THE IMPACT OF FACILITATION BY MUSEUM EDUCATORS ON FAMILY LEARNING AT INTERACTIVE MATH EXHIBITS: A QUASI-EXPERIMENTAL STUDY

Scott A. Pattison,[1] Andee Rubin,[1] Marcie Benne,[2] Ivel Gontan,[3] Elizabeth Andanen,[2] Todd Shagott,[2] Melanie Francisco,[4] Smirla Ramos-Montañez,[2] Crosby Bromley,[5] and Lynn D. Dierking[6]

[1] TERC, Boston, MA, USA

[2] Oregon Museum of Science and Industry (OMSI), USA

[3] Fleet Science Center, San Diego, California, USA

[4] Kaiser Permanente Center for Health Research Northwest

[5] Salem-Keizer School District, Salem, Oregon, USA

[6] Oregon State University, Corvallis, Oregon, USA

This is the pre-publication version of the following article: Pattison, S. A., Rubin, A., Benne, M., Gontan, I., Shagott, T., Francisco, M., Ramos-Montañez, S, Bromley, C., & Dierking, L. D. (2018). The impact of facilitation by museum educators on family learning at interactive math exhibits: A quasi- experimental study. *Visitor Studies, 21*(1), 4–30.

https://doi.org/10.1080/10645578.2018.1503879

Abstract

Although discussions of museums often revolve around exhibits, educators in these spaces have the potential to create in-depth, social learning experiences beyond what is possible at exhibits alone. There is still little empirical research, however, to inform how we understand, approach, and improve museum facilitation practices. In this study, we sought to address this gap by quantifying the impact of facilitation by trained educators working with visitors at interactive museum exhibits and comparing this to visitor engagement and learning outcomes for families without educator support. Using a quasi-experimental design, we measured the impact of staff facilitation on visitor engagement time, mathematical reasoning, math awareness, satisfaction, and intergenerational communication across three different exhibits, four trained educators, and two experimental conditions. Multivariate regression modeling showed that staff facilitation had a positive impact on engagement time, mathematical reasoning, and satisfaction, a negative impact on intergenerational communication, and no impact on math awareness.

Keywords: museum education; staff facilitation; interactive exhibits; family learning; quasi-experimental design; mathematics

THE IMPACT OF FACILITATION BY MUSEUM EDUCATORS ON FAMILY LEARNING AT INTERACTIVE MATH EXHIBITS: A QUASI-EXPERIMENTAL STUDY

Although discussions of museums often revolve around exhibits, educators in these informal learning institutions have the potential to create in-depth, social learning experiences beyond what is possible at exhibits alone. With varying roles across different organizations, these educators lead school group programs and classes, enhance visitor experiences with demonstrations and tabletop activities, and facilitate visitor learning at exhibits throughout the museum. Through these experiences, they find ways to connect with visitors on a personal level and serve as teachers, learning guides, and inspirational role models (Astor-Jack, Whaley, Dierking, Perry, & Garibay, 2007; National Research Council, 2015; Pattison & Dierking, 2012; Tran, 2007). Although the role of museum educator require many of the same skills as a classroom teacher, they also must navigate the complex and unique dynamics of informal/free-choice learning (Pattison & Dierking, 2013). For example, in addition to pursuing the educational goals of their program and institution, museum educators must also prioritize participants' enjoyment, choice and control, social interactions, and life-long learning relevance and interests (Bevan & Xanthoudaki, 2008; National Research Council, 2009, 2015).

Over the past several decades, there has been a growing number of projects, professional development resources, and funding opportunities focused on supporting the work and professional learning of museum educators (Patrick, 2017), including several nationally recognized professional development programs.[1] Despite this increased attention, however, there is little empirical research to inform how we understand, approach, and improve museum facilitation practices with regard to visitor engagement and learning (Bevan & Xanthoudaki, 2008). As an example, in the seminal synthesis report on learning science in informal environments, which focused heavily on science centers and museums, the National Research Council (NRC) (2009) did not cite a single study specifically on the work of museum educators. Six years later, in the more recent report on quality STEM programs in out-of-school settings, the NRC (2015) referenced some studies of effective professional development (e.g., National Research Council & Institute of Medicine, 2002; Peter, 2007) and synthesized general characteristics of effective programs but again made no mention of research specifically on the practices and impacts of educators in these settings. The little research that has been done over the last several decades has focused primarily on school groups (e.g., Cox-Petersen, Marsh, Kisiel, & Melber, 2003; Jarvis & Pell, 2005; Tal & Morag, 2007), which represent only a small fraction of the varied ways that educators in museums engage adults and children (see Pattison, Randol, et al., 2017, and Pattison & Dierking, 2013, for reviews). A few studies have also looked at the professional learning of educators (e.g., Ash & Lombana, 2013; Tran, 2007), but have not empirically connected this work to the impact of museum staff on visitor learning. Therefore, it is not clear what educational strategies and approaches should be the focus of educator trainings, and managers and policy-makers have little empirical evidence to guide their support for educators.

In this study, we sought to address these gaps by quantifying the impact of facilitation by trained educators working with visitors at interactive museum exhibits and comparing this to visitor engagement and learning outcomes for families without educator support. We focused on unscripted conversations between educators and visitors that often occur in exhibit spaces or at informal activity stations. In our work, we refer to these as *unstructured* interactions, to distinguish them from the more structured, pre-planned experiences that museum educators might lead, such as a school group program or stage demonstration (Pattison, Randol, et al., 2017; Pattison & Dierking, 2012). During unstructured interactions, the educator is often entering an existing social interaction among visitors and the structure and directions of the interaction are co-negotiated between the visitor and the educator (Cunningham, 2004; Pattison & Dierking, 2013). We believe these interactions are common and central to the practice of museum educators and represent a type of facilitated interaction unique to informal/free-choice learning environments—when an educator has the opportunity to support learning but still preserve the choice and control held by participants over the learning experience (Bevan & Xanthoudaki, 2008; National Research Council, 2015).

The study was part of the three-year, National Science Foundation (NSF)-funded *Researching the Value of Educator Actions on Learning* (REVEAL) project, which was designed to develop and test effective staff facilitation strategies for family learning at interactive math exhibits in a science center.[2] Initially, we conducted a design-based research (DBR) study to identify promising staff facilitation strategies and develop a theoretical model of staff-facilitated family learning at interactive exhibits (Pattison, Randol, et al., 2017). In this article, we report on the subsequent quasi-experimental study in which we tested the impact of this facilitation approach on visitor learning and engagement with four educators at three different exhibits and compared outcomes between facilitated and unfacilitated interactions. Before describing the methods and results of the study, we lay out the REVEAL facilitation model that was developed during the initial phase of the project and served as the guiding framework for the facilitation approach used in this study.

REVEAL Facilitation Approach

The REVEAL facilitation approach was developed through a DBR study with two expert museum educators, hundreds of museum visitor groups, and extensive data collection and analysis using observations, video and audio recording, and educator debriefs and reflections (Pattison, Randol, et al., 2017). The study used three interactive exhibits designed to engage visitors in algebraic thinking—a type of mathematical reasoning, similar to scientific inquiry, involving the exploration of mathematical relationships in the world around us and the use of these relationships to understand and create (Greenes & Rubenstein, 2008; Kaput, Carraher, & Blanton, 2008; Moses, 1999). These three exhibits had been developed and tested to not only engage visitor groups alone, but also to support interactions between visitors group and educators (Pattison, 2011). The exhibits, therefore, represented a best-case scenario for understanding the value an educator might contribute to visitor learning at an exhibit.

The DBR study focused on families, given that they represent a critical audience for museums (Dierking & Falk, 1994; Ellenbogen, Luke, & Dierking, 2007; National Research Council, 2009) and an especially challenging group for museum educators (Pattison & Dierking, 2013). As research has repeatedly shown, families come to museums with their own patterns of learning together and often engage successfully with museum exhibits without staff support (e.g., Ash, 2002; Crowley & Jacobs, 2002; Ellenbogen, 2002; Garibay Group, 2013; Rigney & Callanan, 2011). Museum educators, therefore, must navigate existing social dynamics and goals within the family while finding creative ways to support learning. Aligned with this perspective, the DBR study and our subsequent work has been founded on an asset-based perspective on education and learning (e.g., Garibay, Yalowitz, & Guest Editors, 2015; González, Moll, & Amanti, 2005; Gutiérrez & Calabrese Barton, 2015; Gutierrez & Rogoff, 2003), culturally responsive approaches to education and research (e.g., Brown & Crippen, 2017; Friedman, 2008; Kirkhart, 1995), and a view of museum education as a complex form of social interaction rather than simply a teaching experience (Bevan & Xanthoudaki, 2008; Pattison & Dierking, 2013).

Through iterative refinement and testing of facilitation strategies with the two expert educators, we developed the REVEAL model of facilitation within the context of family learning at interactive math exhibits (Pattison, Randol, et al., 2017). As shown in Figure 1, the model identifies the three goals that guided the approach, the reflective cycle of facilitation that educators used to pursue those goals and respond to the unique needs and interests of family groups, and the various physical, personal, and social factors that influenced the nature of the interactions and the strategies and techniques used by the educators with each family. Based on our view that museum facilitation is a complex social interaction between educators and visitors that requires acknowledging and building on the interests and assets of visitor groups, the three learning goals identified in the model guided educators in a continuous process of balancing the more content-focused goal of the interactions (mathematical reasoning) with the more visitor-centered goals (satisfaction and intergenerational communication) that are essential to informal/free-choice learning (Falk & Dierking, 2013; National Research Council, 2009, 2015). As the model indicates, to achieve this balance the educators followed an ongoing cycle of *responsive facilitation*, observing visitors and their interactions with the exhibit, making choices about their facilitation strategies, and then reflecting on the impact their facilitation had on the families.

