

AISL NYSCI Summative Evaluation Report

Designing for Narrative and Cultivating Empathy in Museum-Based Settings: Supporting Girls' Engagement in Engineering Practices

Kylie Peppler, Anna Keune, & Maggie Dahn
The Creativity Labs
University of California, Irvine

March 2020



This work was supported in part through a grant by the National Science Foundation [#1712803] awarded to Dorothy Bennett, Katherine McMillan, Susan Letourneau, and Peggy Monahan. Any opinions, findings, and conclusions or recommendations expressed are not those of the National Science Foundation.

I. Introduction and background	4
Underrepresentation of girls in engineering and the role of museum settings	4
A promise for inclusive engineering activities: Cultivating empathy through narrative	4
II. Project goals and research questions	6
III. Museum settings and activities	9
The New York Hall of Science Queens, NY	10
The Tech Interactive San Jose, CA	11
Scott Family Museum Amazeum Bentonville, AR	12
Introducing the engineering activities	13
Help Grandma/Invention Challenge	14
Chain Reaction	15
Air-Powered Vehicles	16
IV. Scope of research	21
V. Methods	23
Research instruments	23
Visitors	23
Narrative practices, engineering practices, and empathy markers	23
Data sources & analytical approach	25
VI. Findings	26
Narrative: To what extent does the condition type, activity type, or museum impact the diversity of narrative practices that visitors demonstrate?	26
Narrative: How correlated are the narrative practices across all visitors?	28
Engineering: To what extent does the condition type, activity type, or museum impact the diversity of engineering practices that visitors demonstrate?	29
Engineering: How does dwell time relate to diversity of engineering practices?	31
Engineering: How correlated are the observed engineering practices across all visitors?	33
Empathy: To what extent does the condition type, activity type, or museum impact the diversity of empathy markers that visitors demonstrate?	34
Empathy: How correlated are the empathy markers across all visitors?	37
What was the relationship between engineering practices and empathy markers?	37
How does dwell time relate to diversity of empathy markers and engineering practices?	39
How are activities designed to evoke empathy in engineering?	39
VII. Summary of findings	42
Summary of narrative findings	42
Summary of engineering findings	43
Summary of empathy findings	43
Summary of relationship between engineering and empathy	44
Summary of design recommendations	44

VIII. Questions for consideration	46
VIII. References	47
Appendix A: Observation protocol	50
Appendix B: Semi-structured interview protocol	51

I. Introduction and background

Underrepresentation of girls in engineering and the role of museum settings

Making and engineering design spaces have become prevalent in out-of-school settings such as museums, schools, libraries, and community settings and are often framed as places where underrepresented youth can be invited into engineering career pathways (Kalil, 2012; Keune, Peppler, & Wohlwend, 2019; Vossoughi & Bevan, 2014). However, there is evidence that inequities in engineering participation for underrepresented groups are replicated in these informal settings (Buchholz, Shively, Peppler, & Wohlwend, 2014; Buechley, Peppler, Eisenberg, & Kafai, 2013; Dawson, Seakins, Archer, Calabrese Barton, & Dierking, 2015). Girls in particular are less likely than boys to show interest in engineering in informal making and engineering spaces (Vossoughi, Hooper, & Escudé, 2016), and there is general evidence that girls are less well served than boys in science museums (Bevan, 2016; Feinstein & Meshoulam, 2014). There is additionally a risk that this trend carries over to the broader field of engineering as women continue to be consistently underrepresented in science and engineering professions (Bix, 2014; Buse, 2018; Sax et al., 2016; Varma, 2018).

In part, the underrepresentation of women in engineering is linked to girls' early experiences in STEM, during which youth form gendered assumptions about STEM activities and hone their interests based on the stereotypes they internalize (Meiksins et al., 2017). Middle school and the years immediately preceding are a particularly fitting time to intervene in STEM activities because this is when girls generally begin to lose interest in science and math (American Association of University Women, 2010). However, one potential barrier to entry for engineering for young girls in out-of-school settings is that engineering activity design has typically focused on robotics and electronics experiences that draw on a history of predominantly white male hackerer and tinkerer cultures (Buechley, 2013; Halverson & Sheridan, 2014; Reisslein, Ozogul, Johnson, Bishop, Harvey, & Reisslein, 2013), thus, perpetuating similar inequities that exist in the broader engineering field (Buchholz, Shively, Peppler, & Wohlwend, 2014; Buechley, Peppler, Eisenberg, & Kafai, 2013; Dawson, Seakins, Archer, Calabrese-Barton, & Dierking, 2015). Yet, there is little research evidence to guide effective alternative activity designs that are more inclusive and ensure that a broader spectrum of learners in out-of-school settings are engaged with the core goals of engineering activities.

Sciences museums play a particularly interesting role within larger efforts toward inclusive engineering as these out-of-school settings serve a broad range of visitors (Vossoughi & Bevan, 2014). Given the trend of replicating the same practices that have kept some—particularly girls—at the margins, educators and researchers interested in broadening participation for youth in engineering fields within museum settings are presented with the design challenge of how to generate early design recommendations for equitably engaging all visitors in engineering and making activities.

A promise for inclusive engineering activities: Cultivating empathy through narrative

To address the gender gap and to spark girls' early interest in engineering, promising research has suggested that girls' engagement in engineering may be supported by contextualizing engineering problems in relation to personally meaningful contexts, people, and communities in which learners are invested (Bennett, 2000; Dorie & Cardella, 2013; Eccles, 2005; Eccles & Wang, 2015; Wigfield & Eccles, 2000). Additionally, supporting the development of empathy within engineering activities may be valuable for inviting girls into engineering because demonstrating empathy is an embedded part of what it means to be an engineer (Walther, Miller, & Sochacka,

2017)—indeed, designing for others is central to the problem solving process (Engineering Accreditation Commission, 2015; Walther, Miller, & Keller, 2012).

Aligned with this idea, narrative can be a powerful tool to engage girls in engineering design problems (Pruit & Adlin, 2006) through connecting with stories (Putnam, 2010) and increasing empathy. For example, engineering activities that incorporate narrative elements can cultivate perspective-taking and promise to invite girls in particular to generate engineering solutions matched to stakeholders' needs (Bennett, 2000; Bennett & Monahan, 2013; Bennett, Monahan, & Honey, 2016; Buchholz, Shively, Peppler, & Wohlwend, 2014; Dusold, 2008). Carefully selected narrative design elements may be particularly suited to inviting girls to recognize and seek to respond to stakeholders' problems, needs, and constraints and thus develop empathy and perspective-taking within activities (Bennett, 2000; Bennett, Monahan, & Honey, 2016; Buchholz, Shively, Peppler, & Wohlwend, 2014; Walther, Miller, & Sochacka, 2016). In turn, this can support girls' persistence within activities as well as the critical ideation and iteration phases of engineering design (Atman, Adams, Cardella, Turns, Mosborg, & Saleem, 2007).

Thus, understanding specifically how to integrate narrative elements to evoke empathy in engineering design activities promises to support the design of more inclusive learning spaces that invite girls to engage more readily in engineering activities. While little work has attended to how to intentionally design for cultivating empathy around personally meaningful problems in engineering contexts, the present project brings this design challenge into focus.

II. Project goals and research questions

The present study is broadly concerned with understanding if and how narrative design elements impacted girls' engagement with engineering across different museum activities and settings. Within activities, the research is also specifically concerned with understanding if and how the activity design impacted girls' engagement with narrative and cultivated markers of empathy. Empathy was of particular interest as an outcome because of its promise for inviting girls into the engineering field by tapping into how to design for personally meaningful contexts, problems, and people. To study these related goals involving the intertwined relationship between narrative, engineering, empathy, two conditions (i.e., initially defined as narrative and non-narrative) of three select engineering design activities were observed within one museum, and one activity was implemented across two additional museum sites.

Table 1. Key terms

Key Term	Explanation	Example from the museum floor
Activity	<i>Engineering design exhibits implemented at museum sites</i>	<i>Activities included Help Grandma/Invention Challenge, Chain Reaction, and Air-Powered Vehicles</i>
Guided narrative condition (formerly narrative condition)	<i>The version of the activity that called for visitors to engage with a particular narrative focused on a character or setting</i>	<i>In the Help Grandma/Invention Challenge activity, visitors created an invention for a specific grandma character; in the Air-Powered Vehicles activity, visitors created a vehicle to move across a specific setting such as a desert landscape</i>
Visitor-generated narrative condition (formerly non-narrative condition)	<i>The version of the activity that did not call for visitors to engage with a particular narrative, yet was not void of narrative elements</i>	<i>In the Help Grandma/Invention Challenge activity, visitors created an invention to solve a social or real-world problem; in the Air-Powered Vehicles activity, visitors created a vehicle to move across different textured surfaces</i>
Narrative practices	<i>Practices with which visitors engaged linked to narrative-related elements of activity</i>	<i>Narrative practices included referencing narrative, elaborating narrative, inventing narrative, and inventing user</i> Inventing a user: As Rica makes her vehicle, she explains, "It's going to be something for a robot, it brings stuff to you"
Engineering practices	<i>Practices with which visitors engaged that linked to engineering-related elements of activity</i>	<i>Engineering practices included imagination, iteration, persistence, problem scoping, solution finding, testing, and tinkering</i> Iteration: While making something for grandma, Laura explains: "I've got another idea to keep it straight" and begins working on new concept
Empathy markers	<i>Markers that indicated visitors expressed empathy through how they engaged with the activity</i>	<i>Empathy markers included affective (user and designer), desire to help, familiarity, perspective-taking, societal issue, and UCD criteria</i> Affective (user): Mia explains, "The dog feels lonely"
Dwell time	<i>The length of time visitors engaged with an activity¹</i>	<i>Visitors were timed from the moment they entered the activity space until they left; a visitor's total dwell time was cumulative if a visitor re-entered the space at a later time while evaluation team was still observing</i>

¹On average, visitors spend about 1 minute with science museum activities (Dancstep & Gutwill, 2019). Longer dwell times can thus be indicators that the activity presents a rich context for evaluation research.)

The present study brought to light the idea that any engineering design activity can prompt visitors to engage with narrative elements. Thus, from here forward the so-called narrative and non-narrative conditions are referred to as guided narrative (for the narrative condition), in which specific narrative elements were explicitly embedded in the activity design, and visitor-generated narrative (for the non-narrative condition), in which specific narrative elements were not part of the activity design but where activity design invited the possibility for visitors to craft their own narratives. See Table 1 for explanations and examples of key terms.

Overall, the goals of the evaluation research were to observe 1) whether and how the design of activities in the guided and visitor-generated narrative conditions influenced the engagement of 7 to 14-year-old girls with consequential engineering practices as well as 2) whether and how activities influenced the presence of narrative practices and empathy markers and if so, how those target constructs related to engineering practices. The findings in the present report also support the larger project subgoals, which were to:

- Build a conceptual model of how narrative elements can be integrated into engineering design activities in ways that support empathy markers that aligns with and complements the core engineering emphases of the activities.
- Build practice-based knowledge that can guide the design of narratively-framed engineering design activities to cultivate empathy and encourage the engagement and persistence of girls with the engineering design process.

The project focused on the following evaluation research questions and sub-questions:

A closer look at narrative

- 1) How do narrative design elements influence engagement with narrative practices for 7 to 14-year-old girls across conditions, activities, and museums?
 - a) To what extent does the condition type, activity type, or museum impact the diversity of narrative practices that visitors demonstrate?
 - b) How correlated are the observed narrative practices across all visitors?

A closer look at engineering

- 2) How do narrative design elements influence engagement with engineering practices for 7 to 14-year-old girls across conditions, activities, and museums?
 - a) To what extent does the condition type, activity type, or museum impact the diversity of engineering practices that visitors demonstrate?
 - b) What is the relationship between dwell time and the diversity of engineering practices?
 - c) How correlated are the observed engineering practices across all visitors?

A closer look at empathy

- 3) How do narrative design elements influence engagement with empathy markers for 7 to 14-year-old girls across conditions, activities, and museums?
 - a) To what extent does the condition type, activity type, or museum impact the diversity of empathy markers demonstrated by visitors?
 - b) How correlated are the empathy markers across all visitors?

Diving into engineering and empathy

- 4) What was the relationship between engineering practices and empathy markers?
 - a) How does dwell time relate to the diversity of engineering practices and the diversity of empathy markers?

By addressing the evaluation research questions and sub-questions, the evaluation research aimed to produce 1) evidence regarding the feasibility and appeal of the engineering design activities to a broad audience; 2) evidence regarding the impact of engineering design activities on girls' engagement with engineering practices and the engineering design process; 3) an evidence-based framework for understanding the role of narrative elements and empathy in

shaping visitor participation within drop-in engineering design activities; and 4) evidence-based guidance for the facilitation of both guided narrative and visitor-generated narrative drop-in engineering design activities in science museums. Table 2 presents an overview of all evaluation research goals.

Table 2. Summary of evaluation research goals

Summary of Goals
• Understand whether and how narrative-based elements influence engagement in engineering activities for 7 to 14-year-old girls across three engineering design activities. • Highlight whether and how museum settings influence engagement with engineering practices for 7 to 14-year-old girls. • Investigate the relationship between engineering practices, narrative practices, and empathy markers for 7 to 14-year-old girls within the designed engineering activities. • Research how activity designs influence girls' narrative practices, engineering practices, and empathy markers across activities and museums

III. Museum settings and activities

The New York Hall of Science (NYSCI), The Tech Interactive (The Tech), and Scott Family Amazeum (Amazeum) were museum partner settings for the evaluation research. NYSCI served as the primary hub for activity design and iterations. Activities were piloted and refined at NYSCI before they were implemented at The Tech and Amazeum. The evaluation research focused on studying two activities within the guided and the visitor-generated narrative conditions at NYSCI and one activity designed for both conditions across all three museum settings.

All three museum settings include workshop spaces that focus on designing interactive science and engineering activities for the visitors they serve. The primary activity design site for the present evaluation research, NYSCI, is a hands-on science center in Queens, New York that serves a highly diverse population. The Tech Interactive is a family-friendly science center located in San Jose, California focused on creating unique design challenges and experiences. The Amazeum in Bentonville, Arkansas is a hands-on science museum that aims to make connections to the local Arkansas community through activity and exhibit design. In this section, each museum is described in relation to a number of relevant points of alignment and difference, including the communities in which they are embedded, the flow of visitors through the respective museums, and the overall principles that guide facilitation approaches. See Table 3 for an overview of the museums.

