned to an exhibit to use it again.

[D-head] Holding Time

Definition: Duration of time spent at each exhibit, watching (feet planted and facing) or manipulating (hands-on). This code also included time spent during return visits to each exhibit. Holding time has long been considered a measure of engagement in the museum field (Falk 1983; Humphrey and Gutwill 2005).

Results for Holding Time were:

• ADULTS: On average, adults spent significantly more time at Tabletops (4.05 minutes) than at Immersives (2.89 minutes). The effect size for this difference is considered large.

• CHILDREN: On average, children spent significantly more time at Tabletops (4.06 minutes) than at Immersives (2.89 minutes). Again, the effect size is considered large.
Exhibit Returns

Definition: Number of times participants returned to any of the exhibits, throughout the groups’ participation in the study. Exhibit returns were considered evidence of how compelling the exhibit was to the visitor. It is worth noting that the large nature of immersive exhibits made it difficult to create a reliable coding scheme. Consequently, our coding scheme for “using” the exhibit necessarily included simply walking over or through the exhibit, which could artificially inflate the number of Exhibit Returns at immersives.

Results for Exhibit Returns were:

- **ADULTS**: Adults were significantly more likely to return to the Immersive exhibits (40%) than to the Tabletop exhibits (20%). The size of this effect is considered small, but notable.
- **CHILDREN**: Children were similarly likely to return to Immersives (50%) as to Tabletops (42%). The effect size of any difference is negligible.

Affective Variables

To assess whether the affective impact of immersives was greater than that of tabletops, we coded participants’ responses to two interview questions. The first asked them to rate how much they enjoyed the exhibits they used and then explain their rating; the second asked whether there was anything unusual or different about the exhibits they used as compared to the rest of their museum experience. We created four codes to
categorize visitors’ responses: Exhibit Attitudes—Positive; Exhibit Attitudes—Negative; Social Attitudes—Positive; and Social Attitudes—Positive.

**[D-head] Enjoy Using**

Definition: Visitors’ ratings, on a scale of 1-5, regarding how much they enjoyed using each exhibit (photos of the exhibits accompanied each rating request).

Results for Enjoy Using were:

- **ADULTS**: On average adults rated their enjoyment at the Immersives (3.83) similarly to adults’ ratings of enjoyment at the Tabletops (3.70). However, the size of the effect, while small, is considered notable, in favor of the Immersives.

- **CHILDREN**: Children, on average, rated their enjoyment at the Immersives (4.09) similarly to children’s’ ratings of enjoyment at the Tabletops (3.98). Again, the size of the effect is considered small yet notable, in favor of the Immersives.

**[D-head] Attitudes About Exhibit Experiences**

Definition: Visitors’ positive and negative comments regarding the exhibit experience. A coding scheme was developed using keyword analysis, a linguistics tool that identified potentially distinguishing words between the immersives and tabletops data sets, as well as words that were unique to either data set (Caldas-Coulthard and Moon 2010; Sindorf 2011; Kucera and Francis 1967). Examples of positive attitudes about exhibit experiences (Exhibit Attitudes—Positive) include mentions of fun (“It was
hilarious. When I [sat] down it gave me the shortest legs I've ever seen in my life.”), creative and visual aspects (“I liked the lights and different shapes you could make.”), and intellectual aspects (“. . . we figured it out,” and “It held the kids’ attention.”). Examples of negative attitudes about exhibit experiences (Exhibit Attitudes—Negative) involve talk about the boring or difficult aspects (“It didn’t hold my son’s attention,” and “I don't think in these terms. I couldn't figure out why it was happening.”).

Results for Exhibit Attitudes—Positive were:

- **ADULTS:** Adults were similarly likely to express positive thoughts about the exhibit experience at Immersives (45%) and Tabletops (33%). However, the effect size, in favor of the Immersives, is small but notable.

- **CHILDREN:** Children expressed positive thoughts about the exhibit experience at Immersives (43%) about as often as at Tabletops (32%). Again, the effects size is considered small but notable, in favor of the Immersives.

Results for Exhibit Attitudes—Negative were:

- **ADULTS:** Adults were similarly likely to mention negative thoughts about the exhibit content at Immersives (68%) as at Tabletops (80%). However, the effect size is considered small but notable, with more negative comments occurring at Tabletops.
• CHILDREN: Children mentioned significantly more negative thoughts about the content at Tabletops (70%) than at Immersives (50%); the effect size is small yet notable.