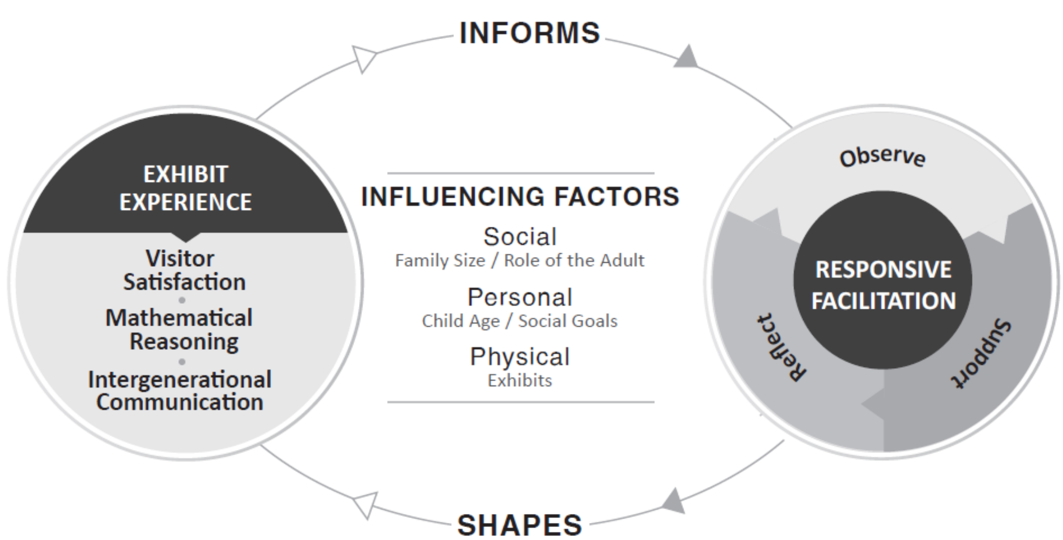

Figure 1. REVEAL responsive facilitation model for families.

In the DBR study, we identified a number of indicators that the expert educators used when observing families and assessing their needs and interests, as well as a core set of facilitation strategies that the educators frequently drew from when supporting family learning and engagement. Cues that facilitators looked for when tailoring their strategies and approaches to each family included (a) how well families oriented to the exhibits and were able to figure out the basic goals and functions of the interactives, (b) whether or not families found ways to go beyond basic use and explore the exhibits more deeply, (c) the level and type of mathematical reasoning being used by family members, (d) the extent to which families took control and ownership over the experiences based on their own goals and interests, and (e) how well adults and children were interacting with each other at the exhibits. Depending on what facilitators noticed for each family, they used a variety of strategies to support the intersection of educator and family goals, such as orienting visitors to the exhibits, proposing different types and levels of challenges, providing just-in-time explanations (Crowley et al., 2001; Palmquist & Crowley, 2007), showing appreciation, and helping visitors establish ownership over their experiences. For example, when a family approached one of the exhibits and seemed unsure of how to begin, the facilitator might offer some basic orientation and pose an initial, entry-level challenge. In contrast, if a family immediately began exploring and experimenting with one of the exhibits on their own, the educator might instead hang back and then offer an extension challenge when she saw that the family was looking for different ways to explore the exhibit.[3]

Research Question

In the present study, we trained four new educators in the REVEAL facilitation approach and tested the impact of their facilitation through a quasi-experimental research design. The research was guided by the following question: What is the impact of staff facilitation on family learning at interactive exhibits, compared to a “greeting only” condition? Although not the focus of this article, we also investigated associations between the measured impact of staff facilitation and the degree to which facilitators used a responsive facilitation approach, matching their strategies to the specific characteristics, needs, and interests of families. In addition, we explored how well the theoretical model developed during the previous phase of the REVEAL project matched the quantitative patterns observed in the quasi-experimental study. At the end of this article, we briefly discuss findings and challenges associated with these secondary study goals.

Based on the REVEAL facilitation model, we measured five indicators of family learning and engagement: mathematical reasoning, visitor satisfaction, intergenerational communication, math awareness, and positive math affect. These were intended to represent important learning outcomes from both the perspectives of the educators and research team (e.g., mathematical reasoning) and of families (e.g., satisfaction). Similarly, the range of math-related measures were intended to provide a more complete picture of how families engaged with the mathematical content. As control variables, we also considered five factors hypothesized to influence the nature and outcomes of the interactions: exhibit, age of child, visitor social goals, group size, and adult visitor roles. These were intended to represent the physical (exhibit), social (group size, adult visitor roles), and personal (visitor social goals, age of child) factors that emerged as important during the DBR study (Figure 1), aligned with Falk and Dierking's contextual model of free-choice learning (Falk & Dierking, 2000, 2013).

Method

This study used a quasi-experimental design (Shadish, Cook, & Campbell, 2001) with a carefully chosen control group to test the impact of staff facilitation on family learning at exhibits. To broaden the generalizability and external validity of the results, the impact of facilitation was tested with four trained educators and three different exhibits. Throughout the study, we adopted a culturally responsive research (CRR) approach (e.g., Allen et al., 2007;

Frechtling, 2010; Hacker, 2013; Kirkhart, 1995; Okazaki & Sue, 1995). Building on Kirkhart's notion of cultural validity (Kirkhart, 1995; Kirkhart & Hopson, 2010), we developed a CRR framework specific to the project and identified strategies for data collection, analysis, interpretation, and dissemination, such as reflecting on cultural assumptions underlying study measures, assembling a research team of bilingual/bicultural researchers who could analyze and interpret data in the original language of participants, and working with two external CRR "coaches" to guide the team throughout the study and help ensure accountability to our CRR framework.[4]

Data Collection Context

The study was conducted at the Oregon Museum of Science and Industry, a large science center in Portland, Oregon, USA. Three exhibits, developed through the NSF-funded Design Zone project,[5] were included: Balancing Art, Drawing in Motion, and Designing for Speed. As described in detail in Pattison, Randol, et al. (2017), all three are intended to engage visitors in using functional reasoning and algebraic thinking in creative, design contexts. The exhibits used in the study were developed with specific "facilitation affordances" (Pattison & Dierking, 2012), such as extension challenges, to give educators opportunities to deepen and extend visitor learning beyond what was possible when the exhibits were unfacilitated.

In brief, Balancing Art challenges visitors to create balanced sculptures by hanging weights on each side of a rod suspended on a central fulcrum. Distances from the center of the rod and the weight of each sculpture piece are quantified so that visitors can explore, discover, and use both qualitative and quantitative understandings of the mathematical relationship between distance, weight, and balance. The Drawing in Motion exhibit is, in essence, a giant Etch A Sketch, in which visitors collaborate to draw images on a screen, with one visitor using a slider to control horizontal motion and a second visitor using a different slider to control vertical motion. The activity includes a number of built-in challenges to introduce visitors to the relation between the relative motion of each slider and the shape and slope of the line on the screen. Finally, the Designing for Speed activity is a mathematical version of the classic science center exhibit about rotational motion. Two tracks allow visitors to roll wheels down an inclined ramp and test how the distribution of the weight from the center of the wheels influences how quickly the wheels accelerate and reach the bottom of the ramp. The wheels are labeled and the time the wheel takes to reach the bottom is automatically recorded so that visitors can discover the qualitative relationship between weight distribution and acceleration (e.g., the farther the weight is from the center of the wheel, the longer it takes for the wheel to accelerate).

Participants

Study participants included both educators and visiting family groups. During data collection, all visitor groups were allowed to interact with the exhibits but only intergenerational family groups were included in the final analyses. Families were defined broadly to include any group of visitors that came to the museum together and included at least one adult (18 years or older) and one child (under the age of 18). Age and family relationships (i.e., who was included in the family group) were estimated by observers and confirmed using survey responses. During the six months of data collection, 369 eligible families were videotaped and completed the post-interaction survey, evenly distributed between the two experimental conditions. Based on a power analysis,[6] we analyzed a subset of families within each combination of experimental factors (educator, exhibit, day of week, time of day), randomly selecting from the participant groups when a greater number of data points were available for a specific combination. This strategy resulted in a total of 263 family groups included in the final analyses (171 in the facilitation condition, 92 in the control condition).

The final sample of families represented a broad range of OMSI visitors, based on gender, age, group size, number of children, education level, ethnicity, race, languages spoken at home, and visits within the last 12 months to the museum. Demographics and visitor group characteristics were reported by the individual that completed the post-interaction survey. Adult respondents were evenly split between males and females and were 41 years old on average. The majority (67%) had at least a bachelor's degree and had visited the museum between three and four times on average in the last 12 months. Most respondents identified as white, not Hispanic/Latino (82%), although the sample included over 7% of participants identifying as Hispanic/Latino and almost 9% identifying as Asian. Nearly a quarter of participants (22%) reported speaking a language other than English at home. The typical group size was 3 to 4 visitors, including about two children or youth on average.