Table 3. Overview of museum settings

	NYSCI	The Tech	Amazeum
Visitors (annually 2018)	Approximately 500K	Approximately 500K	Approximately 250K
Demographics	Serves a racially and ethnically diverse population, including a large local immigrant population	Serves surrounding community, notably immersed in Silicon Valley tech culture	Serves the local Bentonville community, including many local homeschoolers
Facilitation approach	Design-Make-Play approach (Honey & Kanter, 2012) to guide “explainers” who facilitate with freedom regarding available materials and facilitation style	High energy facilitation style, promoting spirited competition through design challenges; 1-on-1 facilitation emphasized when possible within activities	Activities are designed to specifically engage visitors with topics relevant to the local Arkansas industries and making connections to employment opportunities
Museum space	Activities occur in a 10,000 square foot <i>Design Lab</i> Space that includes separate, partially enclosed activity hubs on the lower floor	Activities occur across large three floor building within separate activity hubs	Activities occur in the 50,000 square foot 3M Tinkering Hub that uses repurposed materials and displays past youth projects

The New York Hall of Science | Queens, NY



The New York Hall of Science (NYSCI) is an interactive, hands-on science museum that is committed to long-term collaboration with schools and community organizations. Additionally, a close connection to the local community is an integral part of how the museum operates, key to both its programming and facilitation. NYSCI is located in Queens, New York, the most diverse county in the United States (Lobo, Salvo, & Alvarez, 2013) and is embedded in the neighborhood of Corona, a bustling community mainly comprised of newly arrived immigrants from around the world, most prominently from Central and South America. In 2018, about 500,000 teachers, parents, and youth visited NYSCI.

Museum professionals and researchers at NYSCI design activities that are aligned with theories of learning that emphasize building from learners' prior knowledge and experiences in the world (e.g., Papert, 1980). By building on what young people bring with them to science learning experiences, NYSCI hopes to inspire active participation—rather than passive observation—in STEM learning. Examples of this include biology demonstrations, in which visitors are asked to shout hypotheses to a facilitator, a roller coaster wall, in which visitors mix and match wooden and magnetic pieces to accelerate a ball, or rotating activities, such as, crafting dresses for wooden drawing mannequins. For their facilitation practices, NYSCI uses the Design-Make-Play (Honey & Kanter, 2012) approach across all museum activities and experiences to support young people in making and creating new information. The Design-Make-Play approach includes five core principles to support NYSCI's specific strategies of engagement: 1) People and play at the center; 2) kids as creators; 3) problems you think are worth solving; 4) divergent solutions; 5) open invitation (for more detailed information on the Design-Make-Play approach, see Honey & Kanter, 2012).

Activity development, implementation, and iteration for the present evaluation research took place in NYSCI's Design Lab, a 10,000 square foot teaching and learning space built in 2014 dedicated to offering visitors hands on design and engineering experiences. Design Lab includes five distinct, yet open, activity hubs to facilitate activities simultaneously throughout the lower floor of the museum. The facilitation style at NYSCI is aptly described as minimalist, focused on supporting visitors' tinkering and encouraging open exploration with materials. Facilitators are dubbed "Explainers," and wear red aprons so they can be easily identified within the activity spaces in the Design Lab. NYSCI's Explainers are mostly local high school and college students. Although they receive training in the Design-Make-Play approach used in the Design Lab and are offered general guidelines for supporting the activity design (e.g., how to set up the activity, which materials to include, how to engage with visitors and when), Explainers have a fair amount of freedom in how they operate and interact with visitors within activities. Additionally, Explainers rotate across activities within the Design Lab so within each shift, they get experience facilitating a range of different activities. Explainers also have agency to adjust each activity design a bit to align with their personal styles, including the exact materials made available to visitors and how they choose to introduce the activities when visitors enter the activity spaces.

The Tech Interactive | San Jose, CA



The Tech Interactive (The Tech) is a hands-on science and technology museum in downtown San Jose, California driven by a mission to inspire the innovator in everyone. As part of a wide range of programs, The Tech facilitates a popular event called The Tech Challenge, in which youth engage in design processes that require collaborative problem solving. Regular floor

programs aim to support visitors' design processes and decenter competition among and across groups. Additionally, through The Tech Awards, the museum honors community members who have done work to benefit humanity. The Tech is a community resource that has more recently shifted their focus to developing programming that emphasizes getting low-income students and girls engaged in engineering and other STEM learning activities. The Tech is deeply rooted in the history of Silicon Valley and its spirit of innovation, welcoming approximately half a million visitors annually.

Activities are set up in thematic areas that visitors explicitly enter to engage with and focus on engineering design and innovating with technology, at times through scheduled programming. For example, The Tech offers programs on bio design, including artificial intelligence augmented tangible design of fantasy creatures and a virtual reality augmented Body Worlds activities, as well as spaces that focus on tinkering with a range of low- and high-technology, including robotics and mechanical designs. Facilitators actively engage with visitors through high-energy strategies, including recruiting visitors to join the activity, asking questions, and supporting visitors to overcome challenges in their designs. Similar to NYSCI, explainers rotate across activities throughout the day.

Scott Family Museum Amazeum | Bentonville, AR



The Scott Family Museum Amazeum (Amazeum) is an interactive science museum in Bentonville, Arkansas that engages visitors with themes across the local community, including the industries that are integral to the cultural fabric of Arkansas and especially the area around Bentonville. The Amazeum is committed to impacting the community through creating active and vibrant STEAM-focused museum experiences. Activities that highlight this include a space that invites visitors to paint glass walls with watering colors as well as the Hershey's Lab, which combines learning of the science of commercial chocolate making with playful and expressive activities. Additionally, the Amazeum provides spaces for young people to create and test their ideas, including large branded grocery store playshop that includes a bakery, a vegetable vendor, and a café as well as a life-size semi-truck and agricultural assembly lines that visitors get to operate and engage with. To support a fun and playful approach to science learning, random happenings often occur at Amazeum such as a remote-controlled giraffe riding on a unicycle across the museum floor. The Amazeum also provides various camps, programs, events, and other educational resources for visitors. Event highlights include its Maker Fashion celebration and Tinkerfest, the largest festival of its kind in Arkansas, which brings together makers, artists, and tinkerers of all ages to showcase design work and celebrates the creative, curious, and innovative spirit in everyone. Amazeum has a commitment to serving the areas surrounding the museum, including the local Hispanic and Marshallese communities. In addition to Arkansas, visitors from across the United States travel to visit the Amazeum.

Engineering design activities place in The Amazeum's 3M Tinkering Hub, a 50,000 square foot activity and learning space, which invites visitors to engage in hands on science and crafting activities connected to activities currently on display. In the 3M Tinkering Hub visitors use provided materials and their own ideas to create artifacts inspired by the designed activities. Activities often focus on topics relevant to the local Arkansas community, including agricultural concerns and topics supported by commercial sponsors like Walmart. Facilitation within this space is just on time, with museum staff supporting visitors by pointing out materials and asking targeted questions. Within the 3M Tinkering Hub, materials are stored in cardboard storage bins donated from local shops and visitor projects are displayed across the room to inspire making and building on prior work. Generally, visitors are encouraged to leave their designs behind in the space to support repurposing of materials.

Introducing the engineering activities

Activity designs were informed by tenets of constructionism, following the tradition that materials and activities ought to be thoughtfully designed to support engagement and the development of personal relationships within domain learning (Harel & Papert, 1991). By centering the value of relationships and personally meaningful problems in design, the needs of others were an embedded part of creating user-centered design experiences (Nelson & Stolterman, 2003). Additionally, through their designs, museum partners at NYSCI aimed to incorporate narrative elements within each activity to provide a personally meaningful context for engineering problems (i.e., in the guided narrative condition, specific characters and settings were included).

Activities included: 1) Help Grandma/Invention Challenge in which visitors used a diverse assortment of random materials to design an invention to “help grandma” with a variety of tasks from opening a jar to hearing the doorbell, in the guided narrative condition, or to design an invention to solve a real world problem in the visitor-generated narrative condition. 2) Chain Reaction, in which visitors created Rube Goldberg machines using everyday objects to help a dog that contained circuit components in a targeted way such as by feeding it or playing with it in order to trigger a response (i.e., wagging its tail) in the guided narrative condition, or, in the case of the visitor-generated narrative condition, to invent a contraption that created a “chain reaction” to reach a goal such as ring a bell or get a ball in a cup; and 3) Air-Powered Vehicles in which visitors designed vehicles to travel over a variety of terrains, including specific landscapes in the guided narrative condition and different textures in the visitor-generated narrative condition.

In guided and visitor-generated narrative versions of activities were used to study engineering engagement for female (age 7-14) visitors. The three activities for evaluation (i.e., Help Grandma/Invention Challenge; Chain Reaction; Air-Powered Vehicles) were selected from initial design and implementation of guided and visitor-generated narrative versions of six different engineering activities. The three activities that were studied were chosen from the total of six activities because they were activities that all museum partners agreed on as being possible to implement across all three contexts.

Help Grandma/Invention Challenge

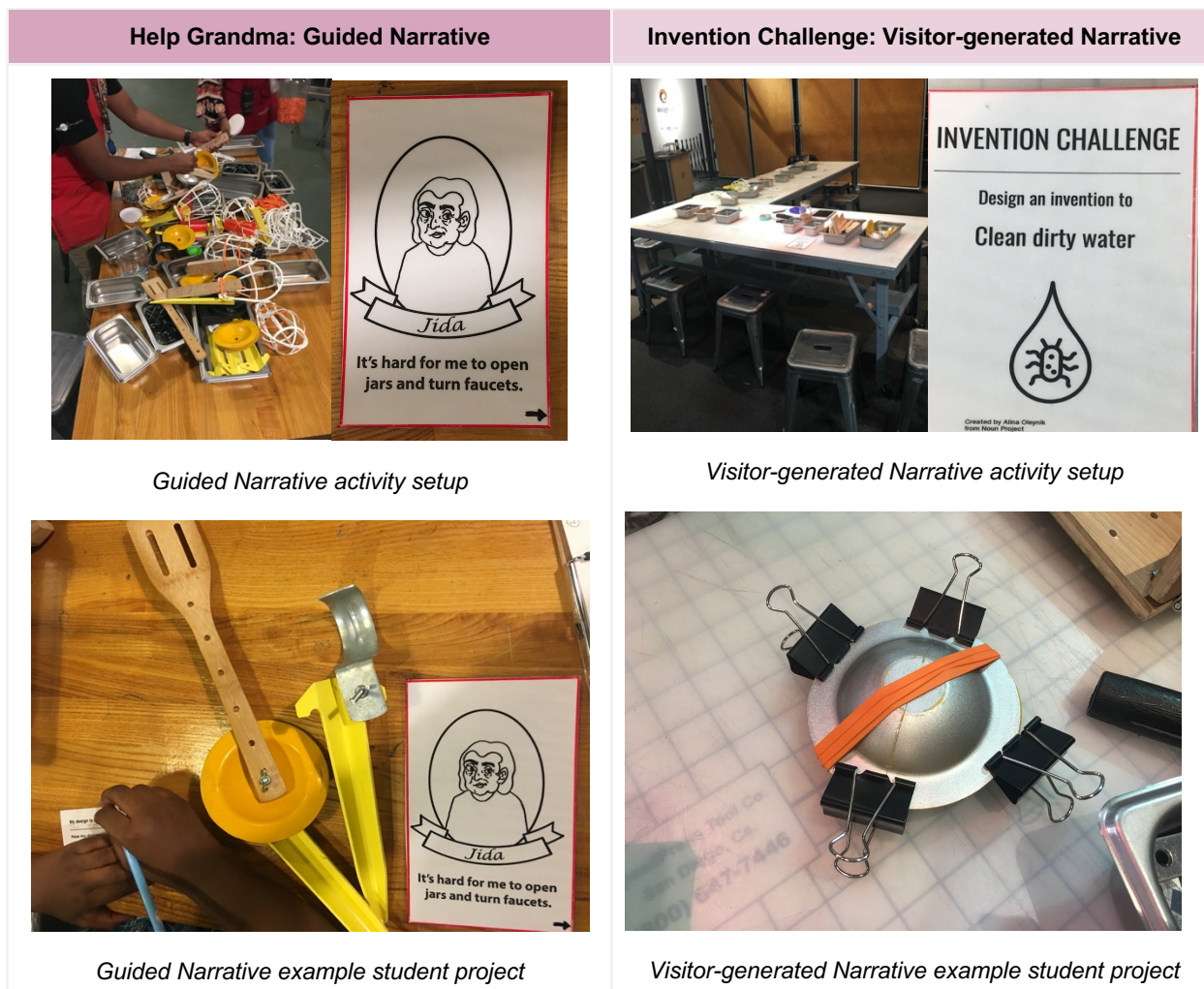


Figure 1. Guided narrative condition (Help Grandma, left) and the visitor-generated narrative condition (Invention Challenge, right) of the Help Grandma/Invention Challenge activity.

Help Grandma/Invention Challenge was an activity implemented only at NYSCI (not across museum settings). The purpose of the Help Grandma/Invention Challenge activity was to use available materials to design an invention to help with a proposed problem. In the guided narrative version of the activity, visitors were tasked with designing an invention to “Help Grandma” with a variety of tasks such as opening a jar, crossing the street, or hearing a doorbell (see Figure 1, left). In this guided narrative condition, tasks were written on activity cards that included a cartoon drawing of a Grandma and an explanation written from her point of view. For example, one card read, “I’m always losing my glasses!” with a sketch of “Nonna” on the front. The back of the card explained the details of the task.

In the visitor-generated narrative version (see Figure 1, right) of the activity visitors were to design an invention to solve a real-world problem. Like in the guided narrative version, activity card prompts offered guidance for engaging in the activity, but in the case of the visitor-generated narrative condition, prompts were not tailored around designing for a specific user. Example prompts included “design an invention to make snow” and “design an invention to help clean dirty water.” While the guided narrative condition included a human user, the visitor-generated narrative version focused on designing for a general problem in the world.