[Definition]
Attitudes About Social Aspects

Definition: Visitors’ positive and negative comments about using the exhibits with other people. Again, a coding scheme was developed using a keyword analysis, as described above. Positive attitudes about the social aspects of exhibits (Social Attitudes—Positive) included mentions of using an exhibit with others or enjoying watching others use it (“I enjoyed not just doing it myself but watching these [kids] do it, so I liked doing it as a family.”). Negative attitudes about the social nature of the exhibits (Social Attitudes—Negative) referred to mentions of crowding (“We tried to make a square but couldn't because there were too many people.”).

Results for Social Attitudes—Positive were:

• ADULTS: Adults were significantly more likely to express positive attitudes about the social aspects of the Immersives (60%) than the Tabletops (28%), with a medium sized effect.

• CHILDREN: Children were significantly more likely to express positive attitudes about the social aspects of the Immersives (22%) than the Tabletops (5%), with a small yet notable sized effect.

Results for Social Attitudes—Negative were:
• **ADULTS:** Adults were similarly likely to say something negative about the social nature of the Immersives (23%) and the Tabletops (20%), with a negligible effect size.

• **CHILDREN:** Children were significantly more likely to say something negative about the social aspects of the Immersives (18%) than the Tabletops (3%), with a small yet notable effect size.

[C-head] **Geometry Thinking Variables**

To explore whether learners’ geometry-related thinking was greater at immersive than at tabletop exhibits, we coded interview and video data. In the interviews, we coded the two questions mentioned above (how much they enjoyed the exhibits they used and why, and whether there was anything unusual or different about the exhibits they used as compared to the rest of their museum experience) for references to participants’ selves or bodies in relation to the exhibit (an aspect of embodied learning). The video data were analyzed for talk involving spatial language, which has been related to spatial reasoning, and found to help improve spatial reasoning skills (Casasola, Bhagwat and Burke 2009; Gentner and Christie 2008; Pruden, Levine and Huttenlocher 2011).

[D-head] **Self-referencing Talk**

Definition: Visitors’ talk about themselves or their bodies in relation to the exhibits. We considered self-referencing talk to indicate that participants saw themselves and their bodies as integrated parts of the exhibit phenomenon, which is an important component of embodied learning (Nemirovsky, Tierney and Wright 1998). A coding
scheme was developed using a keyword analysis, described above. The Self-referencing Talk code included any mentions of the self when describing aspects of the exhibit, and could include the whole self (e.g., “You're able to see yourself on the screen and make the shape.”) or only body parts (e.g., “when you stand and extend your arms, you get the idea. Your arms look squished because they went to the side of the cylinder”).

Results for Self-referencing Talk were:

- ADULTS: Adults were significantly more likely to reference themselves at the Immersive exhibits (68%) than at the Tabletop exhibits (0%), and the size of the effect is large.
- CHILDREN: Children were significantly more likely to talk about themselves and their bodies at the Immersives (90%) than at the Tabletops (2%), with a large effect size.

[D-head] Low-level and High-level Spatial Reasoning Language

Definition: Visitors’ spatial reasoning language (SRs) at the exhibits, including average number of SRs per exhibit and average duration of SRs per exhibit. Spatial reasoning involves using one’s spatial perception, mental rotation capacities, and spatial visualization abilities (Linn and Petersen 1985) to understand, manipulate, reorganize and interpret shapes, space and their relationship (Tartre 1990). Based on the National Research Council’s (2006) types of spatial reasoning, we categorized visitors’ spatial reasoning language into three levels: Static, Dynamic and Causal, which we eventually collapsed into two levels. Low-level Spatial Reasoning Language (Low SRs) includes
spatial utterances that are *structural* by nature. Combining the static and dynamic, Low SRs describe the order, relation and pattern of parts (“It’s a square”) or processes (“Spin your body around”). High-level Spatial Reasoning Language (High SRs) includes spatial utterances that are *functional*; they describe how and why something works (regardless of accuracy). They tend to be cause and effect statements, such as “spin to make a cylinder.” The full coding manual is available on the Spatial Intelligence & Learning Center website under Tests & Instruments (Dancu, Gutwill and Sindorf 2015).