The team of museum educators that engaged with visitors during the study was made up of four part-time educators at the science center (three women and one man). These four were selected because they had all worked at the museum for at least two years and were all considered by managers and other full-time educators to be effective, skilled exhibit facilitators. Once they agreed to take part in the study, they participated in five half-day training sessions over two weeks, including practice facilitating visitor engagement at the three exhibits with guidance and coaching from members of the project team. The first session introduced the project and the REVEAL facilitation approach, the next three focused on each of the three exhibits and the mathematics and facilitation strategies specific to those exhibits, and the final session gave the group time to reflect on what they had learned and prepare for data collection. Each session included opportunities for the educators to watch and discuss example videos of expert educators facilitating family learning at the Design Zone exhibits, as well as to practice facilitating with visitors at these exhibits, supported by coaching from the project team. The training was aligned with the REVEAL responsive facilitation approach and was designed to build the educators’ skills using facilitation strategies identified through the DBR study. Although not as long as some professional development programs in the field, such as Reflecting on Practice,[7] the training was intended to be more intensive than what is typically provided to educators for a specific exhibit, while still of a scope and scale that might be realistically provided by a museum without grant funding.

An important and unique goal of the training sessions was to orientate educators to a different way of thinking about the exhibit content. Instead of communicating messages about math or algebra, the REVEAL facilitation strategies and approaches highlighted in the training were designed to deepen visitor engagement with algebraic thinking and the use of mathematical relationships to solve the creative challenges posed by the exhibits. For example, at the Balancing Art exhibit, facilitators did not focus on whether or not visitors understood the physics of angular momentum, as might be the case at a typical phenomenon-based activity, but instead practiced strategies for helping visitors explore the mathematical relationships between weight, distance, and balance. This included posing new challenges that either simplified the mathematical relationships for families who were just orienting themselves (e.g., using one weight on each side) or encouraged deeper exploration for other families (e.g., presenting a mystery weight or challenging families to create a non-symmetrical balanced configuration). Facilitators also took a broad and flexible approach to the mathematics in the exhibits. At Balancing Art, facilitators might urge some visitors to consider the qualitative relationships (e.g., "If we move this piece closer to the center, will the bar move up or down?"), while encouraging others to investigate more precise quantitative relationships (e.g., weight times distance equals force).

Similarly, the training provided educators with strategies for balancing the three REVEAL facilitation goals of supporting algebraic thinking, intergenerational communication, and visitor satisfaction. For example, if some members of a visitor group were engaging with the Balancing Art activity but others were hanging back, the educator might provide a role for the hesitant members, such as picking the location of the mystery weight to challenge the rest of the group. Or if some group members appeared not to be interested in the mathematical focus of the exhibit, the facilitator might emphasize aesthetic goals (e.g., creating a colorful mobile at the Balancing Art exhibit) or encourage intuitive mathematical strategies (e.g., systematic guess-and-check). Most importantly, the training prepared educators to use the responsive facilitation cycle as a framework for observing visitors and tailoring their facilitation approach to the unique goals and needs of each group.

Data Collection

During almost every Saturday and Sunday between January and June 2015, the research team collected video and audio data and post-interaction surveys from visitor groups engaging with one of the three Design Zone exhibits, either facilitated by an educator or not. Data collection was divided into morning and afternoon shifts each day. Each shift represented a unique combination of museum educator (four different staff members), exhibit (three different exhibits), experimental condition (greeting or facilitation), and time of day (morning or afternoon).

In the facilitation condition, the museum educators were instructed to facilitate all visitor groups at the exhibits as they felt appropriate, based on the REVEAL training they had received. In the greeting condition, the museum educators verbally greeted visitor groups within 30 seconds after they had engaged with the exhibits and then remained within the informed consent area without providing additional facilitation. The greeting condition was intended to serve as a control for potential sampling bias within the study, assuming some visitors might be more or less likely to approach an exhibit when a facilitator is present.

Data for the primary study variables were collected through videotaping of family interactions in both conditions and post-interaction surveys with one adult from each participating group. During data collection, a stationary video camera was set up to one side of the exhibit to record all staff-family interactions during that data collection shifts. A wireless microphone was attached to the exhibit itself to capture high-quality audio of the interactions. While the interactions were being videotaped, a researcher stood nearby and identified the primary adult in each eligible visitor group, defined as the adult who was most directly engaged with the exhibit and the child or children in the group, as estimated by the researcher based on the entire interaction. When this adult left the area, the researcher approached and recruited the individual for the post-interaction survey. If more than one adult was highly engaged, the data collector recruited the last highly engaged adult to leave the stanchioned informed consent area.[8]

Measures

The analyses reported in this article included six outcome variables (visitor satisfaction, math enjoyment, math awareness, mathematical reasoning, intergenerational communication, and engagement time), three experimental variables (facilitation condition, exhibit, and educator), and five control variables (number of prior visits in the last 12 months, adult respondent age, group size, average child age, and engagement time). Engagement time was analyzed as both an outcome and control variable based on our assumption that educators likely influenced the nature and outcomes of family interactions directly, through their use of facilitation strategies, as well as indirectly, by encouraging families to spend longer at the exhibits and thus have more time to deepen and extend their engagement with the activities. We also included adult respondent age and number of visits to the museum within the last 12 months as controls because both were found to be related to the outcome measures during initial exploratory data analysis.

Visitor social goals and adult visitor roles, which emerged as important influencing factors in the DBR study, were originally intended to be assessed through both observation and self-report measures of the degree to which educator facilitation strategies were perceived to respond to the needs and goals of each family—which we labeled *family-facilitator match*. However, the measures of family-facilitator match developed and piloted in the study suffered from issues of reliability and validity and were ultimately dropped from the analysis. (We discuss the challenges of measuring facilitator-family match more at the end of this article.)

All survey items were developed in Spanish and English and all video coding measures were designed to assess visitors speaking either language, following the project's CRR approach and OMSI's guidelines for handling data in multiple languages. Final versions of the Spanish and English survey instruments and video coding rubrics are available on the project website.

Visitor satisfaction. Adult survey respondents were asked to rate their level of satisfaction with the facilitated exhibit through a 5-item self-report index adapted from Packer (2004). Respondents rated their level of agreement to each item using a 7-point scale with five anchor categories (*strongly disagree* to *strongly agree*). For example, visitors were asked to rate their level of agreement with the statement: "I feel I benefited greatly from my experience with the exhibit today." Analysis of survey responses indicated a high level of internal consistency, or reliability among questions, across the five items (Cronbach's $\alpha = 0.95$, 95% CI = 0.90, 1.00, range = 1–7). For the final analyses, the responses were averaged across the five items for each participant to generate a general satisfaction score for each family group.

Math enjoyment. Level of math enjoyment for each family group was also measured using a set of five self-report items. Each adult respondent was asked to rate their level of enjoyment for each item on a 7-point scale with five anchor categories (*not at all enjoyable* to *very enjoyable*). For each item, respondents could also mark that "we did not do this," which was treated as a missing response. Because this set of items was intended to measure visitor level of enjoyment for the mathematical aspects of the activity, regardless of whether or not they perceived these aspects to be mathematical, the items described specific components or activities within the exhibit experience without using the terms *math* or *mathematics*. Also, because the items were specific to the activities, separate item sets were developed for each of the three Design Zone exhibits. For example, participants at the Balancing Art exhibit were asked to indicate how much they enjoyed "testing out different combinations and locations of weights on each side of the bar." Initial analyses of the item responses indicated high internal consistency for all three exhibits.[9] Therefore, the responses were averaged across the five items for each participant to generate an overall math enjoyment score for the family group.

Math awareness. Although encouraging visitors to explicitly associate their interactions at the exhibits with the topic of mathematics was not a goal of the REVEAL facilitation approach, we were interested in how the facilitation impacted this aspect of the experience, especially given ongoing debates in the field about the importance of associations with specific STEM topic domains during informal STEM learning experiences (e.g., Pattison, Rubin, & Wright, 2017). To assess the extent to which families associated their exhibit experiences with mathematics, we coded responses to the open-ended survey question: "What would you tell a friend this exhibit is about?" Coding categories were developed inductively, based on multiple reviews of participant responses, and tested by two members of the research team, in close consultation with the project team's math education expert.

The final coding rubric assessed whether or not (i.e., "yes" or "no") responses made an explicit connection to mathematics (e.g., using the term *math* or a math-related concept, such as multiplication or graph). Interrater reliability for this code was extremely high for the final dataset (rater agreement = 98.9%, Cohen's kappa = 0.97). After assessing reliability, discrepancies between the two coders were discussed and resolved.

Mathematical reasoning. The level of mathematical reasoning related to the algebraic relationships in the exhibits was assessed by coding video and audio data of each staff-family interaction. This process included iterative rounds of development and testing by members of the research team and formal testing with four other researchers, including two bilingual (Spanish/English) staff members who had not previously been involved in the project. Similar to the math enjoyment measure, we chose to develop separate but parallel indicators for each of the three exhibits.