Both conditions the guided narrative and the visitor-generated narrative condition included documentation cards that prompted visitors to name their invention, to draw a sketch of their invention, and to capture how the invention was intended to work. Visitors were then invited to display their inventions along with these documentation card on a sideboard for other visitors to see, get inspired by, and build on. The materials for both guided and visitor-generated narrative conditions included plastic lacrosse baskets, rubber bands, wooden pieces with perforated holes (e.g., spatulas with holes along the grip), screws, nuts, plastic pieces (e.g., yellow bowls with holes at the center), binder clippers, and long wooden skewers. This made it possible for visitors to connect materials through hinging and twisting mechanisms.

Chain Reaction

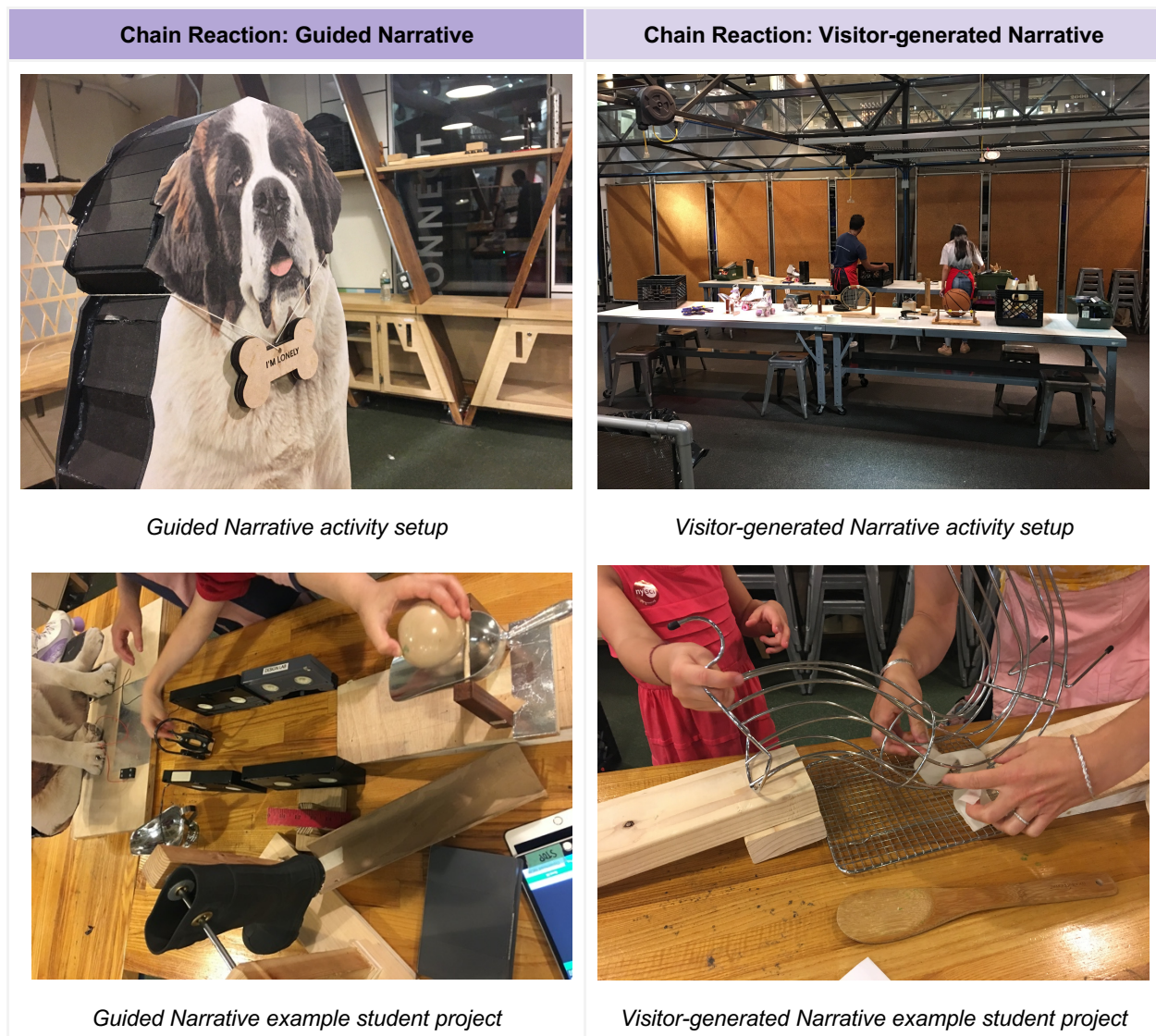


Figure 2. Guided narrative condition (left) and visitor-generated narrative condition of the Chain Reaction activity (right)

Chain Reaction was a second activity that was only implemented at NYSCI. The activity also took place in the Design Lab space and Explainers generally followed the same facilitation

style as the Help Grandma/Invention Challenge activity. In the guided narrative version of Chain Reaction (see Figure 2, left) visitors were prompted to design a chain reaction contraption that could help take care of a pet, diverging from the human user central to the Help Grandma/Invention Challenge activity, especially the guided narrative condition with the grandmother characters. Visitors were prompted to design for a cardboard cutout of a realistic looking dog or cat that were part of the activity. In the visitor-generated narrative version of Chain Reaction (see Figure 2, right) visitors were to create a chain reaction contraption to reach a goal (e.g., ring a bell, get a ball in a cup).

Apart from the inclusion of the animal characters within the guided narrative condition, the materials across both conditions were the same. The activities included VHS tapes, ramp like structures, wooden blocks, balls of different sizes and weights, kitchen utensils, such as saucers and spatulas, as well as materials that could set things in motion, including a hammer and a boot that was connected to a bar to make it swing, as well as roller skates. Typically, at the start of the activity the contraption building materials were set up on the sideboard so that visitors could collect the materials they wanted to use in their designs without cluttering the tables. However, explainers often set up example contraptions that they could use to demonstrate the activity and for visitors to build onto. To make it possible for visitors to build long contraptions that could flow around corners, tables were set up in L-shapes. This was also intended to encourage idea sharing across visitor groups.

Air-Powered Vehicles

Air-Powered Vehicles was implemented across the three museum sites. Across all museums, the guided narrative condition of the activity invited visitors to design an air-powered vehicle to help them travel around the world across different landscapes, such as a desert with sand, a grassland with plant obstacles, and a tundra with cracks and slippery spots. In the visitor-generated narrative condition, visitors were asked to design an air-powered vehicle that could move across different textured surfaces, including smooth plastic, a grid where wheels could get caught, a mattress, and resistant extra-thick carpet. Across the three museum sites, there were notable differences with respect to facilitation styles, space designs, available materials, and activity set ups, some of which we highlight in the present report.

Air-Powered Vehicles at NYSCI. Figure 3 shows the guided narrative condition (left) of the Air-Powered Vehicles activity (e.g., grassland) and the visitor-generated narrative condition (right) with a grid textured surface. After building their air-powered vehicles, visitors used large fans to simulate air flow that would push the vehicles forward. Visitors could operate the fans on their own across three settings (low, medium, and high air flow) or ask an explainer for support. At NYSCI, visitors were provided with a variety of materials to use to create their vehicles, including bobbins, wooden skewers, rubber bands, small pieces of cloth, mylar blankets, binder clippers, wooden boards with holes, and foam pieces.

At NYSCI, the guided narrative version of the activity was in a space that included a large pole in the center, naturally separating the space into two workstations with tables that included a range of materials that were neatly arranged in separate bins (see Figure 4, left). Tables were also accompanied by one landscape each. On the first day of our observations the guided narrative condition of the Air-Powered Vehicles activity included the desert landscape and the tundra. On the second day, the desert was swapped with the grassland, but the tundra remained. The workstation next to the tundra was slightly smaller and fit fewer visitors than the other workstation. Researchers aimed to spend approximately equal amounts of time observing visitors near each of the landscapes.



Figure 3. Set-up of guided narrative (left) and visitor-generated narrative (right) condition of Air-Powered Vehicles activity at NYSCI

The visitor-generated narrative condition was facilitated in an enclosed space as well, however, it was larger than the space where the guided narrative condition occurred. The space included two larger workstations with chairs and a similar range of materials as in the guided narrative condition, also neatly arranged in bins (see Figure 4, right). The visitor-generated narrative added foam blocks, which were not part of the activity facilitation in the guided narrative condition. The visitor-generated narrative set-up included four surfaces, 1) a mattress, 2) a shiny smooth surface, 3) a grid, and 4) a thick carpet. The space included a ceiling-tall shelf as well as shelves alongside the ceiling that held additional materials and displayed quirky projects (including a paper-folded bust of a googly-eyed dinosaur).



Figure 4. Air-Powered Vehicles activity at NYSCI included guided narrative materials (left) and visitor-generated narrative materials (right)

In both conditions at NYSCI, visitors generally sat down on chairs while engaging with the activities. The activity set-up included extra chairs to accommodate larger groups of visitors. Across conditions, the landscapes/surfaces were set up on shelves that were an integrated part of the space design. The guided narrative condition also included a mobile table on wheels for one of the landscapes.

Air-Powered Vehicles at The Tech Interactive. At The Tech, visitors created air-powered vehicles using two broad types of materials: 1) building materials, such as paper, food baskets, and straws, and 2) connecting materials, such as colorful hooks, binder clippers, and rubber bands. Figure 5 shows the different activity set ups for the guided narrative condition of the Air-Powered Vehicles activity (e.g., desert) and the visitor-generated narrative condition (e.g., smooth

surface). Three pointed differences between The Tech and the set-up of the same activity at NYSCI were 1) the use of color in materials, 2) how materials were organized and separated, and 3) the presence of printed backdrops in the guided narrative condition. Another strong difference from NYSCI was that at The Tech, visitors were generally discouraged from operating the fans on their own; facilitators were instructed to support visitors in testing their inventions across landscape surfaces.



Figure 5. Set-up of guided narrative (left) and visitor-generated narrative (right) conditions of Air-Powered Vehicles at The Tech Interactive

At The Tech, we observed the guided and the visitor-generated narrative conditions of the Air-Powered Vehicles activity in the same space. To elaborate on select differences highlighted above, a main difference between conditions was the use of backdrops and different landscape surfaces in the guided narrative condition. Across conditions, materials were divided into building materials and connecting materials. Building materials (Figure 6, left) included, thin paper, food baskets, thin and thick paper and plastic straws, foam disks, popsicle sticks, wooden and plastic bobbins, CDs, fabric squares, paper pipes, and plastic balls. The building materials were neatly separated in blue boxes and made available to all visitors on a table that was positioned in the center of the room. Connecting materials (Figure 6, right) were available in a box with several compartments at each of the 2-3 workstations where visitor groups were invited to build. Connecting materials included plastic clothespins, colorful hooks, binder clips, paper clips, pipe cleaners, and rubber bands. Each workstation also included chairs that invited visitors to sit down. In the guided narrative, two copies of three landscapes (i.e., desert, grassland, and tundra) were distributed at each side of the room (apart from the entrance) and in the visitor-generated narrative condition, the surfaces were lined up on a wall covered in mural of a skyline. These surfaces included thick carpet, rough acrylic glass, and smooth acrylic glass. The space also included a sofa and two sofa chairs, intended as a place where group members could rest and recharge.



Figure 6. Air-Powered Vehicles at The Tech Interactive included building materials (left) and connecting materials (right) for both conditions

Air-Powered Vehicles at Amazeum. At Amazeum, visitors' air-powered vehicles were created in the 3M Tinkering Hub that was located close to the museum workshop and that is typically only accessible by youth 7 years and older. Visitors were invited to work with a range of materials that were on display in a rotating material bin as well as a paper shelf that was donated to the museum by a retail business. Figure 7 shows the guided narrative condition (e.g., tundra) and the visitor-generated narrative condition (e.g., thick carpet) of the Air-Powered Vehicles activity as set-up at Amazeum.



Figure 7. Set-up of guided narrative (left) and visitor-generated narrative (right) condition of Air-Powered Vehicles at Amazeum

Amazeum staff set up organically curved tables as one long workstation where visitors could use available materials to create their inventions. The workstation was equipped with a range of cutting and piercing tools, including scissors and holepunches. Other pieces of this table puzzle were arranged on the left side of the room along the wall and were used to hold landscapes and surfaces as well as fans that visitors could use to test their inventions. The same set-up was used for both the guided narrative and the visitor-generated narrative conditions of the Air-Powered Vehicles activity at Amazeum. The guided narrative condition featured a tundra that included a 3D printed ice bear, a grassland with artificial plants, and a desert landscape. The visitor-generated narrative version used different surfaces including a thick carpet, a grid, a mattress, and a smooth, shiny surface. For both conditions at the Amazeum, the surfaces and landscapes used in the activities were similar to those used at NYSCI apart from the addition of the icebear figure for the tundra surface at NYSCI.



Figure 8. Available materials for Air-Powered Vehicles activity at Amazeum across both guided and visitor-generated narrative conditions

Within the room, a laser cutter and other high-tech tools were on display alongside low-tech materials including fabrics, cardboard, and felt. Additionally, prior projects by visitors who had engaged in other activities in the space were on display on shelves, benches, and walls. The Amazeum had an explicit rule that visitors could not take home any projects created in the workshop setting. Aligned with this rule, projects that visitors created as part of the Air-Powered Vehicle activity were displayed at the entrance of the activity for visitors to view and get inspired by. Both of the material containers were freely accessible by the youth. The rotating storage container (Figure 8, left) included corks, straws, popsicle sticks, wooden clothespins, rubber bands, and plastic bobbins. The paper retail shelf (Figure 8, right) stored cut up pool noodles, plastic cups and plates, empty egg cartons, plastic balls, and paper tubes. Visitors could use as many different materials in the building process as they wanted. It is interesting to note that only a few chairs were provided for visitors to sit while designing, and therefore, most visitors stood. A large screen attached to one of the walls of the workshop environment featured videos that facilitators had selected as relevant to the activity, including a video that showed adult males speeding on a range of water vehicles through Florida's marshes.

IV. Scope of research

Activity design and innovation at NYSCI. All activities were implemented at NYSCI with NYSCI acting as the lead activity designer and innovator throughout the study. Air-Powered Vehicles was implemented at both The Tech and Amazeum after it was piloted on the NYSCI museum floor. Differences in results across sites might be attributed to unique museum cultures as well as how the design of activities and observation protocols evolved over time given that new settings create new possibility spaces. It might actually be expected that activities as well as the protocol iterations would get better over time with each subsequent iteration.