Results for Low-level Spatial Reasoning Language (Low SRs) were:

- **ADULTS:** Adults made a similar number of Low SRs per Immersive exhibit (8.97) as per Tabletop exhibit (9.95), with a small but notable effect size in favor of Tabletops. The average duration of adults’ Low SRs was also similar at Immersives (32.92 seconds) and Tabletops (38.90), again with a small but notable the effect size in favor of Tabletops.

- **CHILDREN:** Children expressed significantly more Low SRs per Tabletop exhibit (7.11) than per Immersive exhibit (5.61), which is a large effect. The duration of children’s Low SRs was also significantly longer at the Tabletop exhibits (21.78) than at the Immersive exhibits (17.05), with a small yet notable effect size.

Results for High-level Spatial Reasoning Language (High SRs) were:
• ADULTS: Adults expressed significantly more High SRs at the Tabletop exhibits (1.35) than at the Immersive exhibits (.88), with a large effect size. The duration of adults’ High SRs was also significantly longer at the Tabletops (11.61) than at the Immersives (8.26), with a small but notable effect size.

• CHILDREN: Children did not utter High SRs often enough to conduct the same analyses as above; instead we coded the presence or absence of High SRs for each child at each exhibit. Children were significantly more likely to express High SRs at the Tabletops (55%) than at the Immersives (23%), with a medium effect size.

[C-head] Memorability Variables

To evaluate whether immersive exhibits are more memorable than tabletops, we analyzed video data for moments indicating that a learner may have wished to remember the experience; and we surveyed participants 6 weeks after the visit and coded how well they remembered the exhibits and content.

[D-head] Intent to Remember

Definition: Number of photos taken at each of the research exhibits. We considered photos to indicate visitors’ intent to remember the exhibit experience. Visitors may have additional reasons for taking the photos; however we felt this was a worthwhile proxy measure to explore.
Results for Intent to Remember were:

- **ADULTS with CHILDREN**: Participating adults with the participating children were similarly likely to take photos at the Immersives (20%) as at the Tabletops (12%). However, the effect size is considered small but notable, in favor of Immersives.

### [D-head] Recalled Research Exhibits

**Definition:** Whether or not the participants recalled all three research exhibits in the study 6 weeks after their museum visit. These were simple counts and did not require reliability coding.

Results for Recalled Exhibits were:

- **ADULTS with CHILDREN**: Participating adults with the participating children who experienced the Immersive exhibit group were similarly likely to recall all of the research exhibits (42%) as those who experienced the Tabletop exhibit group (52%). However, the size of the effect is considered small yet notable, in favor of the Tabletops.

### [D-head] Experience Recall

**Definition:** Number of words used to describe what participants recall having done at each of the exhibits, six weeks after using them. Word counts were used as an exploratory measure, as we were unable to create a detailed coding scheme and felt they
would provide a glimpse into visitors’ recall. Word counts were obtained from visitor responses in follow-up surveys and did not require reliability coding.

Results for Experience Recall were:

- ADULTS with CHILDREN: Participating adults with the participating children used a similar number of words to describe what they did at the Immersive exhibits (12.62) as at the Tabletop exhibits (9.86). However, the size of the effect is considered small but notable, in favor of the Immersives.

**Content Recall**

Definition: Number of words used to describe *what they thought each of the exhibits was about*, six weeks later. Word counts were used as an exploratory measure, as we were unable to create a detailed coding scheme and felt they would provide a glimpse into visitors’ recall. Word counts were obtained from visitor responses in follow-up surveys and did not require reliability coding.

Results for Content Recall were:

- ADULTS with CHILDREN: Participating adults with the participating children used a statistically similar number of words to describe what the exhibit was about at the Immersive exhibits (4.54) as at the Tabletop exhibits (10.29), when taking the variability into
account. However, the size of the effect is considered medium, in favor of the Tabletops.

**Table 1.** Results for All Study Variables

[Insert Table 1 ~ here.]

**[Head A] Discussion**

The study's results suggest that immersive and tabletop exhibits each have relative strengths with respect to the visitor experience. In terms of our hypotheses, we found:

- Immersives fostered more positive attitudes (affect) as expected, particularly in the area of using the exhibits with others (social interactions).
- Tabletops held learners' attention for longer periods of time, counter to our expectations.
- Tabletops were more intellectually engaging, counter to our expectations.
- There were few differences in memorability between the two exhibit types.

Below, we discuss each finding and its implications.