The final coding rubric for each exhibit assessed the level of mathematical reasoning across four different dimensions: (a) talking about mathematical quantities, (b) describing mathematical relationships among those quantities, (c) exploring mathematical relationships in the exhibit, and (d) achieving mathematical goals. The first two dimensions focused on verbal indicators of mathematical reasoning, and the second two highlighted visitor behaviors and interactions with the exhibits. For each of these dimensions, coders rated the level of mathematical reasoning from one (no indicator behaviors present) to five (highest level of indicators present). For example, at the Balancing Art exhibit, visitors who mentioned the weight or distance labels on the exhibit were rated as a two or three for talking about mathematical quantities, depending on if they mentioned only one or both of these labels. Visitor groups that went beyond mentioning these labels and described the importance of weight, distance, and balance at the exhibit were rated at level five. Because our unit of analysis was the family, comments and behaviors from any family member, child or adult, counted towards the overall math reasoning ratings.

For the videos selected for final analysis, two of the coders from the initial testing team coded all of the videos. Ratings from the two researchers were then averaged for each dimension of mathematical reasoning. Every few weeks, the two coders met with members of the research team to review results, discuss and resolve major discrepancies, and ensure that application of the rubrics remained consistent. Intraclass correlation coefficients, used as a measure of interrater reliability, were extremely high for the four dimensions.[10] The internal consistency was also strong among the four dimensions of mathematical reasoning for each exhibit.[11] For the final analyses, ratings for each dimension were averaged to create an overall measure of mathematical reasoning for each family group. The potential range of the rating was from one (no indicators observed across any of the four dimensions) to five (all indicators present for all four dimensions). Actual ratings ranged from 1.0 to 4.9.

Intergenerational communication. We used the video and audio data to assess three aspects of intergenerational communication: (a) frequency of adult communication, (b) frequency of child communication, and (c) interactivity between adult and child family members. The rubric development process was identical to that described above for mathematical reasoning.

The final coding rubric assessed each of the three dimensions of communication on a scale of one to seven. For frequency of adult communication, a one rating indicated that adult family members almost never spoke throughout the entire interaction, whereas a seven indicated that adults talked consistently and participated continuously in conversations throughout the entire interaction. The ratings were similar for the frequency of child communication dimension. For the interactivity dimension, a one rating was defined as when adult and child family members spoke, they almost never spoke to each other throughout the entire interaction. In other words, adult and child talk was directed at the staff facilitator, or adults primarily spoke to adults and children primarily spoke to other children. At the other extreme, a seven rating was characterized by adult and child members speaking to each other almost every time they spoke, rather than to the facilitator or to family members of the same generation (i.e., adults to adults, children to children).

Overall, interrater reliability was strong for this measure during testing and final coding.[12] Because these three measures were conceptualized as distinct and independent aspects of intergenerational communication, scores for each, averaged across the two coders, were used separately for the final analyses.

Engagement time. To understand how facilitation impacts the length of time that families spend at interactive exhibits, we used the video data to assess total engagement time for each family group. This measure began when the first child or adult member of the group engaged with or stepped within 5 feet of the exhibit and ended when the last adult or child family member left the informed consent space or stopped engaging with the exhibit for at least 30 seconds. This measure of engagement time was also used as a control variable in the analyses to explore how much of the impact of staff facilitation was due to keeping families at the exhibit longer and how much was attributable to the strategies of the facilitator, above and beyond the impact of longer engagement times.

Data Analysis

To understand not only how outcomes compared between the two experimental conditions but also how all the variables related to each other and the unique contribution of each variable to explaining the impacts of facilitation on family learning, we used linear and logistic regression analyses (Tabachnick & Fidell, 2007). For example, for levels of mathematical reasoning, we simultaneously assessed the extent to which variation across families for this measure was explained by the two facilitation conditions, differences across the three exhibits, differences among the four educators, and visitor group characteristics. We explain the interpretation of these regression analyses in more detail below.[13]

Results

We begin by outlining basic descriptive and bivariate statistics for the measures used in our study. We then introduce the multivariate regression models that were designed to directly address our research question.

Bivariate Results

Table 1. *Descriptive statistics for continuous and interval outcome variables, by condition*

	Facilitation (*n* = 171)		Greeting (*n* = 92)	
Measure	*M*	*SD*	*M*	*SD*
Video coding variables				
Time (min.)	9.05	5.41	5.22	3.71
Math reasoning	3.38	0.81	2.88	1.01
IC child	2.91	1.24	2.83	1.38
IC adult	3.24	1.41	3.14	1.40
IC interact	4.41	1.88	5.44	1.69
Survey variables				
General satisfaction	5.97	1.06	5.51	1.22
Math enjoyment	6.12	0.87	5.88	1.00

Note. IC = intergenerational communication. The sample size for math enjoyment was 170 for the facilitation condition and 91 for greeting.

Table 1 outlines the means and standard deviations for the eight continuous or interval outcome variables included in the study. In general, engagement time was high for all participants, with mean time spent at the exhibit just under twice as high for the facilitating condition (9.05 minutes) compared to the greeting condition (5.22 minutes). Levels of mathematical reasoning fell towards the middle of the 5-point scale, with mean scores slightly higher for the facilitating condition (3.38) compared to the greeting condition (2.88). Ratings for levels of intergenerational communication (IC) were more similar across the two conditions, with mean IC child and IC adult falling on the lower half of the rubric (between almost never speaking and speaking about half the time), while IC interact mean values fell closer to the center of the rubrics (i.e., when adults and children verbally communicated, more than half the time they communicated with each other).

For the survey measures, mean levels of visitor general satisfaction and math enjoyment were very high across both experimental conditions. For general satisfaction, mean scores were between *agree* and *strongly agree* for the five satisfaction items. Similarly, for math enjoyment, mean scores were between *enjoyable* and *very enjoyable*. As shown in Table 2, about a third of participants in each condition were coded as making connections to mathematics in the survey responses.

Table 2. *Percentage of math awareness coding, by experimental condition*

Connections to math	Facilitation (n = 127)	Greeting (n = 70)
Yes	29.92%	34.29%
No	70.08%	65.71%

Note. Percentages exclude missing responses.

Many of these outcome variables were correlated with each other. The strongest correlations were between IC interact and IC adult (r = 0.49, p < 0.001), mathematical enjoyment and general satisfaction (r = 0.47, p < 0.001), engagement time and mathematical reasoning (r = 0.44, p < 0.001), and IC adult and mathematical reasoning (r = 0.31, p < 0.001). Engagement time was negatively correlated with IC interact (r = -0.26, p < 0.001), perhaps because the more time families spent at the exhibits, the more likely it was that during some portion of the interaction children were engaging with the facilitator or other children, or adults were engaging with the facilitator or other adults.

As one approach to testing the cultural validity of the outcome measures used in the study, values for each of the primary outcome variables were compared across visitor demographic groups (gender, race, ethnicity, education level, and languages spoken at home). Only a few of the comparisons were statistically significant, including relationships between ethnicity and math enjoyment,[14] math reasoning and languages spoken at home,[15] and education level and general visitor satisfaction.[16] We speculate on the implications of these findings in the discussion section.

Multivariate Models

Next, we describe the results of the multivariate regression models for each of the outcome variables. All of these models included facilitation condition, exhibit, and educator as the experimental variables and number of visits in the last 12 months, adult survey respondent age, family group size, and average child age within the family group as control variables. These controls were added to account for variation among visitor groups across the experimental conditions and to assess the importance of the key visitor characteristics hypothesized during the previous DBR study (Figure 1). For the experimental variables, the greeting condition, the Balancing Art exhibit, and "Educator A" served as the reference categories. For these variables, the regression models show the predicted impact on the outcome variable of moving from the reference category (e.g., greeting condition) to the predictor category (e.g., facilitating condition). For the continuous visitor characteristic variables, the models show the impact on the outcome of increasing the continuous variable by one unit (e.g., increasing the age of the adult survey respondent by one year). All of the results show the unique contribution of each variable, controlling for all other variables in the model.

In the results below, we provide the full specifications for the regression models in the tables but focus on describing the comparison between the facilitating and greeting conditions, as aligned with our primary research question.

Engagement time. Table 3 shows the results for the linear regression analysis with family engagement time as the outcome variable. Compared to the greeting condition, families in the facilitating condition spent 3.94 minutes longer on average at the exhibit, controlling for all other variables ($\beta = 3.94$, $t(225) = 6.17$, $p < 0.001$, $r_s^2 = 0.14$). Based on the model R^2, all of the variables collectively explained approximately 23% of the variance in engagement time across families.