In their pilot testing, NYSCI initially played with how narrative elements were embedded in designs in both guided and visitor-generated conditions. Ultimately, a choice was made to embed the narrative element of character within the guided narrative versions of the Help Grandma/Invention Challenge and Chain Reaction activities, while the narrative element of setting was embedded within the guided version of the Air-Powered Vehicles activity. In the case of the Help Grandma/Invention Challenge guided narrative condition, visitors were tasked with designing for a person (i.e., a grandma), and in the case of Chain Reaction, visitors designed something for a pet (i.e., a dog). For Air-Powered Vehicles, the narrative element of setting defined as different landscapes (e.g., desert, grassland, tundra) was embedded in the activity design for the guided narrative condition. The visitor-generated narrative conditions were more open-ended, yet still invited visitors to craft their own narratives.

Observations. A total of 138 observations of girls were conducted across the three activities at NYSCI. The breakdown of observations by condition and activity is included in Table 4. For Air-Powered Vehicles, 37 observations were conducted at The Tech and 27 observations were conducted at Amazeum (see Table 5). (Note: Both tables include Air-Powered Vehicles observations for NYSCI).

Table 4. Observations and observation hours by activity and condition at NYSCI

	Help Grandma/Invention Challenge (NYSCI)		Chain Reaction (NYSCI)		Air-Powered Vehicles (NYSCI)		TOTAL
	<i>guided</i>	<i>visitor-generated</i>	<i>guided</i>	<i>visitor-generated</i>	<i>guided</i>	<i>visitor-generated</i>	
Observations (n)	19	19	41 ²	19	20	20	138 obs.
Total observation hours	11.5	10	15	8	5	6	55.5 hours
TOTAL	38 obs. / 21.5 hours		60 obs. / 23 hours		40 obs. / 11 hours		

Table 5. Number and hours of observation hours³ by museum and condition for Air-Powered Vehicles (APV)

	NYSCI (APV)		The Tech (APV)		Amazeum (APV)		TOTAL
	<i>guided</i>	<i>visitor-generated</i>	<i>guided</i>	<i>visitor-generated</i>	<i>guided</i>	<i>visitor-generated</i>	
Observations (n)	20	20	19	18	13	14	104 obs.
Total observation hours	5	6	6	6	7	9	39 hours
TOTAL	40 obs. / 11 hours		37 obs. / 12 hours		27 obs. / 16 hours		

² 21 additional observations for the Chain Reaction activity were conducted because initial results encouraged further investigation of a particular design feature of the activity (i.e., the presence of a dog collar).

³ Observation hours are inclusive of a short meet-and-greet between the evaluation researcher and museum professionals at respective museum sites. Meet-and-greet sessions included a brief tour of the museum, orientation with the activities, and introduction to facilitators.

Evaluation research aimed to understand broadly how narrative, engineering, and empathy were related across the designed activities, conditions, and museum settings. The evaluation research was organized as follows: For Help Grandma/Invention Challenge and Chain Reaction, analysis looked at data for NYSCI only because NYSCI was the only museum to implement these two activities. For Air-Powered Vehicles, research looked at data across NYSCI, The Tech, and Amazeum because all three museum sites implemented this one activity in the two conditions. Analyses looked at what was happening across all three activities at NYSCI, what was happening with Air-Powered vehicles across all museums, and what was happening when all data were compiled together.

Overall, the evaluation research was concerned with making sense of narrative practices, engineering practices, and empathy markers across activities, conditions, and museum settings. To do so, the research isolated each construct of interest (i.e., narrative, engineering, and empathy) to understand how visitors' overall engagement in practices/markers within each construct was influenced by narrative design elements. Additionally, the research examined (a) how the condition, activity, and museum impacted the practices/markers demonstrated; (b) how dwell time and the practices/markers were related; and (c) how the practices/markers correlated with one another across observations. Finally, the research focused on understanding the relationship between engineering and empathy within the designed activities.

Overarching research questions

- 1) How do narrative design elements influence engagement with narrative practices for 7 to 14-year-old girls across conditions, activities, and museums?
- 2) How do narrative design elements influence engagement with engineering practices for 7 to 14-year-old girls across conditions, activities, and museums?
- 3) How do narrative design elements influence engagement with empathy markers for 7 to 14-year-old girls across conditions, activities, and museums?
- 4) What was the relationship between engineering practices and empathy markers?

V. Methods

Research instruments

Research instruments included an **observation protocol** and a **semi-structured, informal interview protocol** to collect data aligned with the overall goals of evaluation research (see Appendices A and B). The observation protocol guided evaluation to mark evidence of engagement in engineering through a number of key engineering practices noted in sustained visitor observation. Engineering practices included on the observation protocol were determined by museum professionals and researchers at NYSCI. In addition to engineering practices, narrative practices, markers of empathy, and other practices with which girls engaged in the context of the activities were noted (please see Tables 6, 7, and 8 for definitions of practices/markers with aligned examples from data). Total dwell time was recorded as well as how often students had their hands on or hands-off materials. An evaluation researcher recorded qualitative notes in the observation protocols as well, including what students were doing, resonant things they said as they were working on projects, and responses to the informal interview questions. Questions included “Is there a backstory to this [what you made]?” and “How will anyone use your design?” When possible, an evaluation researcher took photographs of students’ projects in process and at the end of their engagement in the activity space to include with the observation and interview protocol notes.

Visitors

A total of 202⁴ girls ages 7-14 were observed as they engaged in the museum activities. 21 additional observations were conducted for the Chain Reaction activity to further investigate a particular design feature of the activity (i.e., the presence of a dog collar), the details of which are described in the findings.

Narrative practices, engineering practices, and empathy markers

Table 6. Practices closely connected to narrative elements of the designed activity

Narrative practices	Explanation	Example from the museum floor
Referencing narrative	References narrative in the activity prompt	Sara touches collar of dog and explains, “It’s lonely” as she begins her design
Elaborating narrative	Elaborates and extends the given activity narrative	Zahara explains that her grandfather teaches karate and does not need help opening jars
Inventing narrative	Creates own narrative based on prior knowledge or experiences	As she is making something, Traci explains that she saw an art piece that made music with rainwater so that’s her inspiration
Inventing user	Creates a user for whom the project is intended	As Rica makes her vehicle, she explains, “it’s going to be something for a robot, it brings stuff to you”

⁴ At NYSCI, observations included 38 girls in the Help Grandma/Invention Challenge activity, 60 girls in the Chain Reaction activity, and 40 in the Air-Powered Vehicles activity. At The Tech, observations included 37 in the Air-Powered Vehicles activity. At Amazeum, observations included 27 in the Air-Powered Vehicles activity.

Table 7. Practices closely connected to the engineering design process

Engineering practices	Explanation	Example from the museum floor
Imagining	Imagining new projects and possibilities	As Yesenia creates vehicle, she explains: "If I had actual wheels, it would be better"
Iteration	Improving a design function through implementing	While making something for grandma, Laura explains: "I've got another idea to keep it straight" and begins working on new concept
Persistence	Solving problems with materials	As Natalie tries to connect plastic sticks, she asks, "How do you connect this?" and tries herself three times then gathers additional materials to help
Problem scoping	Multiple aspects of a problem	London explains that her vehicle "did not work because of the wheels they would not move in the wind"
Solution finding	More than one solution (ideas)	As Ava is building her vehicle she settles on a mix of her ideas: "It's going to be a car with a sail"
Testing	Testing a design function	While building a vehicle, Maria tries different ideas and explains, "I am learning how, once I test it and it does not work I can try again and modify it, change it to make it better"
Tinkering	Elaborating, adding features	As Lucy makes a structure, it falls so she adds wood to try to make it stand

Table 8. Markers of expressing empathy within the engineering design process

Empathy marker	Explanation	Example from the museum floor
Affective (designer)	How designer feels	While making something, Bella explains, "I would like to keep adding to it and see what it turns into"
Affective (user)	How user feels	Mia references dog and explains, "The dog feels lonely"
Desire to help	Expressing a wish to help	Sonia explains about what she made: "It will probably help people with disabilities lift something up to help themselves"
Familiarity	Prior experience or knowledge	Trinity makes a connection: "My grandma has a hard time opening jars"
Perspective-taking	Talking about/acting out design	Valerie acts out how grandma would use what she made: "If it were real the person would push it up"
Societal issue	Related to societal issue	Malia makes a vehicle and explains that it's "better for the environment"
UCD criteria	User-centered design criteria	Olivia adds a "painstick" as a handle so "grandma can hold this"

Data sources & analytical approach

Data sources included completed observation protocols using the instrument described above, qualitative interview notes, recorded dwell times, and photographs of visitors' projects. After collecting data for each activity at each museum, raw data from observation protocols were organized by noting all narrative practices, engineering practices, and empathy markers per visitor and transcribing all qualitative notes in a single spreadsheet to prep for analysis. Data was organized by different configurations of museums, activities, and conditions for various parts of our analysis to address each research question and sub-question.

To explore the construct of narrative, frequencies of narrative practices were counted across conditions, activities, and museum settings. Phi correlation coefficients⁵ were calculated for narrative practices across all data and for data at just NYSCI. **To explore the construct of engineering**, frequencies of engineering practices were tallied across conditions, activities, and museum settings. A linear regression model explored whether there were any differences in the diversity of the overall levels of engineering practices by activity at NYSCI since all three activities were implemented there and a similar comparison could not be made for the other museums since only APV was implemented across all three. The relationship between dwell time and engineering practices was explored using linear regression. Phi correlation coefficients were calculated for engineering practices across all data. **To explore the construct of empathy**, frequencies of empathy markers were tallied across conditions, activities, and museum settings. Phi correlation coefficients were calculated for empathy markers across all data. Finally, the **relationship between engineering and empathy** was explored using a Spearman correlation and linear regression model.

Before the start of all observations, the evaluation researcher was guided by mutually agreed upon definitions of the practices and markers described in Tables 6-8. However, due to the nature of the evaluation research design, inter-rater reliability was not possible because there was not a video record of visitors' interactions within activities. Future work should explore a process including a video record to enable revisiting definitions and coding that could be revisited over time.

⁵ The Phi coefficient is a measure of association between two dichotomous variables. Values of Phi can range from -1 to 1, with positive values indicating a positive correlation, negative values indicating a negative or reciprocal correlation, and a value of 0 indicating no correlation. Phi may be interpreted according to the following scale: .70 - 1.00 very strong relationship; .40 - .69 strong relationship; .30 - .39 moderate relationship; .20 - .29 weak relationship; .01 to .19 no or negligible relationship; 0 no relationship.

VI. Findings

The findings below highlight relevant results aligned with constructs of interest beginning with (a) narrative, followed by (b) engineering, and finally, (c) empathy. For each construct, findings first explore the extent to which the condition type, activity type, or museum impacted the diversity of particular practices/markers that visitors demonstrated. For both engineering and empathy constructs, the relationship between dwell time and diversity of constructs is explored. (Note: There was not a clear rationale to pursue and present the relationship between narrative practices and dwell time analysis.) For all three constructs, correlations show how individual practices were related to one another across all observations of visitors. Throughout the presentation of quantitative data, qualitative cases from the museum floor illuminate analysis. The final section of the findings explores the relationship between engineering practices and empathy markers using a linear regression model to preview recommendations for museum activity design based on the present findings.

Narrative: To what extent does the condition type, activity type, or museum impact the diversity of narrative practices that visitors demonstrate?

Because the project was explicitly focused on the role of narrative as a vehicle for design in support of engineering practices for museum visitors (especially girls age 7-14), analysis considered how visitors engaged with the narrative elements of the designed activities through narrative practices. Analysis looked at narrative practices in relation to conditions (i.e., guided and visitor-generated narrative) for all data points (as well as within activities) and for Air-Powered Vehicles, across the three museum sites.

Table 9. Counts and percentages of visitors for whom each narrative practice was observed, across all activity types and museums.

	Guided narrative		Visitor-generated narrative	
	n	% obs.	n	% obs.
Total observations	112		90	
Narrative practice				
<i>Referencing narrative</i>	51	45.5%	1	1.1%
<i>Elaborating narrative</i>	16	14.3%	2	2.2%
<i>Inventing narrative</i>	14	12.5%	24	26.7%
<i>Inventing user</i>	9	8.0%	10	11.1%

To clarify Table 9, the frequencies count the presence of a particular narrative practice, however, one visitor may have demonstrated multiple markers (e.g., referencing and elaborating narrative). Across all museums, activities, and conditions, 120 visitors demonstrated no narrative practices while 82 visitors demonstrated some narrative practices. Given this visitor breakdown, of those visitors who did demonstrate narrative practices, in the guided narrative condition, frequencies of narrative practices clustered around referencing and elaborating narrative. By contrast, in the visitor-generated narrative condition, frequencies of narrative practices clustered

around inventing narratives and inventing users. For example, 45.5% of all visitors who engaged with the guided narrative condition engaged with referencing the narrative and 14.3% of participants elaborated on the provided narrative of the activities. These percentages are in contrast to the 1.1% who referenced narratives in the visitor-generated narrative and 2.2% of visitors who elaborated on narratives in the visitor-generated narrative condition. This trend was somewhat reversed for the visitor-generated narrative condition, where 26.7% of visitors invented narratives and 11.1% of visitors invented users.

However, visitors *did* invent narratives and users in the guided narrative conditions, and therefore, it can be inferred that the guided narrative condition did not preclude visitors from inventing their own narratives or users. In fact, while not a large raw number, visitors invented users at nearly the same level across both conditions. Thus, even though the guided narrative design supported visitors in incorporating a given narrative, visitors still invented their own users at the same rate as they might when no explicit narrative elements are incorporated within activities. Overall, more narrative happened in the narrative condition, which confirms the assumption that the engineering activities were designed to evoke narrative. Table 10 presents a breakdown of trends of observed narrative practices per condition.