**[B-Head] Immersive Exhibits Fostered Slightly More Positive Affect.**
The National Research Council (2009) recently published a seminal report on Learning Science in Informal Environments. The NRC reviewed over 1,200 reports, articles and curricular standards to identify six strands of scientific learning that may occur in informal STEM learning environments. These strands apply to science in general and also to spatial reasoning in particular. The first strand focuses on the affective aspects of learning, highlighting the important function of science museums in enhancing learners’ excitement, interest, and motivation about a topic. Improved STEM affect can extend into future motivation and behavior by increasing interest in school-based topics, and encouraging youth to pursue careers in the fields of science, mathematics, or engineering (Meredith, Fortner and Mullins 1997; Salmi 2002, 2001).

Most of the engagement and affective variables studied suggest that the immersives had a greater affective appeal than the tabletops. (One result—time spent at the exhibits—ran contrary to this conclusion, which we will address below.) For example, immersive exhibits elicited more comments about positive experiences by adults and children than the tabletop exhibits (although to varying degrees).

While the differences between immersives and tabletops in our two measures of enjoyment did not reach the level of statistical significance, the effect sizes for both the enjoyment rating and the positive exhibit comments were small but notable in the direction of the Immersives. Moreover, adults were slightly more likely, and children significantly more likely, to offer negative comments about their engagement (i.e., boring, difficult) at the Tabletops. Taken together, these results support a trend toward greater playful engagement (e.g., enjoyable, fun, creative) at immersive exhibits worth exploring in future research.
We found further evidence that immersives may be seen as playfully engaging to adults in data from an unrelated study, the Exploratorium’s Exit Survey (Alexander, Garcia-Luis, Gardella and Wright 2011). To understand visitors’ overall Exploratorium experience, in-house researchers surveyed a random sample of over 1,900 adults as they exited the museum throughout 2009. Using those data in the current study, we coded visitors’ self-reported “favorite” exhibit as immersive or tabletop, and compared the frequencies of those responses to the number of immersive and tabletop exhibits typically on the museum floor. Adults exiting the museum at the end of their visit mentioned a significantly higher proportion of immersive exhibits as their favorite (32%), than the proportion of immersive exhibits available on the museum floor (4%). This supports the notion that immersive exhibits may be more enjoyable to adults.

Participants were keen on the social interactions at immersive exhibits as well. Adults and children experiencing the Immersive exhibits were significantly more likely to make positive statements about the social aspects of their experience (e.g., “We got to do it together”) than participants who used the Tabletop exhibits. There were no differences in adults’ negative comments about the social experience at the Immersives and Tabletops. However, children who used the Immersive exhibits did mention negative social aspects of their experience (i.e., crowding) significantly more than children at the Tabletop exhibits. This could be a function of having to share an immersive exhibit with other visitors in a way that is less common at tabletop exhibits, and suggests that immersive designs come with frustrations for children as well.

With the exception of holding time (discussed below), the majority of the engagement and affective results tend to support or trend in the direction of supporting
museum professionals’ intuitions about the strengths of immersives (Gilbert, 2000, 2002). However, our results suggest that immersives are less effective at other key aspects of the informal learning experience, namely holding learners' attention and fostering content-related conversations and reasoning. We turn next to these important learning behaviors.

Tabletop Exhibits Held Participants' Attention Longer.

Contrary to our expectations, adults and children at the Tabletop exhibits spent significantly more time there. In fact, the Tabletop users spent over a minute longer per exhibit than the Immersive users. This is particularly notable given that the typical time spent at a single exhibit is between one and three minutes (Humphrey and Gutwill 2005; Serrell 1998). According to the NRC, time spent at a museum exhibit may be viewed either as a prerequisite for learning or as a measure of learning itself (2009). From this perspective, the large results and effect sizes for the difference in holding time at the two types of exhibits indicates that tabletops are more effective than immersives at engaging learners.

Some ambiguity remains, however. While the Tabletop participants spent more time at the exhibits, adults at the Immersive exhibits were significantly more likely to leave and return to an exhibit to use it again. For children, both sets of exhibits were equally alluring, with a substantial number of children returning to all of the research exhibits. Perhaps adults returned to the Immersives to have a more playful and social experience, given the affect results discussed above. A more prosaic possibility is that our coding scheme, which considered simply walking over or through an immersive exhibit to be a "return," may have inadvertently over-counted adults' exiting paths as returns.
Tabletop exhibits were more intellectually engaging than immersive exhibits.