Table 3. *Linear multiple regression models for engagement time*

Variable	β
Intercept	1.78
Experimental variables	
Condition (facilitation)	3.94***
Exhibit (DfS)	-0.81
Exhibit (DiM)	0.81
Educator (B)	0.79
Educator (C)	2.64**
Educator (D)	4.22***
Control variables	
No. of visits	0.04
Adult age	0.02
Group size	-0.01
Avg. child age	0.07
R^2	0.23
F	6.88***
n	236

Note. Standard regression coefficient shown. DfS = Designing for Speed, DiM = Drawing in Motion.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Mathematical reasoning. The multivariate analysis for the mathematical reasoning variable was nearly identical to the engagement time model (Table 4). In this case, we examined the impact of the experimental and control variables on level of mathematical reasoning for each of the three exhibits separately. For each exhibit, we constructed the models with engagement time as a control variable (Model 1) and without (Model 2) to better understand the relations among staff facilitation, engagement time, and level of mathematical reasoning. Collectively, the experimental and control variables were significant predictors of mathematical reasoning scores for all the models except BA and DfS Model 1. Model 2 for DiM explained the most variance in the outcome variable ($R^2 = 0.58$), while Model 1 for BA explained the least ($R^2 = 0.13$).[17]

For the models without engagement time (Model 1 for each exhibit), the impact of facilitation varied by exhibit. Controlling for differences across educators and visitor groups, mathematical reasoning scores were on average significantly higher in the facilitation group compared to the greeting only group for the DiM ($\beta = 0.84$, $t(63) = 4.17$, $p < 0.001$, $r_s^2 = 0.22$) and DfS exhibits ($\beta = 0.42$, $t(71) = 2.17$, $p = 0.033$, $r_s^2 = 0.06$) but not for BA ($\beta = 0.27$, $t(75) = 1.21$, $p = 0.023$, $r_s^2 = 0.02$). For DiM, mathematical reasoning scores in the facilitation group were on average just under one point higher (on the 5-point scale) compared to the greeting only condition, while scores were just under a half point higher on average for families in the facilitation group at the DfS exhibit.

As shown in model 2 for each exhibit, engagement time was a significant predictor of level of family mathematical reasoning for all three exhibits, controlling for all other variables in the model.[18] Based on the regression coefficients, a one-minute increase in engagement time was associated with a 0.10-, 0.10-, and 0.04-point increase in mathematical engagement score for the three exhibits, respectively. Adding engagement time had the largest impact on the BA and DiM regression models, explaining an additional 22% variance in mathematical reasoning scores. For all three exhibits, the addition of engagement time was associated with a decrease in the magnitude of the facilitation condition coefficient, suggesting that the impact of facilitation on mathematical reasoning was partially, but not fully, explained by the impact of facilitation on engagement time.

Table 4. *Linear regression models for mathematical reasoning, by exhibit*

	BA		DiM		DfS	
Variable	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept	1.80**	1.63**	1.00	1.15*	3.24***	3.05***
Experimental variables						
Condition (facilitation)	0.27	-0.21	0.84***	0.55**	0.42*	0.23
Educator (B)	0.27	0.13	0.52	0.34	-0.21	-0.18
Educator (C)	0.32	-0.01	0.30	-0.08	-0.59*	-0.64*
Educator (D)	0.28	-0.39	0.77**	0.39	-0.10	-0.19
Control variables						
No. of visits	0.06*	0.06*	0.01	0.01	0.01	0.01
Adult age	0.01	0.01	0.00	-0.01	-0.00	-0.00
Group size	0.06	0.05	0.16*	0.11*	-0.07	-0.04
Avg. child age	0.04	0.03	0.03	0.03	0.05	0.04
Engagement time	--	0.10***		0.10***		0.04*
R^2	0.13	0.35	0.36	0.58	0.18	0.23
F	1.43	4.51***	4.49***	9.57***	1.91	2.37*
ΔR^2	--	0.22	--	0.22	--	0.06
ΔF	--	25.41***	--	32.35***	--	5.14*
n	84	84	72	72	80	80

Note. Standard regression coefficient shown. BA = Balancing Art, DfS = Designing for Speed, DiM = Drawing in Motion.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Mathematical enjoyment. The structure of the multivariate mathematical enjoyment models was identical to the mathematical reasoning models (Table 5). In this case, very few of the experimental or control variables were significantly related to visitor levels of mathematical enjoyment, after controlling for all other variables in the models, possibly because the vast majority of visitors provided high ratings for this measure (Table 1). Facilitation did not have a significant impact on levels of mathematical enjoyment for any of the three exhibits, with or without engagement time included in the model. Based on the model R^2 values, only the variables in BA Model 2 explained a significant level of variance (24%) in visitor mathematical enjoyment ($R^2 = 0.24$, $F(9, 73) = 2.52$, $p = 0.014$).

Table 5. *Linear regression models for mathematical enjoyment, by exhibit*

	BA		DiM		DfS	
Variable	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept	6.13***	6.03***	6.15***	6.20***	6.28***	6.13***
Experimental variables						
Condition (facilitation)	0.08	-0.19	0.31	0.22	0.07	-0.06
Educator (B)	-0.09	-0.17	0.34	0.28	0.17	0.19
Educator (C)	-0.50*	-0.69**	-0.41	-0.52	-0.33	-0.37
Educator (D)	-0.44	-0.83**	-0.07	-0.19	-0.21	-0.28
Control variables						
No. of visits	0.02	0.02	0.02	0.02	0.01	0.01
Adult age	-0.00	-0.00	-0.01	-0.01	-0.01	-0.01
Group size	0.13	0.12	-0.01	-0.02	0.04	0.06
Avg. child age	-0.00	-0.01	-0.02	-0.02	-0.01	-0.01
Engagement time	--	0.06**	--	0.03	--	0.03
R^2	0.14	0.24	0.11	0.13	0.05	0.07
F	1.47	2.52*	0.99	1.04	0.48	0.60
ΔR^2	--	0.10	--	0.02	--	0.02
ΔF	--	9.55**	--	1.41	--	1.58
n	83	83	72	72	80	80

Note. Standard regression coefficients shown. BA = Balancing Art, DfS = Designing for Speed, DiM = Drawing in Motion. $*p < 0.05$. $**p < 0.01$. $***p < 0.001$.

General satisfaction. The general visitor satisfaction measure was identical for all three exhibits, so the multivariate analyses were conducted across the entire sample (Table 6). The models with engagement time (Model 1) and without (Model 2) were nearly identical, with Model 2 explaining a slightly larger proportion of the outcome variance (18% compared to 16%, $R^2\Delta = 0.02$, $F\Delta = 5.82$, $p = 0.017$). For Model 2, after controlling for all other variables, facilitation condition was significantly related to level of general visitor satisfaction, with scores being about a third of a point higher, on average, in the facilitating condition ($\beta = 0.34$, $t(224) = 2.13$, $p = 0.034$, $r_s^2 = 0.02$). Similar to mathematical reasoning, the addition of engagement time as a variable was associated with a decrease in the magnitude of the facilitation condition coefficient, suggesting that the impact of facilitation on general satisfaction was partially, but not fully, explained by the impact of facilitation on engagement time.

Table 6. *Linear regression models for visitor general satisfaction*

	General satisfaction	
Variable	Model 1	Model 2
Intercept	5.41***	5.35***
Experimental variables		
Condition (facilitation)	0.49**	0.34*
Exhibit (DfS)	-0.75***	-0.72***
Exhibit (DiM)	-0.20	-0.23
Educator (B)	0.22	0.19
Educator (C)	-0.30	-0.40*
Educator (D)	-0.04	-0.19
Control variables		
No. of visits	-0.03	-0.03
Adult age	0.01	0.01
Group size	0.07	0.07
Avg. child age	-0.01	-0.01
Engagement time	--	0.04*
R^2	0.16	0.18
F	4.39***	4.60***
ΔR^2	--	0.02
ΔF	--	5.82*
n	236	236

Note. Standard regression coefficient shown. DfS = Designing for Speed, DiM = Drawing in Motion.
*$p < 0.05$.**$p < 0.01$. ***$p < 0.001$.

Intergenerational communication. The multivariate regression models for intergenerational communication were similar to visitor satisfaction, with one model for each outcome variable across the entire sample and the three exhibits included as experimental variables (Table 7). For the IC adult and IC child models, facilitation condition was not significantly related to level of talk. In contrast, IC interact ratings were a little over half a point lower for groups in the facilitating condition, on average, compared to groups in the greeting condition (β= -0.63, $t(224) = -2.53$, $p = 0.012$, $r_s^2 = 0.03$). Although not shown, removing engagement time had relatively little effect on the three models and did not substantively change the interpretation of the model coefficients.

Table 7. *Multiple linear regression models for intergenerational communication*

Variable	IC child	IC adult	IC interact
Intercept	0.93	3.93***	6.96***
Experimental variables			
Condition (facilitation)	-0.18	0.09	-0.63*
Exhibit (DfS)	0.58**	-0.25	0.58*
Exhibit (DiM)	0.48*	0.06	0.51
Educator (B)	-0.16	-0.10	-0.07
Educator (C)	-0.38	-0.17	-1.09***
Educator (D)	-0.34	0.04	0.07
Control variables			
No. of visits	0.02	0.03	0.05
Adult age	0.01	-0.01	-0.03*
Group size	0.14*	0.03	-0.18*
Avg. child age	0.04	-0.04	0.05
Engagement time	0.06**	0.00	-0.06*
R^2	0.15	0.04	0.23
F	3.48***	0.84	5.93***
n	236	236	236

Note. Standard regression coefficient shown. DfS = Designing for Speed, DiM = Drawing in Motion.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Mathematical awareness. Finally, we analyzed the unique relations among the experimental and control variables and math awareness. For this analysis, we used logistic regression because of the categorical nature of the outcome variable. Similar to the previous multivariate analyses, logistic regression shows the unique impact of each variable on the outcome, controlling for differences across all other variables. For logistic regression, however, the regression coefficients are more complicated to interpret. The *odds ratio*, which we report in conjunction with the model coefficient, indicates the proportional increase in the outcome variable for each one-point increase in the experimental or control variable.