Table 10. Counts (and percentages) of visitors for whom each narrative practice was observed within activity type (NYSCI)

	Help Grandma/Invention Challenge N (% of obs.)			Chain Reaction N (% of obs.)			Air-Powered Vehicles N (% of obs.)		
	guided	visitor-generated	both	guided	visitor-generated	both	guided	visitor-generated	both
Total	19	19	38	41	19	60	20	20	40
Narrative practice									
<i>Referencing narrative</i>	17 (89.5%)	0 (0.0%)	17 (44.7%)	13 (31.7%)	0 (0.0%)	13 (21.7%)	9 (45.0%)	0 (0.0%)	9 (22.5%)
<i>Elaborating narrative</i>	7 (36.8%)	1 (5.3%)	8 (21.1%)	3 (7.3%)	0 (0.0%)	3 (5.0%)	1 (5.0%)	0 (0.0%)	1 (2.5%)
<i>Inventing narrative</i>	3 (15.8%)	7 (36.8%)	10 (26.3%)	5 (12.2%)	5 (26.3%)	10 (16.7%)	3 (15.0%)	4 (20.0%)	7 (17.5%)
<i>Inventing user</i>	4 (21.1%)	2 (10.5%)	6 (15.8%)	1 (2.4%)	4 (21.1%)	5 (8.3%)	1 (5.0%)	1 (5.0%)	2 (5.0%)

When broken down by activity and, in the case of Air-Powered Vehicles also across museums, this trend generally persisted with small variances (see Tables 10 and 11). Visitors who engaged with the guided narrative condition of the Help Grandma / Invention Challenge activity engaged most frequently with referencing narrative, followed by elaborating narrative. In the visitor-generated narrative condition of the same activity, visitors most frequently engaged with inventing narrative. Within the Chain Reaction activity, visitors most frequently engaged with referencing narrative in the guided narrative condition. However, unlike the Help Grandma / Invention Challenge, visitors in the guided narrative of the Chain Reaction activity also frequently engaged with inventing narratives. Within the Chain Reactions visitors invented narratives equally often in the guided and the visitor-generated narrative condition. Within Air-Powered Vehicles, visitors most frequently engaged with referencing narrative in the guided narrative condition and

with inventing narrative in the visitor-generated narrative condition regardless of museum. Only at The Tech did visitors also frequently engage with elaborating narratives in the guided narrative condition of the Air-Powered Vehicles activity.

While the small differences mentioned above were found across conditions, activities, and museums, the overall picture is clear that the guided narrative condition more frequently supported referencing and elaborating narrative while the visitor-generated narrative condition generally prompted more narrative invention. Inventing users was not a common practice, yet remained fairly consistent across conditions, activities, and museums.

Table 11. Counts (and percentages) of visitors for whom each narrative practice was observed in the Air-Powered Vehicles activity within museum settings.

	NYSCI N (% of obs.)			The Tech N (% of obs.)			Amazeum N (% of obs.)		
	guided	visitor-generated	both	guided	visitor-generated	both	guided	visitor-generated	both
Total	20	20	40	19	18	37	13	14	27
Narrative practice									
<i>Referencing narrative</i>	9 (45.0%)	0 (0.0%)	9 (22.5%)	9 (47.4%)	0 (0.0%)	9 (24.3%)	3 (23.1%)	1 (7.1%)	4 (14.8%)
<i>Elaborating narrative</i>	1 (5.0%)	0 (0.0%)	1 (2.5%)	4 (21.1%)	0 (0.0%)	4 (10.8%)	1 (7.7%)	1 (7.1%)	2 (7.4%)
<i>Inventing narrative</i>	3 (15.0%)	4 (20.0%)	7 (17.5%)	2 (10.5%)	2 (11.1%)	4 (10.8%)	1 (7.7%)	6 (42.9%)	7 (25.9%)
<i>Inventing user</i>	1 (5.0%)	1 (5.0%)	2 (5.0%)	2 (10.5%)	1 (5.6%)	3 (8.1%)	1 (7.7%)	2 (14.3%)	3 (11.1%)

Narrative: How correlated are the narrative practices across all visitors?

To extend the findings from frequency counts above, correlations showed across all museums and activities as well as across all activities at NYSCI, that inventing narrative and inventing user had a strong relationship (see Tables 12 and 13 for Phi correlation coefficients). At NYSCI, a strong relationship was also found between referencing and elaborating narrative. Additionally, across all data and at NYSCI only there was a negligible relationship between referencing narrative and inventing narrative or inventing user. These findings suggest that when visitors referenced a given narrative, they engaged in less inventing of their own narratives and users. Interestingly, while referencing narrative occurred almost exclusively in the guided condition (see Table 9 above for frequency counts), inventing occurred across both conditions. This suggests that the guided narrative condition did not prevent visitors from inventing their own narratives and users across activities, however, when visitors did reference a given narrative within an activity, they tended to invent less. **This suggests further exploration of what caused visitors—particularly within the guided narrative condition—to reference the given narrative or invent their own narratives because if they referenced the given narrative, they were less likely to invent narratives and users.**

Table 12. Phi correlation coefficient for each pair of narrative practices across all museums and all activities

Practice	Elaborating narrative	Inventing narrative	Inventing user
Referencing narrative	0.37	0.01	0.08
Elaborating narrative		0.21	0.26
Inventing narrative			0.54

Scale: .70 - 1.00 very strong relationship; .40 - .69 strong relationship; .30 - .39 moderate relationship; .20 - .29 weak relationship; .01 to .19 no or negligible relationship; -1 to 0 no relationship.

Table 13. Phi correlation coefficient for each pair of narrative practices across all activities at NYSCI

Practice	Elaborating narrative	Inventing narrative	Inventing user
Referencing narrative	0.43	-0.03	0.07
Elaborating narrative		0.17	0.16
Inventing narrative			0.53

Scale: .70 - 1.00 very strong relationship; .40 - .69 strong relationship; .30 - .39 moderate relationship; .20 - .29 weak relationship; .01 to .19 no or negligible relationship; -1 to 0 no relationship.

Engineering: To what extent does the condition type, activity type, or museum impact the diversity of engineering practices that visitors demonstrate?

To get a sense of the general performance of engineering practices across the data, frequency counts illuminated how engineering practices were clustered in relation to conditions (i.e., guided and visitor-generated narrative) for all data points as well as to whether and how this differed in relation to activities and, in the case of Air-Powered Vehicles activity, across museums.

For both the guided and visitor-generated narrative conditions, engineering practices clustered around persistence, testing, and tinkering with very small variation across conditions. For example, as can be read in the Table 14 below, 62 out of 112 (55.4%) visitors who engaged with the guided narrative condition demonstrated persistence, whereas 52 out of 90 (57.8%) of visitors who engaged in the visitor-generated narrative condition did. Similarly, testing was demonstrated by 54.5% of visitors in the guided narrative condition and 61.1% of visitors in the visitor-generated narrative condition. Lastly, tinkering was demonstrated by 78.6% of visitors in the guided narrative condition and 75.6% of visitors in the visitor-generated narrative condition. Table 14 presents a detailed breakdown of the counts and percentages of visitors for whom each engineering practice was observed across all activities and museums. **Overall, visitors consistently engaged in similar practices across conditions. Importantly, both narrative conditions richly supported engineering and neither condition precluded visitors from engaging in particular engineering practices. An interesting example includes imagining—a practice that might be more naturally aligned with inventing users and narratives—which was observed at a consistent rate across conditions, further emphasizing that even nuances within the narrative condition did not get in the way of engineering engagement.**

Table 14. Counts (and percentages) of visitors for whom each engineering practice was observed, across all activity types and museums

	Guided Narrative N (% of obs.)	Visitor-generated Narrative N (% of obs.)	Both Conditions N (% of obs.)
Total	112	90	202
Engineering Practice			
<i>Imagining</i>	24 (21.4%)	21 (23.3%)	45 (22.3%)
<i>Scoping</i>	34 (30.4%)	28 (31.1%)	62 (30.7%)
<i>Solution finding</i>	35 (31.2%)	28 (31.1%)	63 (31.2%)
<i>Iteration</i>	46 (41.1%)	42 (46.7%)	88 (43.6%)
Persistence	62 (55.4%)	52 (57.8%)	114 (56.4%)
Testing	61 (54.5%)	55 (61.1%)	116 (57.4%)
Tinkering	88 (78.6%)	68 (75.6%)	156 (77.2%)

When broken down by activity and, in the case of Air-Powered Vehicles also across museums, frequencies again clustered around persistence, testing, and tinkering for Chain Reaction and Air-Powered Vehicles. Iteration was one of the highest frequencies for the Help Grandma/Invention Challenge at NYSCI and APV at NYSCI and The Tech (Table 15). Because there was not a notable difference between guided and visitor-generated conditions in engineering practices (see Table 14 above), frequencies below do not include a breakdown by condition.

Table 15. Counts (and percentages) of visitors for whom each engineering practice was observed within activities and museums combining guided and visitor-generated narrative conditions

	Help Grandma/ Invention Challenge N (% of obs.)	Chain Reaction N (% of obs.)	Air-Powered Vehicles (NYSCI) N (% of obs.)	Air-Powered Vehicles (The Tech) N (% of obs.)	Air-Powered Vehicles (Amazeum) N (% of obs.)
Total	38	60	38	27	37
Engineering Practice					
<i>Scoping</i>	15 (39.5%)	14 (23.3%)	7 (17.5%)	14 (37.8%)	12 (44.4%)
<i>Solution finding</i>	14 (36.8%)	21 (35.0%)	7 (17.5%)	12 (32.4%)	9 (33.3%)
<i>Imagining</i>	13 (34.2%)	14 (23.3%)	6 (15.0%)	7 (18.9%)	5 (18.5%)
Iteration	20 (52.6%)	20 (33.3%)	17 (42.5%)	22 (59.5%)	9 (33.3%)
Persistence	27 (71.1%)	25 (41.7%)	17 (42.5%)	27 (73.0%)	18 (66.7%)
Testing	15 (39.5%)	38 (63.3%)	18 (45.0%)	27 (73.0%)	18 (66.7%)
Tinkering	32 (84.2%)	45 (75.0%)	28 (70.0%)	30 (81.1%)	21 (77.8%)

Where all activities showed high levels of engineering regardless of museum and condition, a linear regression model explored whether there were any differences in the diversity of the overall

levels of engineering practices by activity. The linear regression model assessed the association between the diversity of engineering practices observed and the activity (i.e., Help Grandma, Chain Reaction, and Air-Powered Vehicles) at NYSCI. The regression model included just activities at NYSCI because adjustments for both activities and museums within one regression model would be difficult to parse. Therefore, because engineering practices were so consistent across museums, and because there is no evidence of how Help Grandma or Chain Reaction might have played out at The Tech or Amazeum, the regression at NYSCI can act as a proxy to support a more generalized interpretation of all the data across museums. **Overall, the linear regression model showed some variation between the diversity of engineering practices across activities, ranging from the Help Grandma/Invention Challenge activity encouraging the most diverse number of engineering practices, followed by Chain Reactions, followed by Air-Powered Vehicles (see Table 16). The diversity in engineering practices between Help Grandma and Air-Powered Vehicles was the only comparison found to be statistically significant ($p = 0.0374$). This suggests the need to dig into more qualitative reasons why Help Grandma encouraged a greater diversity of engineering practices.**

Table 16. Estimated difference in mean number of engineering practices among visitors, comparing activity themes (NYSCI data only)⁶

Comparison	Estimate	95% Confidence Interval	P^{α}
Help Grandma compared to Air-Powered Vehicles	1.1	(0.1, 2.1)	0.0374
Chain Reaction compared to Air-Powered Vehicles	0.4	(-0.5, 1.3)	0.3504
Help Grandma compared to Chain Reaction	0.6	(-0.3, 1.6)	0.1716

Engineering: How does dwell time relate to diversity of engineering practices?

Overall, we observed that the activities had an average dwell time above 10 minutes. Overall, this is a much longer dwell time than the museum average of 1 minute. Therefore this is a good indicator for exploring the diversity of engineering practices performed in detail over range of time. Across activities, with the exception of Chain Reaction, regardless of conditions, visitors stayed longer with the guided narrative condition. The difference in this trend for the Chain Reactions activity may be related to the fact that the observations of the guided narrative condition included those in which no collar was included. Within the chain reactions activity, the mean dwell time in the guided narrative condition when the collar was included was about 18 minutes and when the collar was not included, the dwell time was 07 minutes and 25 seconds.

The relationship between dwell times and engineering practices across all data was explored using linear regression. When stratifying the observed dwell times by the diversity of engineering practices observed (Table 17), there is a clear upward trend showing that the longer visitors engaged with the activity, a greater diversity in engineering practices were observed. In

⁶ The p value tests the null hypothesis that there is no difference in the mean number of engineering practices demonstrated by participants at exhibits with the two different themes being compared. There was strong evidence in the data of a higher average number of engineering practices demonstrated among participants at the Help Grandma exhibit, compared to participants at the Air-Powered Vehicles exhibit. Differences in mean number of engineering practices comparing Help Grandma to Chain Reaction, and comparing Chain Reaction to Air-Powered Vehicles, were not significant.

exploring the relationship between engineering practices and dwell time through a linear regression model⁷, it is found that the diversity of engineering practices increased by 0.09 with every 1 minute of engagement. The 95% confidence interval for this estimate was (0.08, 0.10), and the association was statistically significant ($p < 0.0001$). Equivalently, with every 10-minutes of dwell time the diversity of engineering practices increased by a mean of 0.91 of engineering practices demonstrated by visitors (95% CI (0.79, 1.03); $p < 0.0001$). **For every minute of engagement, the number of different engineering practices a visitor demonstrated increased by 0.09, or about 1/10 of an engineering practice. Similarly, for every ten minutes of engagement, the number of different engineering practices a visitor demonstrated increased by 0.91, or nearly one engineering practice. By rounding to the nearest integer, it can be inferred that with some variability across the dataset, beginning from the moment a visitor entered the activity space, approximately ten minutes of engagement equated to approximately 1 additional type of engineering practice demonstrated.**

Table 17. Summary of observed dwell time, stratified by the number of engineering practices demonstrated across conditions, activities, and museums.