Adults and children made more spatial reasoning utterances (SRs) when experiencing the Tabletop exhibits than those experiencing the Immersive exhibits, with effect sizes varying from small to medium. Adults’ Low-level SRs—those about structure—did not differ significantly between the two types of exhibits in number or time spent uttering them. However, the small yet notable effect sizes in both measures of Low SRs suggest a trend toward more low-level cognitive engagement at the Tabletops. Results for the High-level SRs, those about cause and effect, were in the same direction and did reach significance. Results for children’s spatial reasoning were similar, indeed showing a slightly larger impact. Children made significantly more Low SRs about spatial structure at the Tabletops, and spent significantly greater time uttering them. Children rarely made High SR utterances, but children at the Tabletops were more likely to make at least one High-level SR than children at the Immersives. For both adults and children, the number of SRs was highly correlated to the holding time. From the current experiment, we cannot determine whether the longer holding times at tabletops led to more cognitive engagement or the cognitive tasks at tabletops encouraged more focused time.

Regardless, the spatial reasoning results for adults and children disconfirm our hypothesis and the beliefs held by many museum professionals (Gilbert, 2000, 2002) that immersive exhibits would foster greater intellectual involvement. The spatial reasoning results also contradict findings in other educational contexts that gross motor exploration
(similar to that found in many immersive exhibits) may build spatial reasoning skills (e.g., Acredolo 1977; Hazen 1983). Perhaps mathematical reasoning language is not a sufficient measure of the kind of kinesthetic/spatial learning that immersives may afford.

These results support other research findings that hands-on interactivity can bolster spatial thinking skills. For example, educational research has shown that physical and spatial explorations of manipulables help build spatial reasoning skills (National Council of Teachers of Mathematics 2000; Casey, Nuttal and Pezaris 2005). This is important because such skills increase performance in more abstract mathematics and the other sciences (Dennis 1995; Pallrand and Seeber 1984). Further, practicing and strengthening scientific reasoning skills is critical to learning STEM in both museum and school contexts (Ansbacher 1999; Ash 1999; Adey and Shayer 1993; White and Frederiksen 2000).

Although the spatial reasoning results were unambiguous, our other measures of Geometry-related Thinking revealed more nuanced differences between the Tabletops and the Immersives. For instance, participants who experienced the Immersive exhibits were significantly more likely to talk about themselves in relation to the exhibits, suggesting that they saw themselves and their bodies as integrated parts of the exhibit phenomenon. Nemirovsky et al. (1998) refer to this type of symbolic extension of the self as "fusion," and suggest that it is important to experiencing and developing embodied learning. This may indicate a recognition of the kinesthetic learning opportunities that immersive exhibits afford, or reflect visitors’ feelings of being immersed, which other researchers have suggested is important to the visitor experience of immersives (Bitgood, et al. 1990; Jones and Wageman 2000). Still, our finding that immersives prompted
learners to talk about themselves within the exhibit does not necessarily mean that they were more intellectually engaged at the Immersives.

**[B-head] There were few differences in memorability.**

Finally, we explored the widely held belief that immersive exhibits are more iconic and memorable than tabletop exhibits (Gilbert, 2000, 2002). As most of our memorability measures were exploratory, they do not provide conclusive results, but do help us identify areas for future research. Counter to our hypothesis, the results did not reveal significantly greater memorability of immersive exhibits. Visitors did not take more photos—our measure of intent to remember—in either condition. From the follow-up survey mailed out 6 weeks after participants' museum visit, we found that participants in both conditions were similarly likely to remember all three exhibits. However, the effect sizes support the affective and cognitive findings above. Specifically, there was a trend that learners used more words to describe their experience at the exhibits in the Immersive condition, exhibits which provoked more positive affect. In contrast, learners used more words to describe the content at the exhibits in the Tabletop condition, where learners spent more time engaging in spatial reasoning. These trends are likely due to the exploratory nature of the measures and the smaller sample size given the response rate, and suggest that future study is warranted with a fully developed coding scheme.

**[A-head] Limitations**

There are several limitations to consider when interpreting the results of this study. First, this study focused on a single type of immersive exhibit—the Whole-body
Interactive. These results may not hold for other types of immersives, such as Recreated Immersive exhibits, or even Whole-body Non-interactive exhibits. Further, this study took place in a single museum setting. Perhaps there is something about this particular museum—with its large number of tabletop exhibits—that attracts visitors who prefer tabletop experiences, or that allows its visitors to practice and become familiar with the tabletop experience. While some evidence suggests that general visitors to the Exploratorium preferred a larger proportion of available immersive exhibits than tabletops (Alexander, et al. 2011), it will be important to replicate these findings in other museums. It is our hope that this study will contribute to a larger body of work, focusing on each type of immersive exhibit in different contexts.