Table 8 outlines results for the analyses with engagement time included. As with intergenerational communication, the results were nearly identical for the model without engagement time. Based on a variety of recommended estimation techniques (Field et al., 2012), the predicators collectively explained between 26 and 38 percent of the variance in coded level of mathematical awareness. Controlling for all other variables, facilitation condition was not significantly related to the outcome.

Table 8. Logistic regression predicting coded level of visitor math awareness

Variable	β	Odds ratio
Intercept	-2.72*	--
Experimental variables		
Condition (facilitating)	0.51	0.60
Exhibit (DfS)	-3.80***	0.02
Exhibit (DiM)	-0.91*	0.40
Educator (B)	0.03	1.03
Educator (C)	-0.67	0.51
Educator (D)	-1.35*	0.26
Control variables		
No. of visits	0.15*	1.16
Adult age	0.02	1.02
Group size	0.25	1.29
Avg. child age	0.05	1.05
Engagement time	0.09	1.09
R^2 (Hosmer-Lemeshow)	0.26	
R^2 (Cox-Snell)	0.27	
R^2 (Nagelkerke)	0.38	
Model χ^2	56.03***	
n	177	

Note. Unstandardized logistic regression coefficients shown. DfS = Designing for Speed, DiM = Drawing in Motion.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Discussion

In this study, we set out to rigorously investigate the impact of staff facilitation on family learning using the REVEAL responsive facilitation model, compared to a “greeting only” condition. Using a quasi-experimental design and data collected from visitors through both videotaped observations and post-interaction surveys, we found strong evidence that facilitators trained in the REVEAL responsive facilitation model positively influenced some family learning outcomes at interactive math exhibits, although the outcomes varied by exhibit.

First, facilitation by a trained staff member greatly increased the amount of time that families spent at exhibits. Even after controlling for differences among exhibits, staff members, and visitor characteristics, families spent almost 4 minutes longer, on average, in the facilitation condition than in the greeting condition. Furthermore, the amount of time families spent at the exhibits was positively correlated with a number of other outcome measures: family mathematical reasoning, math enjoyment (for the BA exhibit), and general visitor satisfaction. In other words, the educators using the REVEAL facilitation approach appeared to be quite adept at keeping families engaged at the exhibits and creating opportunities for deeper mathematical reasoning that families found enjoyable and satisfying.

These findings suggest that facilitators enrich family learning not only by increasing engagement time, but also by increasing the quality of that engagement through facilitation strategies. For example, visitor satisfaction was higher in the facilitation condition than the greeting condition, even after controlling for engagement time. Visitors appeared to be more satisfied with the facilitated experience both because it was longer and because the facilitators provided richer experiences. In the Drawing in Motion exhibit, levels of mathematical reasoning were significantly higher in the facilitating condition, even after controlling for engagement time. At least at this exhibit, families appeared to have higher levels of mathematical reasoning both because they had a longer time to engage and because the facilitator was able to deepen their mathematical experience.

The study results also highlighted limitations of the REVEAL facilitation approach. Despite the project's focus on supporting engagement for the whole family, levels of interaction among children and adults (one of the three measures of intergenerational communication) were lower in the facilitation condition compared to the greeting condition. Either by unintentionally disrupting intergenerational communication, or simply by creating a new focus of attention, the presence of facilitators appeared to lower the relative amount of time that adults and children spent interacting with each other. Other aspects of visitors' experience were not affected by facilitation. Facilitators had no impact on reported levels of mathematical enjoyment or visitors' awareness of the mathematics in the exhibit. In the case of enjoyment, this may be due to a ceiling effect, as most visitors reported high levels of enjoyment. In the case of math awareness, this lack of impact was likely due to the fact that the REVEAL facilitation approach did not emphasize explicitly mentioning the terms *math* or *mathematics*. As we discuss below, these findings suggest trade-offs to incorporating facilitators into an exhibit space, as well as the importance of carefully choosing educational goals for facilitators to pursue.

The REVEAL facilitation model hypothesized that a variety of physical, personal, and social factors would influence the nature and outcomes of facilitator-visitor interactions. Although not the focus of this article, the regression models demonstrated that all of the control variables except average child age showed a significant relationship with at least two of the outcome measures in the regression models, although none of the variables was a consistently strong predictor of outcomes. Therefore, the study indicates that the model is promising for describing the factors and processes influencing staff-facilitated interactions at museum exhibits. More work is needed, however, to understand when and why different factors are important.

One of the strongest effects on the outcome variables was the exhibit itself. Differences across the three exhibits not only directly influenced several of the outcome variables (visitor satisfaction, intergenerational communication, and math awareness) but also interacted with the impact of facilitation. For example, facilitation had a significant effect on mathematical reasoning at Drawing in Motion but not at the other two exhibits. We speculate that this was due to the mathematical relationships in the exhibit being less accessible for visitors than those in the two other REVEAL exhibits. Drawing in Motion required the visitors controlling the horizontal and vertical sliders to consider the relative speed of their movements when drawing a diagonal line, which was a fairly complicated mathematical idea. Therefore, facilitation may have been key to helping visitors see and use this relationship. In contrast, the relationship in the Designing for Speed exhibit between mass distribution and wheel acceleration was relatively straight forward (at least as qualitatively represented in the activity) and likely required less facilitator support to help visitors engage with the mathematics.

Implications for Practice

For museum professionals, this study provides some of the first rigorous evidence that staff facilitation by trained educators can have a positive impact on family engagement and learning during unstructured interactions. Although some prior studies have suggested that staff facilitation can occasionally be unwelcome (Marino & Koke, 2003), and researchers have speculated that staff members might interfere with learning in informal learning environments (National Research Council, 2009), this study found that facilitation using the REVEAL approach was positively related to engagement time, mathematical reasoning, and visitor satisfaction. Of course, these results are limited to a specific museum context, facilitation approach, and set of exhibits, and more research is needed to understand how the results might generalize beyond the REVEAL context. Nonetheless, the findings provide preliminary motivation for managers to consider how they invest in and support front-line staff.

The study also emphasizes the importance of high quality, research-based professional development. The facilitators who participated in the research engaged in five days of intensive training over two weeks. Although two weeks is relatively short compared to many professional development programs for teachers, it is significantly more training than most museum educators receive, especially around a single topic or exhibit. The training was based on a year-long collaborative investigation between researchers, expert facilitators, and a math educator (Pattison, Randol, et al., 2017); built on research about family learning outside of school; and focused on strategies appropriate to informal settings. The REVEAL facilitation approach placed equal weight on educators' goals (e.g., supporting mathematical reasoning) and visitors' goals (e.g., supporting visitor satisfaction), recognizing that the fulfillment of personal goals is central to informal learning (Falk, 2009; Falk & Dierking, 2013). Although future research is needed to directly test the impact of staff training, such as what was provided during this study, we believe that the positive impacts of facilitation demonstrated through the REVEAL project depended on high-quality professional development.

Beyond providing more professional development, there is a need for educators and managers to carefully consider their goals for families. No facilitation approach can improve all visitor outcomes simultaneously; there are always trade-offs. The REVEAL facilitation approach was no exception, as it improved some outcomes, decreased one, and had no effect on others. Therefore, managers should make decisions about facilitation resources based on specific goals. For example, if intergenerational communication is the priority, putting resources into facilitation may not be as useful as focusing on stand-alone exhibits, at least when using an approach similar the REVEAL facilitation model.

Finally, there are also lessons to be learned from the REVEAL study about exhibit design. The REVEAL exhibits were intentionally created to provide facilitators with opportunities to add value to visitor interactions, and the exhibit prototyping process explicitly focused on these facilitation affordances (e.g., props or tools to enhance the exhibit experience that only museum educators could make available). Again, we believe this was an essential component of the success of the REVEAL facilitation approach, although more research is needed to compare the impact of facilitation at a broader range of exhibits, including those not intentionally designed to support staff-family interactions. In the meantime, exhibit developers should consider the role of staff and their interactions with visitors in the design of exhibits that are likely to be facilitated. Similarly, education managers can look across their exhibit spaces and make strategic choices about which exhibits are likely to benefit from staff facilitation and which are better left for families to engage with on their own.

Implications for Research

This study was an important step in our ongoing efforts to understand staff facilitation and its impact on family learning at exhibits and in other designed informal learning environments. As educators and policymakers increasingly recognize the importance of out-of-school learning, it is critical that researchers provide evidence and insight to guide the roles of educators in these spaces. This study provides a model for future investigations to rigorously assess the impact of facilitation while remaining true to the informal/free-choice nature of these experiences. Using a quasi-experimental design, we attempted to balance the rigor of the study, including eliminating potential alternative explanations to observed differences between the two conditions, and the naturalistic quality of the experiences, such as visitors' freedom to approach the experience without being recruited. This study design was ideal for understanding the overall impact of the REVEAL facilitation approach, but many questions remain about which aspects of the approach are more or less essential.