Number of engineering practices	Number of participants	Mean	(Standard Deviation)	Median	(IQR) ^α
0	34	1.9	2.8	0.8	(0.4 to 2.9)
1	30	6.1	5.5	3.9	(1.8 to 9.0)
2	14	11.9	10.7	7.8	(4.0 to 17.6)
3	27	18.9	8.6	16.0	(13.8 to 22.7)
4	24	25.8	14.9	21.2	(14.3 to 33.9)
5	40	31.8	15.9	30.4	(20.4 to 40.4)
6	23	36.3	18.1	30.4	(22.6 to 43.5)
7	10	36.5	17.2	32.7	(25.9 to 47.6)

IQR^α -- interquartile range. The interquartile range is the interval from the value at the 25th percentile to the 75th percentile of the variable.

As the data show, the mean dwell time across all activities was greater than ten, thus meaning that all activities provided rich opportunities to observe visitors engaging in engineering practices. Overall, the regression model exploring the relationship between dwell time and engineering supports the conjecture that all activities were rich spaces for visitors to engage in engineering practices regardless of condition.

Data further suggest that after about 30 minutes of engagement in activities (see medians in Table 17), the number of engineering practices remains about the same. The graph in Figure 9 visually reinforces this point as it shows the relationship between minutes of engagement and number of engineering practices demonstrated for all visitors at all museums and activities. This means that after 30 minutes of engagement, while some visitors would engage in a diverse range of engineering practices, the increase was negligible. **This suggests that after 30 minutes of engagement, visitors may instead be engaging more deeply with a few practices or engaging in some of the same practices repeatedly rather than continuing to diversify in engineering practices. This finding presents an important discussion point on balancing**

⁷ The dependent variable in the linear regression model was the number of engineering practices (an integer from 0 to 7), and the independent variable was total minutes.

the diversity and range of engineering practices with depth of engineering practices within activities.

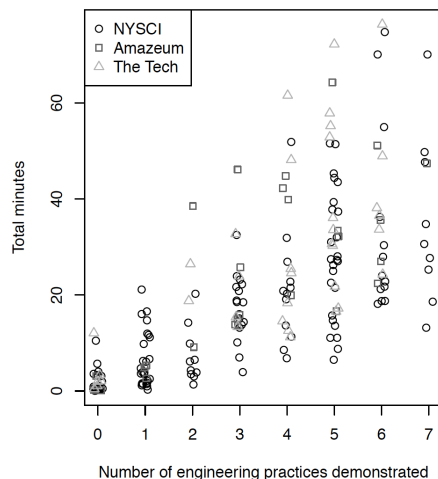


Figure 9. Dwell time for visitors at all museums and all activities in relation to the diversity of engineering practices exhibited.

Engineering: How correlated are the observed engineering practices across all visitors?

We correlated the observed engineering practices in order to get a sense of whether and how these concepts were observed as separate or overlapping constructs. The analysis was intended to guide future iterations of the observation instrument and to identify related sub-constructs of overarching ideas related to engineering within the context of the museum activities. Across museums and activities, some of the seven engineering practices observed co-occurred strongly (Table 18, bolded correlations). For example, iteration strongly correlated with scoping, solution finding, persistence, and testing. Solution finding strongly correlated with imagining and persistence was also strongly correlated with testing and tinkering. Additionally, scoping and imagining were not at all correlated.

Table 18. Phi correlation coefficient for each pair of engineering practices, across all museums and all activities

Practice	Solution finding	Imagining	Iteration	Persistence	Testing	Tinkering
Scoping	0.32	0.08	0.48	0.45	0.38	0.31
Solution finding		0.44	0.44	0.40	0.32	0.29
Imagining			0.30	0.33	0.24	0.21
Iteration				0.53	0.45	0.38
Persistence					0.43	0.52
Testing						0.39

Scale: .70 - 1.00 very strong relationship; .40 - .69 strong relationship; .30 - .39 moderate relationship; .20 - .29 weak relationship; .01 to .19 no or negligible relationship; -1 to 0 no relationship.

The higher correlation between persistence and iteration as well as persistence and tinkering suggests that future observation protocol iterations may want to consider testing whether

it's worthwhile to combining tinkering, persistence, and iteration or measure persistence through tinkering or iteration as these practices may be inclusive of one another. However, further discussion among researchers about the qualitative aspects of these engineering practices may result in their continued observance as separate constructs. That is, it may be the case that solution finding and imagining have a strong correlation but are still separate observable constructs. **Overall, the correlations indicate that additional and deeper analysis is warranted to further develop the observation instrument through a psychometric analysis method to determine whether or not particular practices ought to be collapsed into one another.**

Empathy: To what extent does the condition type, activity type, or museum impact the diversity of empathy markers that visitors demonstrate?

To better understand how visitors engaged with empathy markers, we analyzed empathy markers in relation to conditions (i.e., guided and visitor-generated narrative) for all data points as well as whether and how this differed in relation to activities and, in the case of Air-Powered Vehicles, activity across museums.

Table 19. Counts (and percentages) of visitors for whom each empathy marker was observed, across all data.

	Guided Narrative N (%)	Visitor-generated Narrative N (%)	Both Conditions N (%)
Total	112	90	202
Empathy marker			
<i>How designer feels</i>	1 (0.9%)	4 (4.4%)	5 (2.5%)
<i>Societal issue</i>	1 (0.9%)	5 (5.6%)	6 (3.0%)
<i>How user feels</i>	2 (1.8%)	4 (4.4%)	6 (3.0%)
<i>Desire to help</i>	11 (9.8%)	6 (6.7%)	17 (8.4%)
<i>UCD criteria</i>	11 (9.8%)	6 (6.7%)	17 (8.4%)
<i>Familiarity</i>	18 (16.1%)	19 (21.1%)	37 (18.3%)
<i>perspective-taking</i>	27 (24.1%)	28 (31.1%)	55 (27.2%)

Across both the guided and visitor-generated narrative conditions, observations of empathy markers were lower than engineering practices. This may be the case because most often empathy was not performed in action but had to be observed through speech during the semi-structured interviews facilitated while visitors were designing. Across both conditions, frequencies of empathy markers clustered around perspective-taking and familiarity. Additionally, within the guided narrative condition, empathy markers around desire to help and user-centered design (UCD) were also frequent. For example, in the guided narrative condition, 24.1% of all observed visitors showed evidence of perspective-taking, 16.1% showed evidence of familiarity and 9.8% showed evidence of desire to help as well as user-centered design criteria. In the visitor-generated narrative condition, 27.2% of all observed visitors showed evidence of perspective-taking and 18.3% showed evidence of familiarity. Both of these frequencies were proportionally

slightly higher for the visitor-generated narrative condition than the guided narrative condition. Table 19 presents a detailed breakdown of the counts and percentages of visitors for whom each empathy marker was observed across all activities and museums. **Overall, the frequencies show that like engineering practices, empathy markers were overwhelmingly consistent across conditions with slight variation.**

When broken down by activity and, in the case of Air-Powered Vehicles also across museums, frequencies also clustered around perspective-taking and familiarity (Tables 20 and 21). Additionally, in the Help Grandma / Invention Challenge activity frequencies clustered around the empathy markers of desire to help and user-centered design criteria (Table 20). **It is particularly interesting to note that overall visitors exhibited many more empathy markers within the Help Grandma/Invention Challenge activity overall, less in the Chain Reactions activity, and even less in the Air-Powered Vehicles activity. This suggests that there were explicit and implicit elements of the Help Grandma/Invention Challenge activity design that supported the cultivation of empathy.**

Table 20. Counts (and percentages) of visitors for whom each empathy marker was observed within activities at NYSCI

	Help Grandma/Invention Challenge N (% of obs.)			Chain Reaction N (% of obs.)			Air-Powered Vehicles N (% of obs.)		
	guided	visitor-generated	both	guided	visitor-generated	both	guided	visitor-generated	both
Total	19	19	38	41	19	60	20	20	40
Empathy marker									
<i>How designer feels</i>	0 (0.0%)	2 (10.5%)	2 (5.3%)	1 (2.4%)	0 (0.0%)	1 (1.7%)	0 (0.0%)	1 (5.0%)	1 (2.5%)
Societal issue	0 (0.0%)	4 (21.1%)	4 (10.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (5.0%)	1 (2.5%)
<i>How user feels</i>	0 (0.0%)	1 (5.3%)	1 (2.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Desire to help	7 (36.8%)	4 (21.1%)	11 (28.9%)	2 (4.9%)	0 (0.0%)	2 (3.3%)	0 (0.0%)	2 (10.0%)	2 (5.0%)
UCD criteria	8 (42.1%)	4 (21.1%)	12 (31.6%)	1 (2.4%)	0 (0.0%)	1 (1.7%)	0 (0.0%)	2 (10.0%)	2 (5.0%)
Familiarity	8 (42.1%)	7 (36.8%)	15 (39.5%)	3 (7.3%)	3 (15.8%)	6 (10.0%)	3 (15.0%)	0 (0.0%)	3 (7.5%)
perspective-taking	9 (47.4%)	9 (47.4%)	18 (47.4%)	6 (14.6%)	5 (26.3%)	11 (18.3%)	1 (5.0%)	4 (20.0%)	5 (12.5%)

Table 21. Counts (and percentages) of visitors for whom each empathy marker was observed in the Air-Powered Vehicles activity within museum

	NYSCI N (% of obs.)			The Tech N (% of obs.)			Amazeum N (% of obs.)		
	guided	visitor-generated	both	guided	visitor-generated	both	guided	visitor-generated	both
Total	20	20	40	19	18	37	13	14	27
Empathy Marker									
<i>How designer feels</i>	0 (0.0%)	1 (5.0%)	1 (2.5%)	0 (0.0%)	1 (5.6%)	1 (2.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
<i>Societal issue</i>	0 (0.0%)	1 (5.0%)	1 (2.5%)	1 (5.3%)	0 (0.0%)	1 (2.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
<i>How user feels</i>	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (10.5%)	1 (5.6%)	3 (8.1%)	0 (0.0%)	2 (14.3%)	2 (7.4%)
<i>Desire to help</i>	0 (0.0%)	2 (10.0%)	2 (5.0%)	1 (5.3%)	0 (0.0%)	1 (2.7%)	1 (7.7%)	0 (0.0%)	1 (3.7%)
<i>UCD criteria</i>	0 (0.0%)	2 (10.0%)	2 (5.0%)	2 (10.5%)	0 (0.0%)	2 (5.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
<i>Familiarity</i>	3 (15.0%)	0 (0.0%)	3 (7.5%)	4 (21.1%)	4 (22.2%)	8 (21.6%)	0 (0.0%)	5 (35.7%)	5 (18.5%)
<i>Perspective taking</i>	1 (5.0%)	4 (20.0%)	5 (12.5%)	8 (42.1%)	7 (38.9%)	15 (40.5%)	3 (23.1%)	3 (21.4%)	6 (22.2%)

The differences in empathy markers across activities—particularly for Help Grandma/Invention Challenge—suggest that more about cultivating empathy can be understood by closely examining more specific features of engineering activity design. (Note also the previously presented finding of a significant difference in engineering practices between Help Grandma and Air-Powered Vehicles). For example, in taking a closer look at the observations within the Chain Reactions activity, a difference was found within the guided narrative version of the activity. While about half of the guided narrative observations included a dog collar, about half did not. Of the visitors who engaged with the guided narrative dog collar version, 5 people demonstrated at least some empathy, while only one visitor of all observed demonstrated empathy in the non-collar version. **While more people showed empathy in the visitor-generated narrative condition of the same activity, it is interesting to consider why the design choice of the dog collar may have prompted some empathy in the guided narrative version. Subtle shifts in design may have supported shifts in if and how visitors demonstrated particular markers. Additionally, although diversity of empathy markers was about the same in guided and visitor-generated, overall, visitors engaged much longer in the guided narrative version of Chain Reaction (18 minutes versus 7:25 minutes) and so perhaps the depth of empathic engagement is worth further exploration.**

Empathy: How correlated are the empathy markers across all visitors?

Similar to the engineering practices, we ran correlations for the observed empathy markers to analyze their relationships and raise the question of whether the individual markers should continue to be observed as separate constructs. This analysis can guide future iterations of the observation protocol and support the identification of sub-constructs of overarching ideas related to engineering within the context of the museum activities. Across museums and activities, some of the six empathy makers co-occurred strongly (Table 22, bolded correlations). For example, desire to help strongly correlated with user-centered design across all data. Additionally, many of the empathy markers did not correlate much at all across all museums and activities. For example, perspective-taking, one of the most frequently occurring empathy markers, was not strongly correlated with societal issues or the affective empathy marker of how the designer feels. Additionally, familiarity, the other most frequently occurring empathy marker, was not correlated with desire to help, affective empathy markers (i.e., how the designer feels and how the user feels), and societal issues.

Table 22. Phi correlation coefficient for each pair of empathy markers across all museums and all activities

Marker	UCD criteria	How designer feels	How user feels	perspective-taking	Familiarity	Societal issue
Desire to help	0.61	-0.05	0.16	0.22	0.18	0.37
UCD criteria		0.07	0.26	0.34	0.27	0.16
How designer feels			-0.03	0.12	0.09	-0.03
How user feels				0.29	0.14	-0.03
perspective-taking					0.23	0.02
Familiarity						0.14

Scale: .70 - 1.00 very strong relationship; .40 - .69 strong relationship; .30 - .39 moderate relationship; .20 - .29 weak relationship; .01 to .19 no or negligible relationship; -1 to 0 no relationship.

The particularly high correlation between the desire to help and user-centered design suggests that expressed empathy in relation to both empathy markers could be observed together in future observations. For these empathy makers, additional analysis is warranted to further develop the observation instrument. The dominant lack of relationship between most empathy markers suggests that the empathy markers used for the evaluation research are likely distinct constructs and can continue to be observed separately.

What was the relationship between engineering practices and empathy markers?

Because it was found that regardless of narrative condition, visitors engaged in ample engineering practices and showed evidence of empathy, the relationship between engineering and empathy became a central point of interest for analysis to support evaluation research goals. Initial descriptive analysis laid the foundation for examining how engineering and empathy related to one another in the context of the narrative activities.