Another limitation to this study is that we were unable to assess kinesthetic learning, which is potentially a content-related advantage of immersive exhibits. We attempted several coding schemes, bringing in experts in embodied learning, spatial cognition, and Laban Movement Analysis. We found it incredibly difficult to create a reliable coding scheme that focused on kinesthetic learning, and in the end had to abandon these efforts. However, we were able to get a glimpse of embodied learning by looking to visitors’ self-referencing talk, or verbalizations about themselves and their bodies as part of the exhibit experience, which were more common to the immersive experience. It is our hope that the field will develop other methods for accurately assessing kinesthetic learning in this kind of research.

Similarly, we had a difficult time creating a scheme for our memorability measures, in part due to the low response rate. We turned to word counts to get a general
sense of the depth of responses. Conclusions cannot be drawn from these exploratory measures; however, we feel that they do indicate that future research is deserved.

Finally, the between-subjects nature of this research design presented visitor groups with either immersive or tabletop exhibits. Therefore, we were not able to answer questions about differences in the visitor experience when faced with both immersive and tabletop exhibits. The next step in our research agenda will be a within subjects design to understand how the immersive and tabletop visitor experience differs when both types of exhibits are present.

[A-head] Conclusion

Despite this study’s limitations, the greater intellectual engagement found at tabletop exhibits complements Chabon’s (2013) view that a more direct and intuitive grasp of a topic comes through hands-on manipulation. However, immersives appear to have an important positive effect on visitors’ socio-emotional experience. Perhaps practitioners’ beliefs about the benefits of immersive exhibits stem from this impact. Indeed, some in the museum field argue that our educational power lies heavily in the affective domains of inspiration, enjoyment and social interaction (Anderson and Roe 1993; Dewitt and Storksdieck 2008; Paris, Yambor and Wai-Ling Pacard 1998; Russell 1990). Our hope is that this research helps inform museum practitioners as they determine which type of exhibit merits their time, funding, and space to create.

Notes:
1. For the purposes of this study, Personal Space was loaned to the Exploratorium by Snibbe Interactive.

2. Inter-rater agreement for codes of interview responses was above 90%, which is considered high (Multon 2010). The two coders discussed all disagreements and determined the final codes. Coding of time spent at exhibits (as captured on video) led to 90% agreement about exact start and stop times on the 20% of the data that were double-coded, which is also considered high (Multon 2010). Coding of participants’ talk captured on video resulted in Cohen’s Kappa statistics—a stringent measure of interrater reliability that controls for chance agreement—of .76 for adults and .72 for children (on the 25% of the data that were double-coded). These statistics indicate good-to-excellent levels of agreement (Fleiss, et al. 2004). The video coders discussed disagreements until final codes were agreed upon. To ensure that the remaining 75% of the single-coded video data were of a similarly high quality to the double-coded data, coders were allowed to flag questionable utterances and discuss them together and with the research coordinator.

3. Comparisons were conducted separately for adults and children when appropriate. T-tests determined differences in continuous variables. Chi square tests of independence determined differences in categorical and dichotomized variables (with the exception of a few analyses that required Fisher’s Exact tests due to an extremely small cell size). Whenever the variable distribution was non-normal, data were Winsorized. When Winsorization did not produce normal variation, the sensitivity of the parametric results was checked using non-parametric analyses. Further, when Levine’s Test revealed the
variances to be significantly unequal, the degrees of freedom were adjusted to yield more
accurate test results. All tests employed a two-tailed significance level of .05.

4. Effect size d (used when comparing averages) is considered small but notable at .2,
medium at .5, and large at .8 or greater (Cohen 1992). Effect size phi (used when
comparing proportions) is considered small but notable at .1, medium at .3, and large at .5
(Green and Salkind 2004).

5. $\chi^2(1) = 127.07$, $p < .001$.

6. The correlation between Low-level SRs and Holding Time was .74 for adults and .64
for children. For High-level SRs and Holding Time, the correlation was .48 for adults
and .31 for children (recall children rarely made High-level SRs, so the latter correlation
may not be an accurate reflection of that relationship).