We are particularly interested in further research on the complex concept of visitor-facilitator match. In the first phase of the REVEAL project, we hypothesized that visitor-facilitator match was essential to effective facilitation in informal/free-choice settings (Pattison, Randol, et al., 2017). This concept of match refers to how well facilitators notice the needs and interests of different family groups and respond accordingly. In this study, we attempted to operationalize the concept of match both through an index of post-interaction survey items, prompting visitors to rate how well staff members matched their needs, and through a video coding rubric assessing observed indicators of match. In the case of the survey items, the vast majority of visitors rated the facilitators as providing a perfect match, perhaps because they were predisposed to give the educators the benefit of the doubt or because they had little prior experience with these types of facilitated experiences. This resulted in minimal variation for the measure, making it essentially useless for analysis. The video coding rubric had a different problem. Although raters were able to observe differences in match (e.g., level of visitor comfort using the interactive exhibit and corresponding level of orientation support from educators), they were not able to achieve an acceptable level of reliability, suggesting that the measure as we described it was too imprecise or subjective. We believe that the notion of match between family needs and facilitator strategies is critical for museum educators and recommend that the field pursue the development of a measure to be used in future research.

Additional work is also needed to continue to assess and improve the variety of measures and instruments developed through this study to investigate staff-family interactions at interactive math exhibits. This includes the cultural validity of each measure (Kirkhart, 1995; Kirkhart & Hopson, 2010). Following a culturally responsive research (CRR) approach, rather than using demographic variables in our analyses without theoretical justification, we explored associations between family demographics and the study variables in order to provide an initial test of the cultural validity for our outcome measures. The analyses found a few statistically significant associations. Hispanic/Latino families reported higher levels of math enjoyment on average compared to non-Hispanic/Latino families, families that spoke only English at home were coded for higher levels of math reasoning on average compared to other families, and adult education level was negatively associated with general satisfaction. Each of these results could mean either that important relationships exist between the family characteristic and the outcome variable or that there are culturally based assumptions or biases in the measure. For example, it may be that because the mathematical reasoning coding rubrics rely on both verbal and non-verbal visitor indicators, the measure underestimates the levels of mathematical reasoning of families that reported speaking languages other than English at home. To ensure that approaches to exhibit facilitation support all families, these and other measures should be continually tested and revised so that they are culturally responsive and inclusive.

References

Allen, S., Gutwill, J. P., Perry, D., Garibay, C., Ellenbogen, K., Heimlich, J., … Klein, C. (2007). Research in museums: Coping with complexity. In J. H. Falk, L. D. Dierking, & S. Foutz (Eds.), *In principle, in practice: Museums as learning institutions* (pp. 44–56). Lanham, MD: AltaMira.

Ash, D. (2002). Negotiations of thematic conversations about biology. In G. Leinhardt, K. Crowley, & K. Knutson (Eds.), *Learning conversations in museums* (pp. 357–400). Mahwah, NJ: Erlbaum.

Ash, D., & Lombana, J. (2013). Reculturing museums: Working toward diversity in informal settings. *Journal of Museum Education, 38*(1), 69–80.

Astor-Jack, T., Whaley, K., Dierking, L. D., Perry, D., & Garibay, C. (2007). Understanding the complexities of socially-mediated learning. In J. H. Falk, L. D. Dierking, & S. Foutz (Eds.), *In principle, in practice: Museums as learning institutions* (pp. 217–228). Lanham, MD: AltaMira.

Bevan, B., & Xanthoudaki, M. (2008). Professional development for museum educators: Unpinning the underpinnings. *Journal of Museum Education, 33*(2), 107–119.

Brown, J. C., & Crippen, K. J. (2017). The knowledge and practices of high school science teachers in pursuit of cultural responsiveness. *Science Education, 101*(1), 99–133.

Cox-Petersen, A. M., Marsh, D. D., Kisiel, J., & Melber, L. M. (2003). Investigation of guided school tours, student learning, and science reform recommendations at a museum of natural history. *Journal of Research in Science Teaching, 40*(2), 200–218.

Crowley, K. D., Callanan, M. A., Jipson, J. L., Galco, J., Topping, K., & Shrager, J. (2001). Shared scientific thinking in everyday parent-child activity. *Science Education, 85*(6), 712–732.

Crowley, K. D., & Jacobs, M. (2002). Building islands of expertise in everyday family activity. In G. Leinhardt, K. D. Crowley, & K. Knutson (Eds.), *Learning conversations in museums* (pp. 333–356). Mahwah, NJ: Erlbaum.

Cunningham, M. K. (2004). *The interpreter's training manual for museums*. Washington, DC: American Association of Museums.

Dierking, L. D., & Falk, J. H. (1994). Family behavior and learning in informal science settings: A review of the research. *Science Education, 78*(1), 57–72.

Ellenbogen, K. (2002). Museums in family life: An ethnographic case study. In G. Leinhardt, K. Crowley, & K. Knutson (Eds.), *Learning conversations in museums* (pp. 81–101). Mahwah, NJ: Erlbaum.

Ellenbogen, K., Luke, J. J., & Dierking, L. D. (2007). Family learning in museums: Perspectives on a decade of research. In J. H. Falk, L. D. Dierking, & S. Foutz (Eds.), *In principle, in practice: Museums as learning institutions* (pp. 17–30). Lanham, MD: AltaMira.

Falk, J. H. (2009). *Identity and the museum visitor experience*. Walnut Creek, CA: Left Coast Press.

Falk, J. H., & Dierking, L. D. (2013). *The museum experience revisited*. Walnut Creek, CA: Left Coast Press.

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Thousand Oaks, CA: Sage.

Frechtling, J. (2010). *The 2010 user-friendly handbook for project evaluation*. Washington, DC: National Science Foundation. Retrieved from http://informalscience.org/documents/TheUserFriendlyGuide.pdf

Friedman, A. J. (2008). *Framework for evaluating impacts of informal science education projects: Report from a National Science Foundation workshop*. Retrieved from http://informalscience.org/documents/Eval_Framework.pdf

Garibay, C., Yalowitz, S., & Guest Editors. (2015). Redefining multilingualism in museums: A case for broadening our thinking. *Museums & Social Issues, 10*(1), 2–7.

Garibay Group. (2013). *Design Zone exhibition summative report*. Portland, OR: OMSI. Retrieved from http://informalscience.org/evaluation/ic-000-000-008-817/Design_Zone_Exhibition_Summative_Report

González, N., Moll, L. C., & Amanti, C. (2005). *Funds of knowledge: Theorizing practice in households, communities, and classrooms*. Mahwah, NJ: Erlbaum.

Greenes, C. E., & Rubenstein, R. (Eds.). (2008). *Algebra and algebraic thinking in school mathematics: Seventieth yearbook*. Reston, VA: National Council of Teachers of Mathematics.

Gutiérrez, K. D., & Calabrese Barton, A. (2015). The possibilities and limits of the structure-agency dialectic in advancing science for all. *Journal of Research in Science Teaching, 52*(4), 574–583.

Gutierrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher, 32*(5), 19–25.

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

Hacker, K. (2013). *Community-based participatory research*. Thousand Oaks, CA: SAGE.

Jarvis, T., & Pell, A. (2005). Factors influencing elementary school children's attitudes toward science before, during, and after a visit to the UK National Space Centre. *Journal of Research in Science Teaching, 42*(1), 53–83.

Kaput, J. J., Carraher, D. W., & Blanton, M. L. (2008). *Algebra in the early grades*. New York, NY: Erlbaum/National Council of Teachers of Mathematics.

Kirkhart, K. E. (1995). 1994 conference theme: Evaluation and social justice seeking multicultural validity: A postcard from the road. *American Journal of Evaluation, 16*(1), 1–12.

Kirkhart, K. E., & Hopson, R. (2010, June). *Strengthening evaluation through cultural relevance and cultural competence*. Presented at the American Evaluation Association/Centers for Disease Control Summer Institute, Atlanta, GA.

Marino, M., & Koke, J. (2003, February). Face to face: Examining educational staff's impact on visitors. *ASTC Dimensions*. Retrieved from http://astc.org/pubs/dimensions/2003/jan-feb/index.htm

Moses, B. (Ed.). (1999). *Algebraic thinking, grades K-12: Readings from NCTM's school-based journals and other publications*. Reston, VA: National Council of Teachers of Mathematics.

National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: National Academies Press.

National Research Council. (2015). *Identifying and supporting productive STEM programs in out-of-school settings*. Washington, DC: The National Academies Press.

National Research Council & Institute of Medicine. (2002). *Community programs to promote youth development*. Washington, DC: National Academies Press.

Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment, 7*(3), 367–375.

Packer, J. (2004). *Motivational factors and the experience of learning in educational leisure settings* (Doctoral dissertation). Queensland University of Technology, Queensland, AU. Retrieved from http://eprints.qut.edu.au/15911/1/Jan_Packer_Thesis.pdf

Palmquist, S., & Crowley, K. (2007). From teachers to testers: How parents talk to novice and expert children in a natural history museum. *Science Education, 91*(5), 783–804.

Patrick, P. G. (Ed.). (2017). *Preparing informal science educators: Perspectives from science communication and education*. Cham, Switzerland: Springer.

Pattison, S. A. (2011). *Access Algebra staff facilitation: A formative evaluation report*. Retrieved from http://www.omsi.edu/sites/all/FTP/files/evaluation/algebrastafffacilitation.pdf

Pattison, S. A., & Dierking, L. D. (2012). Exploring staff facilitation that supports family learning. *Journal of Museum Education, 37*(3), 69–80.