Table 23 descriptively illuminates the relationship between engineering practices, dwell time, and the presence of empathy across conditions, activities, and museums. As the table shows, the presence of any empathy (1 marker or more as noted on the observation protocol) corresponded to a mean of 4.5 engineering practices across visitors and an average dwell time of 29:49 minutes. Demonstrating no empathy corresponded with a mean of 2.3 engineering practices and an average dwell time of 13:20 minutes. This results in an additional 16:29 minutes of engagement in activities on average for visitors who showed any empathy. While these

differences are only descriptive, they hint at a very strong relationship between the constructs of empathy, engineering, and dwell time. Furthermore, the average number of empathy markers for visitors showing any empathy was just 1.8, meaning that understanding how to engage visitors in demonstrating empathy—even just a little—might be valuable for how they engage with engineering and how long they stay with activities. Because only about 39.6% (80) visitors showed any empathy, the data here presents an opportunity for understanding how to design for empathy within engineering activities so that some of the other 60.4% (122) visitors might be prompted to demonstrate any empathy within activities.

Table 23. Relationship between mean engineering practices, dwell time, and presence of empathy markers across all conditions, activities, and museums

	Mean empathy markers (1-7 markers)	Mean engineering practices (1-7 practices)	Average dwell time (minutes)	Frequency N (% of obs.)
Empathy (1-7 markers)	1.8 markers	4.5 engineering practices	29:49 minutes	80 (39.6%)
No empathy (0 markers)	0 markers	2.3 engineering practices	13:20 minutes	122 (60.4%)
TOTAL	0.71 markers	3.2 engineering practices	19:51 minutes	202

To statistically test the relationship between engineering and empathy, the Spearman correlation coefficient was used to assess the strength of the correlation (i.e., the monotonic association) between the number of engineering practices and the number of empathy markers demonstrated by visitors. Specifically, the Spearman correlation assessed whether the number of empathy markers and engineering practices changed together over time together, yet not necessarily at a constant rate (as a Pearson correlation would test). Spearman correlations are used most often with continuous or ordinal variables as is the case here (i.e., a visitor could demonstrate 0-7 empathy markers and 0-7 engineering practices). Additionally, all data were used to test this correlation so that the relationship between engineering and empathy could be understood across our dataset and more data offered more opportunities to better understand the relationship across a variety of contexts. A Spearman correlation was also used for just NYSCI data for the sake of comparison.

Overall, among visitors **at all museums and all activities**, the Spearman correlation between diversity of engineering practices and diversity of empathy markers was 0.48 (a moderately strong association). When stratified by condition, the estimated correlation was slightly stronger among visitors at the guided narrative condition of the activities (0.53) than at the visitor-generated narrative condition of the activities (0.41). Among visitors at all **NYSCI activities**, the correlation between number of engineering practices and number of empathy markers was 0.50 (a moderately strong association). When stratified by condition, the estimated correlation was slightly stronger among visitors at the guided narrative condition (0.53) than at the visitor-generated narrative condition (0.46).

Across the data, the findings suggest that there is a moderately strong relationship between empathy and engineering, regardless of condition, museum context, and activity. Furthermore, while the difference in the correlation coefficients among conditions is slight, it might be interesting to probe what about the design of the guided narrative conditions fostered a slight uptick in empathy and engineering correlations.

The relationship between empathy and engineering is further emphasized in a linear regression model with all data with the number of engineering practices as the dependent variable and number of empathy markers as the independent variable. Within this model,

there were statistically significant linear associations between empathy markers and engineering practices. The details of this model showed that among visitors at all museums and all activities, the mean number of engineering practices demonstrated was estimated to be 0.8 higher among visitors demonstrating one additional empathy marker (95% CI (0.6, 1.1); $p < 0.0001$). Among visitors at all NYSCI activities, the mean number of engineering practices demonstrated was estimated to be 0.9 higher among visitors demonstrating one additional empathy marker (95% CI (0.6, 1.2); $p < 0.0001$). The association was not substantively different when the model was adjusted for condition (guided vs. visitor-generated).

How does dwell time relate to diversity of empathy markers and engineering practices?

Across the data, observations also showed that like with engineering practices, when participants demonstrated empathy markers, their mean dwell time was higher (see Tables 23 and 24). While this difference is interesting enough, to statistically test how much dwell time mattered, linear models were used to assess the association between the number of engineering practices demonstrated and the number of empathy markers demonstrated, controlling for visitors' dwell time at the activities. Among visitors at all museums and all activities, one additional empathy marker was associated with a 0.3 higher mean number of engineering practices, when comparing visitors with the same dwell time at the activities. The 95% confidence interval for this estimate was (0.1, 0.6), with $p = 0.0009$. Among visitors at NYSCI (all activities), one additional empathy marker was associated with a 0.3 higher mean number of engineering practices, when comparing visitors with the same dwell time at the activities. The 95% confidence interval for this estimate was (0.1, 0.6), with $p = 0.0089$. **In summary, this means that even when controlling for dwell time for visitors across activities, the difference in engineering practices between those who demonstrated empathy and those who did not was still statistically significant.**

Table 24. Summary of observed dwell time, stratified by the number of empathy markers demonstrated. Here, data is pooled from all museums and exhibits.

Number of empathy markers	Number of participants	Mean	(Standard Deviation)	Median	(IQR) ^{α}
0	122	13.3	14.7	6.7	(1.8 to 20.7)
1	43	31.1	20.1	25.5	(14.5 to 38.3)
2	20	28.0	13.8	24.4	(15.9 to 36.6)
3	9	24.7	11.3	25.3	(16.0 to 32.0)
4	7	31.4	11.2	26.4	(22.9 to 41.9)
5	1	43.5	NA	43.5	(43.5 to 43.5)

IQR ^{α} = interquartile range. The interquartile range is the interval from the value of the 25th percentile to the value of the 75th percentile of the variable.

How are activities designed to evoke empathy in engineering?

The present study found promising possibilities for fostering empathy in engineering design activities across museum contexts. Recommendations for design possibilities that can foster empathy are offered for activity frames, choice of characters, and facilitation strategies.

Design learnings from Help Grandma/Invention Challenge. Results suggest how informal engineering activities can be designed to **cultivate empathy through purposeful activity frame choices** like asking learners to design something to help a familiar character (like a grandmother), and by including descriptions of relatable problems (as with the dog collar)

to support girls' engagement in challenges. For example, within the guided narrative condition of the Help Grandma/Invention Challenge activity, one visitor (Lexi) pulled an activity card about inventing something to help a grandmother open a jar. Lexi demonstrated the familiarity empathy marker when she explained that “her grandma has a hard time (opening jars).” As she used rubber bands, a spatula, nuts and bolts, and a metal hook to construct her original invention, she demonstrated the perspective-taking empathy marker for a grandma user and showed how a grandmother might use what she made, demonstrating how to open a jar with the invention. Lexi also showed the engineering marker of persistence within the activity when she could not initially get her design to work and continued adding rubber bands to get it to function in the way she wanted it to. Across conditions, the frame of the engineering design activity was to produce something that could solve a practical challenge that a particular character (e.g., a grandmother, a nonna, an amama, etc.) might face or a more general challenge that can be observed across people in the real world (e.g., carrying something across the street). As inherent to the engineering design process, **a small change toward a more specific design challenge in ways that are relatable and familiar to museum visitors can support empathy within engineering activities.** For the Help Grandma/Invention Challenge activity, it is important to not that the activity assumed a particular fragility of older women, a societal discourse likely familiar to museum visitors whether young or old. In moving forward, **it may be even more worthwhile to consider ways in which people could invent their own familiar characters that would require support in a variety of tasks (e.g., such as opening jars) to bring about empathy through designing for familiarity within engineering activities.**

The example of Lexi further speaks to **the usefulness of the explicit inclusion of relatable and familiar characters within engineering design activities to cultivate empathy.** Because participants were invited to create something to help a grandmother or to create an invention within the activity, the value of the Help Grandma activity is in how it supported girls in the human-centered design process inherent in engineering. While not all participants explicitly demonstrated empathy within the activity (e.g., some girls made inventions without specific reference to the grandmother user), the design of the activity supported expressions of empathy by linking engineering design to a relatable character.

Design learnings from Chain Reaction and Air-Powered Vehicles. It is interesting to note that other activities also called for visitors to design for particular character, including a photographically represented nearly life-sized dog in the Chain Reactions guided narrative condition and a 3D printed polar bear that was included in the tundra landscape of the guided narrative condition of the Air-Powered Vehicles activity at the Amazeum. For example, one visitor (Ava) engaged with the guided narrative of the Chain Reaction activity. She began the activity by stating her technical understanding of the challenge: “Oh I get it. You’re supposed to connect [the circuit] and it makes the dog move.” As she worked with materials to create her chain reaction, Ava referenced the dog repeatedly, drawing on her prior understandings about dogs’ behavior. She showed *familiarity* when she explained, “dogs usually eat bones and bones are white, so I am trying to make it look like that.” She demonstrated the empathy marker *affective-user* when she said, “the dog is getting really hungry...it’s actually having trouble finding its way to the bone because it’s so hungry.” Her participation within the Chain Reaction activity was consistently driven by the relatable challenge as she continued to describe what she was making and why she was making it, creating a meaningful context for her activity.

Additionally, within the Chain Reaction activity, an interesting implication for design was based around the introduction of a physical prop—a dog collar—that was either present or not present around the cardboard dog’s neck in the activity setup. For roughly half of the observations in Chain Reaction the cardboard dog model wore a collar that read “I want to play,” “I’m lonely,” or “I am hungry.” The second half of observations did not include collars. In the case of Ava, the collar was present, which provided a visual prompt for her to reference the dog’s needs and explain why she was designing. By contrast, throughout the observations of the Air-Powered

Vehicles activity that included a 3D printed polar bear, visitors did not engage with the added polar bear character.

The collars in combination with the photographic representation and nearly life size measures of the dog collar may have provided the necessary prompt to support empathy by asking girls to help solve a specific problem that the “user” of the design (in this case, the dog) was facing. The relatability of the dog character and premise of helping it play or eat was increased because dogs are common pets and the museum designers opted to **use a realistic and life-sized model of a dog rather than a cartoon character or illustration.**

VII. Summary of findings

The underrepresentation of girls in engineering is a persistent historical trend. Because museums cater to a broad range of visitors on a daily basis, museum contexts are uniquely positioned to offer engineering engagement opportunities for young girls. Building on promising research that suggests that engineering is connected to empathy (e.g., Walther, Miller, & Sochacka, 2017) and that empathy can be cultivated through narrative (e.g., Bennett & Monahan, 2016), this research had two main aims:

- 1) To investigate whether and how engineering design activities that included a guiding narrative compared to activities that did not guide narratives related to the engagement with engineering practices
- 2) To research whether and how the conditions (i.e., guided narrative, visitor-generated narrative), activities, and museum contexts influenced visitors' demonstration of narrative practices, engineering practices, and empathy markers.

In this section, the major findings of the present report are summarized to provide a high-level overview of the most interesting results and shed light on the research aims.

First, narrative results are summarized, including how guided and visitor-generated narrative conditions, activities, and museum contexts influenced engagement with narrative practices for the target population (i.e., girls 7-14). Discussion of correlated narrative practices is included to get a sense of the reliability of the research instrument to measure separate narrative constructs.

Second, a focused summary includes engineering results. Engineering results are reported related to conditions, activity type, museum. An analysis of engineering related observations further iterate how dwell time relates to the diversity of demonstrated engineering practices to substantiate that all three activities are high-quality research contexts across conditions and museums. Correlations between engineering practices demonstrate the research instrument reliability in measuring distinct constructs.

Third, a summary of the results on empathy highlight how condition, activity, and museum context relate to the diversity of empathy markers to get a sense of how visitors empathy demonstrated empathy across contexts. Similar to narrative and engineering practices, empathy marker correlations demonstrate the co-occurrence of sub-constructs.

Lastly, results show how dwell time, engineering practices, and empathy markers are related. Together, the findings present a compelling case of designing for empathy to foster engineering practices among girls (age 7-14) and support recommendations for design possibilities across activities and museum contexts.

Summary of narrative findings

Narrative practices were demonstrated across activities and museums regardless of condition. This means that whether or not narrative elements were explicitly embedded in activity design, visitors brought their own narratives. Still, more narrative practices were observed in the guided narrative condition overall, which confirms that designs meant to explicitly evoke a particular narrative did, in fact, evoke more narrative practices.

Within the guided condition, narrative practices of referencing narrative and elaborating narrative were more frequently present than inventing narratives and inventing users. Within the visitor-generated narrative condition, inventing narrative and inventing users were most frequently demonstrated. Frequency clusters are an interesting and perhaps unsurprising trend, which illustrate that within the guided condition, visitors engaged with the given narrative more frequently than inventing their own narratives. However, it is important to note that the guided narrative condition did not prevent visitors from bringing their own narratives as well, as can be seen by the

relatively high frequency of inventing narrative practices within this condition (nearly comparable to the frequency in the visitor-generated condition). Thus, even though the guided narrative design supported visitors in incorporating a given narrative, visitors still invented their own users at the same rate as they might when no explicit narrative elements are incorporated within activities. This suggests further exploration of what caused visitors—particularly within the guided narrative condition—to reference the given narrative or invent their own narratives. Importantly, if they referenced the given narrative, they were less likely to invent narratives and users.

Summary of engineering findings

Both conditions across activities and museums supported visitors to engage in engineering practices. There were ample numbers of engineering practices present across the board. Across conditions for all data visitors engaged most in the engineering practices of persistence, testing, and tinkering. This trend remained consistent across activities and museums, with a few exceptions. For instance, the Help Grandma/Invention challenge activity as well as the Air-Powered Vehicle activity at NYSCI and The Tech also frequently prompted iteration. Overall, visitors who engaged with the Help Grandma/Invention Challenge activity demonstrated a higher diversity of engineering practices compared to the Air-Powered Vehicles activity. This was surprising as the Air-Powered Vehicles was hypothesized to be the most engineering-centric during the museum partners' design and activity selection process.

Linear regression showed that for every 10 minutes visitors engaged with an activity (regardless of condition, activity type, or museum context), the diversity of engineering practices visitors demonstrated increased by 1. While 10 minutes seems like a long time, (especially knowing that the average mean engagement time in science museum activities is 1 minute (Dancstep & Gutwill, 2019)), across the data, mean dwell times were greater than 10 minutes. This average supports the assertion that all activities in the present study provided rich opportunities to observe visitors engaging in engineering practices. This means that the longer visitors stayed with the activity, the wider the range of engineering practices they demonstrated across all conditions, activities, and museums. Additionally, after around 30 minutes of engagement time visitors seemed to engage more deeply with a few of the same engineering practices rather than continuing to diversify engineering practices. This suggests that considering the depth of engineering practices may become important in future observations in addition to tracking the diversity of engineering practices.