**Acknowledgments**

The authors would like to thank the incredible group of contributors who made this
research, and its dissemination, possible. The creative and supportive Geometry
Playground team, who created immersive and tabletop exhibit pairs, helped us lug the
exhibits on and off the floor throughout the study, came in on their days off to
troubleshoot and maintain the exhibits in the study, and so much more: Kevin Boyd, Eric
Dimond, Nina Hido, Kristina Larsen,
Sebastian Martin, Peggy Monahan, Joha Pedersen, Tom Rockwell, Susan Sandlin, Eric
Socolofsky, Paul Stepahin, and Mary Elizabeth Yarbrough. The diligent data collection,
data management, and coding team, without whom there would be no study: Adam
Klinger, Debbie Siegel, Sarah Elovich, Tamas Makany, Courtney Brown, Karen
Campbell, Melissa Hempel, Mario Landau-Holdsworth, and Brian Simpson. Our knowledgeable and encouraging advisory council: Maureen Callanan, Doug Clements, Peggy Hackney, and Ricardo Nemirovsky. The Exploratorium community for allowing us to create, and helping us maintain, a controlled environment on the museum floor. And our talented photographers who make it so much easier to share our work: Amy Snyder and Shannon Paras. Finally we’d like to share our appreciation for the charitable contribution by Scott Snibbe & his team at Snibbe Interactive, who not only loaned us an immersive piece to include in the research, but also spent hours setting it up and teaching us how to maintain it.

We are grateful for the generous financial support of the National Science Foundation.

This material is based upon work supported by the National Science Foundation under Grant number 0610436. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References


———. 1989 "The Role of Visualization in the Middle School Mathematics Curriculum ". Focus on Learning Problems in Mathematics 11), 49-60.


Casasola, M., J. Bhagwat, and A. S. Burke. 2009 "Learning to Form a Spatial Category of Tight-Fit Relations: How Experience with a Label Can Give a Boost." Developmental Psychology 45), 711-723.