Pattison, S. A., & Dierking, L. D. (2013). Staff-mediated learning in museums: A social interaction perspective. *Visitor Studies, 16*(2), 117–143.

Pattison, S. A., Randol, S. M., Benne, M., Rubin, A., Gontan, I., Andanen, E., … Dierking, L. D. (2017). A design-based research study of staff-facilitated family learning at interactive math exhibits. *Visitor Studies, 20*(2), 138–164.

Pattison, S. A., Rubin, A., & Wright, T. (2017, March 3). *Math in the Making: Reflections for the Field* (Webpage). Retrieved March 11, 2017, from http://www.informalscience.org/news-views/math-making-reflections-field

Peter, N. (2007). *Promising practices in out-of-school time professional development*. Retrieved from http://repository.upenn.edu/cgi/viewcontent.cgi?article=1098&context=spp_papers

Rigney, J. C., & Callanan, M. A. (2011). Patterns in parent–child conversations about animals at a marine science center. *Cognitive Development, 26*(2), 155–171.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Sindorf, L., Gutwill, J., & Garcia-Luis, V. (2015). Gaining visitor consent for research III: A trilingual posted-sign method. *Curator: The Museum Journal, 58*(4), 369–381.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson/Allyn & Bacon.

Tal, T., & Morag, O. (2007). School visits to natural history museums: Teaching or enriching? *Journal of Research in Science Teaching, 44*(5), 747–769.

Tran, L. U. (2007). Teaching science in museums: The pedagogy and goals of museum educators. *Science Education, 91*(2), 278–297.

Acknowledgments

We gratefully acknowledge the many colleagues and partners who made this project possible, including Leticia Aguilar, Patricia Alvarado, Jaclyn Barber, Karyn Bertschi, Jana Borgen, Summer Brandon, Marta Civil, Michael Coe, Chris Cunningham, Mary Kay Cunningham, Rebekah Elliott, Katie Forbes, Cecilia Garibay, Josh Gutwill, Andrew Haight, Laura Huerta-Migus, Natalie Johnson, Chip Lindsey, Jan Mokros, Aaron Nash, Ricardo Nemirovsky, Veronika Nuñez, Kate Nuhring, Maria Perdomo, David Perry, Jen Powers, Allison Prasad, Scott Randol, David Redburn, Lauren Retzlaff, Bob Reynolds, Cate Rhodes, Saul Rockman, Susan Jo Russell, Jessica Shamek, Sam Siciliano, Mary Soots, Matt Suplee, Omar Vargas, and Barry Walther. This material is based upon work supported by the National Science Foundation under Grant Number DRL-1321666. 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.

About the Authors

Scott A. Pattison, Ph.D., is a Research Scientist at TERC. He works with communities to study and support education, learning, and interest development in free-choice and out-of-school environments, including museums, science centers, and everyday settings. Address correspondence to: Scott Pattison, TERC, 2067 Massachusetts Ave., Cambridge, MA 02140. Email: spattiso@gmail.com.

Andee Rubin is a Senior Scientist at TERC. A mathematician and computer scientist, she has designed and studied innovations in math and science education both in and out of school for over 30 years.

Marcie Benne, Ph.D., is the Director of Engagement Research and Advancement at OMSI. For over 15 years she has advanced studies of people accessing, interacting with, and learning within designed environments.

Ivel Gontan is the Community Programs Senior Manager at the Reuben H. Fleet Science Center. Her focus is primarily on advocating for underserved audiences through informal science education.

Elizabeth Andanen has worked as an educator and assistant manager in the Museum Education department of OMSI since 2006. In her dozen years with OMSI, she has been most interested in how facilitators can and do affect visitors' experiences at activities and exhibits.

Todd Shagott, Ph.D., is a researcher at OMSI. His work is currently focused on exhibit evaluation and visitor interactions.

Melanie Francisco, Ph.D., currently works as a Statistical Research Analyst with Kaiser Permanente Center for Health Research Northwest, providing data and analytic support on various public health studies.

Smirla Ramos-Montañez, Ph.D., is a Research and Evaluation Associate at OMSI. She is interested in culturally responsive research and evaluation and STEM learning in informal and formal settings.

Crosby Bromley is a teacher for Salem-Keizer School District. She has been teaching in traditional and non-traditional settings for over twelve years, continually striving to find the best way to reach all her students.

Lynn D. Dierking, Ph.D., is a Sea Grant Professor in Free-Choice STEM Learning, Colleges of Science & Education, Oregon State University, and Director of Strategy & Partnerships, Institute for Learning Innovation. Her research involves lifelong learning, particularly free-choice, out-of-school time learning (in after-school, home-, community-based, and cultural contexts), with a focus on youth, families, and community, particularly those who have historically not identified with science.

Notes

[1] For example, Reflecting on Practice (http://reflectingonpractice.org/) and REFLECTS (http://www.informalscience.org/news-views/reflects-model-professional-development).

[2] https://REVEAL.TERC.edu

[3] More details about the REVEAL facilitation approach, including videos of educators interacting with families using REVEAL facilitation strategies, can be found on the project website.

[4] More details about the project's CRR approach can be found on the project website.

[5] https://omsi.edu/exhibitions/designzone/

[6] A power analysis allows researchers to determine the sample size needed to detect statistically significant relationships and differences of the desired size within the data, often with an eye to detecting differences that are of practical significance. Based on our original power analysis, our target minimum sample size was 228 family groups, including 152 groups in the facilitation condition and 76 groups in the non-facilitation condition. This sample size provided sufficient power (0.80) to detect a change in R^2 of 0.03 (small effect size) for the entire sample and 0.05 (small to medium effect size) within the facilitation condition, assuming a critical value of 0.05, one test variable, and 12 total variables (Faul, Erdfelder, Lang, & Buchner, 2007).

[7] http://reflectingonpractice.org/

[8] Informed consent procedures for videotaping followed the posted-sign method developed by Gutwill and colleagues (Gutwill, 2003; Sindorf, Gutwill, & Garcia-Luis, 2015).

[9] Balancing Art (α = 0.88 [95% CI = 0.78, 0.98]), Drawing in Motion (α = 0.88 [95% CI = 0.77, 0.98]), and Designing for Speed (α = 0.92 [95% CI = 0.83, 1.01). The scale range across the three exhibits was 3 to 7.

[10] Intraclass correlations ranged from 0.85 to 0.91 during the testing phase and 0.79 to 0.84 for the final video coding. Practically speaking, this means that the vast majority of the variation across ratings (at least 79%) was attributable to actual differences in mathematical indicators across participant groups, rather than differences in coding interpretations among raters.

[11] For the final set of coded videos: Balancing Art (α = 0.78 [95% CI = 0.63, 0.93]), Drawing in Motion (α = 0.82 [95% CI = 0.68, 0.97]), and Designing for Speed (α = 0.67 [95% CI = 0.49, 0.85]).

[12] Intraclass correlation coefficients during initial testing ranged from 0.78 to 0.94. For the final coding, coefficients were 0.70 and 0.69 for adult communication and interactivity, respectively. However, the coefficient for child communication was lower (0.49). Therefore, results with this measure must be interpreted cautiously. Final scores ranged from 1 to 7 for all three measures.

[13] All multivariate models were reviewed for violations of the assumptions of linear and logistic regression analysis, including evidence of abnormal residual distribution (e.g., heteroscedasticity), outliers with undue influence on model parameters, independence of errors, and multicollinearity (Field, Miles, & Field, 2012; Tabachnick & Fidell, 2007). In cases where assumptions were violated, we conducted bootstrapped regression analyses with the data, following Fields and colleagues (2012) and using the recommended 2000 iterations. Unless noted, parameters for the bootstrapped regressions were nearly identical to the original models and interpretation of the model coefficients and parameters was substantively identical, providing strong evidence for the robustness and statistical generalizability of the model results.

[14] Welch two-sample t-test, $t(23.22) = 2.574$, $p = 0.017$, Mean (Hispanic/Latino) = 6.45, Mean (Non-Hispanic/Latino) = 6.01.

[15] Linear model, $F(2, 248) = 7.50$, $p < 0.001$, Mean (English only) = 3.14, Mean (English and other) = 2.74, Mean (Non-English only) = 2.10, $R^2 = 0.057$.

[16] Pearson correlation, $r = -0.124$, p = 0.046, $n = 259$.

[17] The residuals for BA model 2 showed some signs of heteroscedasticity and nonlinearity, possibly because of the positively skewed distribution for the engagement time variable. In the boostrapped regression, the parameter confidence intervals were nearly identical to the original model, with the exception of the number of visits in last 12 months, for which the bootstrapped confidence interval crossed zero. In other words, the model appeared to generalize, although there is not strong evidence for a generalizable, statistically significant relationship between number of visits to the museum and mathematical reasoning at the BA exhibit.

[18] BA exhibit: $\beta = 0.10$, $t(74) = 5.04$, $p < 0.001$, $r_s^2 = 0.26$. DfS exhibit: $\beta = 0.04$, $t(70) = 2.27$, $p = 0.026$, $r_s^2 = 0.07$. DiM exhibit: $\beta = 0.10$, $t(62) = 5.69$, $p < 0.001$, $r_s^2 = 0.34$.