### Table 1. Results for All Study Variables

<table>
<thead>
<tr>
<th>Topic</th>
<th>Outcome variable</th>
<th>People/Source</th>
<th>Immersive mean (SD), or %</th>
<th>Tabletop mean (SD), or %</th>
<th>Test statistic and result</th>
<th>Effect size*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical Engagement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holding Time per Exhibit (mins)</td>
<td>Adult</td>
<td>2.89 (1.01)</td>
<td>4.05 (1.52)</td>
<td>t(103) = 4.92, p &lt; .001</td>
<td>d = .91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>2.89 (.98)</td>
<td>4.06 (1.49)</td>
<td>t(102) = 5.08, p &lt; .001</td>
<td>d = .94</td>
<td></td>
</tr>
<tr>
<td>Exhibit Returns</td>
<td>Adult</td>
<td>40%</td>
<td>20%</td>
<td>$\chi^2(1) = 5.71, p = .02$</td>
<td>phi = .22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>50%</td>
<td>42%</td>
<td>$\chi^2(1) = .84, p = .36$</td>
<td>phi = .08</td>
<td></td>
</tr>
<tr>
<td><strong>Affect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enjoy Using</td>
<td>Adults</td>
<td>3.83 (.53)</td>
<td>3.70 (.66)</td>
<td>t(118) = 1.17, p = .25</td>
<td>d = .22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>4.09 (.50)</td>
<td>3.98 (.58)</td>
<td>t(118) = 1.13, p = .26</td>
<td>d = .21</td>
<td></td>
</tr>
<tr>
<td>Exhibit Attitudes—Pos</td>
<td>Adults</td>
<td>45%</td>
<td>33%</td>
<td>$\chi^2(1) = 1.71, p = .19$</td>
<td>phi = .12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>43%</td>
<td>32%</td>
<td>$\chi^2(1) = 1.74, p = .19$</td>
<td>phi = .12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>68%</td>
<td>80%</td>
<td>$\chi^2(1) = 2.13, p = .14$</td>
<td>phi = .13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>50%</td>
<td>70%</td>
<td>$\chi^2(1) = 5.00, p = .03$</td>
<td>phi = .20</td>
<td></td>
</tr>
<tr>
<td>Social Attitudes—Pos</td>
<td>Adults</td>
<td>60%</td>
<td>28%</td>
<td>$\chi^2(1) = 12.20, p &lt; .001$</td>
<td>phi = .32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>22%</td>
<td>5%</td>
<td>$\chi^2(1) = 7.21, p = .007$</td>
<td>phi = .25</td>
<td></td>
</tr>
<tr>
<td>Social Attitudes—Neg</td>
<td>Adults</td>
<td>23%</td>
<td>20%</td>
<td>$\chi^2(1) = .20, p = .66$</td>
<td>phi = .04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>18%</td>
<td>3%</td>
<td>$\chi^2(1) = 6.99, p = .008$</td>
<td>phi = .24</td>
<td></td>
</tr>
<tr>
<td>Self-referencing Talk</td>
<td>Adults</td>
<td>68%</td>
<td>0%</td>
<td>Fishers exact = p &lt; .001</td>
<td>Phi = .72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>90%</td>
<td>2%</td>
<td>Fishers exact = p &lt; .001</td>
<td>Phi = .87</td>
<td></td>
</tr>
<tr>
<td>Low SRs per exh—Number</td>
<td>Adults</td>
<td>8.97 (3.31)</td>
<td>9.95 (4.93)</td>
<td>t(103) = 1.28, p = .20</td>
<td>d = .24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>5.61 (2.50)</td>
<td>7.11 (3.22)</td>
<td>t(118) = 2.84, p = .005</td>
<td>d = .52</td>
<td></td>
</tr>
<tr>
<td>Low SRs per exh—Length</td>
<td>Adults</td>
<td>32.92 (16.89)</td>
<td>38.90 (25.76)</td>
<td>t(102) = 1.50, p = .14</td>
<td>d = .28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>17.05 (10.10)</td>
<td>21.78 (14.83)</td>
<td>t(104) = 2.04, p = .04</td>
<td>d = .38</td>
<td></td>
</tr>
<tr>
<td>High SRs per exh—Number</td>
<td>Adults</td>
<td>.88 (.66)</td>
<td>1.35 (.98)</td>
<td>t(103) = 3.02, p = .003</td>
<td>d = .56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>8.26 (8.29)</td>
<td>11.61 (9.89)</td>
<td>t(118) = 2.01, p = .047</td>
<td>d = .37</td>
<td></td>
</tr>
<tr>
<td>High SRs—Any Uttered</td>
<td>Adults</td>
<td>23%</td>
<td>55%</td>
<td>$\chi^2(1) = 12.63, p &lt; .001$</td>
<td>phi = .32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children</td>
<td>23%</td>
<td>55%</td>
<td>$\chi^2(1) = 12.63, p &lt; .001$</td>
<td>phi = .32</td>
<td></td>
</tr>
<tr>
<td>Intent to Remember</td>
<td>Video</td>
<td>20%</td>
<td>12%</td>
<td>$\chi^2(1) = 1.56, p = .21$</td>
<td>phi = .11</td>
<td></td>
</tr>
<tr>
<td>Recalled Exhibits</td>
<td>Follow-up</td>
<td>42%</td>
<td>52%</td>
<td>$\chi^2(1) = .47, p = .49$</td>
<td>phi = .10</td>
<td></td>
</tr>
<tr>
<td>Experience Recall</td>
<td>Follow-up</td>
<td>12.62 (10.85)</td>
<td>9.86 (7.90)</td>
<td>t(45) = -.97, p = .34</td>
<td>d = .29</td>
<td></td>
</tr>
<tr>
<td>Content Recall</td>
<td>Follow-up</td>
<td>4.54 (5.02)</td>
<td>10.29 (13.87)</td>
<td>t(24) = 1.81, p = .08</td>
<td>d = .59</td>
<td></td>
</tr>
</tbody>
</table>

*For effect size d, small=.2, med=.5, large=.8; for effects phi, small=.1, med=.3, large=.5
Figure 1:

How do they differ?

Figure 2:

Figure 3:

Immersive: Personal Space.  Tabletop: Dividing Space.
Figure 4:

**Immersive:** Spin a Shape Around You.  
**Tabletop:** Spin a Shape.

![Immersive: Spin a Shape Around You.](image1) ![Tabletop: Spin a Shape.](image2)

Figure 5:

**Immersive:** Distorted Chair.  
**Tabletop:** Distorted Drawing.

![Immersive: Distorted Chair.](image3) ![Tabletop: Distorted Drawing.](image4)

Figure 6:

**Immersives Condition**

![Immersion](image5) $n=60$

**Tabletops Condition**

![Tabletops](image6) $n=60$