Across the data, some engineering practices were strongly correlated, which suggests that future observations may combine some practices. However, further psychometric investigations are needed to better understand and make recommendations to refine the protocol. For example, it may be the case that solution finding and imagining have a strong correlation but are still valuable as separate observable constructs in the context of the study. Overall, the correlations indicate that additional investigation is warranted to further develop the observation instrument through psychometric analysis to determine whether or not particular practices ought to be collapsed into one.

Summary of empathy findings

Across all conditions, activities, and museums the empathy markers were equally clustered around perspective-taking, familiarity, user-centered design criteria, and desire to help. It is particularly interesting to note that overall, visitors exhibited many more empathy markers within the Help Grandma/Invention Challenge activity, less in the Chain Reactions activity, and even less in the Air-Powered Vehicles activity. This suggests that there were explicit and implicit elements of the Help Grandma/Invention Challenge a design that supported the cultivation of empathy. This finding is particularly interesting as the Help Grandma/Invention Challenge activity also prompted higher frequencies of engineering practices.

Within the Chain Reactions activity more people showed empathy in the visitor-generated narrative condition of the activity, yet it is also interesting to consider why the design choice of the dog collar may have prompted some empathy in the guided narrative version. Subtle shifts in design may have supported shifts in empathy. Additionally, visitors engaged much longer in the guided narrative version of Chain Reaction (18 minutes versus 7:25 minutes) and so the depth of empathic engagement is worth further exploration.

Across all data, when visitors showed at least one empathy marker they stayed with the activity longer than when they demonstrated no empathy marker. This suggests that designing to support empathy could encourage visitors to engage longer with an activity.

The particularly high correlation between the desire to help and user-centered design suggests that expressed empathy in relation to both empathy markers could be observed together in future observations. For these empathy makers, additional analysis is warranted to further develop of the observation instrument.

Summary of relationship between engineering and empathy

The evaluation research further focused on examining the relationship between empathy and engineering. The aim was to empirically demonstrate the value and interconnected nature of empathy and engineering within the engineering design process. Across conditions, activities, and museums, when visitors demonstrated any empathy markers (i.e., a number of empathy markers 1-77) the mean number of demonstrated engineering practices nearly doubled (i.e., from a mean of 2.3 engineering practices with empathy markers present to a mean of 4.5 engineering practices without empathy markers present). This was a statistically significant difference between diversity of engineering practices observed with and without empathy present. Additionally, with empathy markers present, the dwell time of observed participants was over twice as long compared to when no empathy markers were demonstrated.

To statistically test the relationship made visible through descriptive statistics, a linear regression model with data across conditions, activities, and museums showed that there was a statistically significant linear association between empathy markers and engineering practices. Across all data there was a moderately strong relationship between engineering practices and empathy markers. These findings suggest that when visitors demonstrated any empathy (regardless of the number of markers) within the engineering design activities, the diversity of their engineering practices increased. These findings hold constant even when controlling for dwell time, meaning that even if dwell time was the same, a statistically significant difference in diversity of engineering practices was found between visitors who demonstrated empathy and those who did not. For the design of activities, this means that small changes in evoking any empathy markers can have a profound impact on engineering practices.

Summary of design recommendations

A number of design recommendations rise to the surface given the high-level findings on the clear relationship between empathy and engineering. Qualitative analyses suggest that empathy can be cultivated through purposeful activity frame choices like asking learners to design something to help a familiar character (like a grandmother) and by including descriptions of relatable problems (as with the dog collar) to support girls' engagement in engineering challenges. Small design changes that result in more specific design challenges can support relatability and familiarity empathy markers within engineering activities. Furthermore, designing for human as well as particular designed animal characters can support the cultivation of empathy in engineering activities. The collars in combination with the photographic representation and nearly life size measures of the dog collar may have provided the necessary prompt to support empathy in the guided narrative version of the activity by asking girls to help solve a specific problem that the "user" of the design (in this case, the dog) was facing. The relatability of the dog character and premise of helping it play or eat increased the presence of empathy because dogs are

common pets and the museum designers opted to use a realistic and life-sized model of a dog rather than a cartoon character or illustration. For future design iterations it may be worth considering how visitors could invent their own familiar characters (perhaps within their own lives) that would require support in a variety of tasks (e.g., such as opening jars) to bring about empathy through designing for familiarity within engineering activities. Designing scaffolds to prompt particular challenges (e.g., a difficult to open cupboard or an unstable stool) could support visitors in relating challenges to their own lives.

VIII. Questions for consideration

While a thoughtful discussion is a logical next step and essential to interrogating the relationship between quantitative and qualitative data in the present report, the evaluation team proposes a few questions that will help identify critical next steps. These questions are grounded in analyses and meant as a springboard for future conversations amongst museum professionals/researchers and the evaluation team.

- How might particular qualitative differences in museum settings, including cultural aspects of the museums, diverse populations, facilitation styles, space arrangements, and material use impact empathy design recommendations?
- How did qualitative differences in activities, including materials, project set-ups, project goals, and types of engineering emphasized influence quantitative outcomes across narrative practices, engineering practices, and empathy markers?
- How did the research instrument constrain and open up conceptualizations of narrative, engineering, and empathy constructs? How might the instrument be revised in future iterations?
- How can empathy be expressed in non-verbal ways that are consequential for engineering engagement? How might an observation protocol track non-verbal expressions of empathy?

VIII. References

- American Association of University Women (2015). *Solving the equation: The variables for women 's success in engineering and computing*. Available for download at: https://www.ehu.eus/documents/2007376/3500574/solving_the_equation.pdf .
- Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of engineering education*, 96(4), 359.
- Bennett, D. (2000). Inviting girls into technology: developing good educational practices. Commissioned paper for the American Association of University Women. Excerpts in American Association of University Women (2000). *Tech-savvy: Educating girls in the new computer age*. Washington, DC: AAUW Educational Foundation.
- Bennett, D. & Monahan, P. (2013). NYSCI Design Lab: No bored kids! In M. Honey & D. Kanter (Eds.), *Design, Make, Play: Growing the Next Generation of STEM Innovators*. New York: Routledge.
- Bennett, D., Monahan, P., & Honey, M. (2016). Museum design experiences that recognize new ways to be smart. In L. Annetta & J. Minogue (Eds), *Connecting Science and Engineering Education Practices in Meaningful Ways*. New York: Springer Publishing.
- Bevan, B. (2016). 4 The museum of pink. *Intersections of Formal and Informal Science*, 41.
- Bix, A. S. (2014). *Girls coming to tech!: A history of American engineering education for women*. Cambridge, MA: MIT Press.
- Buchholz, B., Shively, K., Peppler, K., & Wohlwend, K. (2014). Hands on, hands off: Gendered access in crafting and electronics practices. *Mind, Culture, and Activity*, 21(4), 278-297.
- Buechley, L., Peppler, K. A., Eisenberg, M., & Kafai, Y. B. (Eds.). (2013). *Textile Messages: Dispatches From the World of E-textiles and Education*. New York, NY: Peter Lang.
- Buechley, L. (2013). *Thinking about Making*. Talk presented at FabLearn Conference, Stanford University, Palo Alto, CA. Available to stream at: <http://edstream.stanford.eduNideo/Play/883b61dd951d4d3f90abeec65eead2911d> .
- Buse, K. (2018). Women's Under-representation in Engineering and Computing: Fresh Perspectives on a Complex Problem. *Frontiers in Psychology*, 9, 595.
- Dancstep, T., & Gutwill, J. P. *Towards disrupting the status quo of exhibit design: A data reanalysis to support future research*. San Francisco, CA: Exploratorium.
- Dawson, E., Seakins, A., Archer, L., Calabrese Barton, A., & Dierking, L. (2015). *Equity in informal science learning: a practice-research brief*. Available for download at: <http://stem.oregonstate.edu/files/stemfiles/Equity%20brief%20-%20Youth%20Equity%20Pathways%20in%20ISL.pdf>.
- Dorie, B.L. & Cardella, M. (2013). *Engineering childhood: Knowledge transmission through parenting*. School of Engineering Education Graduate Student Series. Paper 50. <http://docs.lib.purdue.edu/enegs/50>
- Dusold, T.C. (2008). *Applications of narrative to the engineering decision making process and the pedagogy of engineering education*. Digital Depository@ Iowa State University. Graduate Theses and Dissertations.
- Eccles, J. S. (2005). Studying gender and ethnic differences in participation in math, physical science, and information technology. *New Directions for Child and Adolescent Development*, 110 (Winter), 7- 14.
- Eccles, J. S., & Wang, M. T. (2015). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, UDI: 0165025415616201.

- Engineering Accreditation Commission (2015). Criteria for Accrediting Engineering Programs, Accreditation Board for Engineering and Technology (ABET), <http://www.abet.org/>, Baltimore, Maryland, latest version.
- Feinstein, N. W., & Meshoulam, D. (2014). Science for what public? Addressing equity in American science museums and science centers. *Journal of Research in Science Teaching*, 51(3), 368-394.
- Halverson, E.R. & Sheridan, K. (2014, December). The maker movement in education. *Harvard Educational Review*, 84(4), 495-504.
- Harel & Papert, S. (Eds.) (1991). *Constructionism*. Westport, CT: Ablex Publishing.
- Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Washington, DC: American Association of University Women.
- Honey, M., & Kanter, D. E. (Eds.). (2013). *Design, Make, Play: Growing the Next Generation of STEM Innovators*. London, UK: Routledge.
- Kalil, T. (June 12, 2012). *Extreme Marshmallow Cannons! How the Government and Private Sector Can Turn American Kids on to Science Through "Making."* Slate special issue on science education. Available for download at: http://www.slate.com/articles/technology/future_tense/2012/06/every_child_a_maker_how_the_government_and_private_sector_can_turn_kids_on_to_science_and_engineering_through_making_.html.
- Keune, A., Peppler, K., & Wohlwend, K. (2019). Recognition in makerspaces: Supporting opportunities for women to “make” a STEM career. *Computers and Human Behavior*. 99, pp. 368-380.
- Lobo, A.P., Salvo, J.J., & Alvarez, J. (2013). *The newest New Yorkers: Characteristics of New York's foreign-born population*. Population Division of the New York City Department of City Planning, Document #NYC DCP 13-10.
- Meiksins, P., Layne, P., Beddoes, K., Martini, G., McCusker, M., Rideau, R., & Shah, Y. (2016). Women in engineering: A review of the 2015 literature. *SWE Magazine*, 44-65.
- Nelson, H. G., & Stolterman, E. (2003). *The design way: Intentional change in an unpredictable world: Foundations and fundamentals of design competence*. Educational Technology.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York, NY: Basic Books, Inc..
- Pruitt, J. & Adlin, T. (2006). *The Persona Lifecycle*. San Francisco, CA: Elsevier.
- Putnam, C. (2010). *Bridging the gap between user experience research and design in industry: An analysis of two common communication tools*. Ph.D. Dissertation, University of Washington, Seattle.
- Reisslein, J., Ozogul, G., Johnson, A.M., Bishop, K.L., Harvey, J. & Reisslein, M. (2013). Circuits kit K-12 outreach: impact of circuit element representation and student gender. *IEEE Transactions on Education*, 56(3), 316-321.
- Sax, L. J., Kanny, M. A., Jacobs, J. A., Whang, H., Weintraub, D. S., & Hroch, A. (2016). Understanding the changing dynamics of the gender gap in undergraduate engineering majors: 1971–2011. *Research in Higher Education*, 57(5), 570-600.
- Varma, R. (2018). US Science and Engineering Workforce: Underrepresentation of Women and Minorities. *American Behavioral Scientist*, 62(5), 692-697.
- Vossoughi, S., & Bevan, B. (2014). Making and tinkering: A review of the literature. *National Research Council Committee on Out of School Time STEM*, 1-55. Available for download at: <http://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse089888.pdf>

- Vossoughi, S., Hooper, P. K., & Escudé, M. (2016). Making through the lens of culture and power: Toward transformative visions for educational equity. *Harvard Educational Review*, 86(2), 206-232.
- Walther, J., Miller, S.E., & Keller, N.N. (2012). Exploring the role of empathy in engineering communication through a transdisciplinary dialogue. In 119th *American Society for Engineering Education Annual Conference and Exposition*, AC-2012-4670
- Walther, J., Miller, S.E., & Sochacka, N.W. (2016). *Fostering empathy in an undergraduate mechanical engineering course*. Paper presented at the American Society for Engineering Education (ASEE) Annual Conference and Exposition, New Orleans, LA.
- Walther, J., Miller, S. E., & Sochacka, N. W. (2017). A model of empathy in engineering as a core skill, practice orientation, and professional way of being. *Journal of Engineering Education*, 106(1), 123-148.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary educational psychology*, 25(1), 68-81.

Appendix A: Observation protocol



EXHIBIT:

Girl: Condition: N NN Interview: Y N
 Date: Crowdedness: H M L Exhibition: __ ppl. __ groups. __ girls.
 Exhibition time: Facilitation: H M L Group: __ ppl. __ girls.
 Group Composition (start): ☐ Changed:

Narrative

- ☐ Referencing narrative
- ☐ Elaborating narrative
- ☐ Inventing narrative
- ☐ Inventing user

Quotes & Notes

H-on / H-off

Empathy

- ☐ Desire to help
- ☐ UCD criteria
- ☐ How DESIGNER feels (affective)
- ☐ How USER feels (affective)
- ☐ Perspective-taking: Talking about/acting use
- ☐ Familiarity: Prior experience of knowledge
- ☐ Societal issue mentioned

Engineering

- ☐ Problem scoping: Multiple aspects
- ☐ Solution finding: More than one idea
- ☐ Imagining new projects / possibilities
- ☐ Iteration: Implementing improved function (larger cycle)
- ☐ Persistence: Repeating attempts to solve problems with materials
- ☐ Testing
- ☐ Tinkering: elaborating & adding

Other Practices

- ☐ Adding to prior projects
- ☐ Capturing (also adults)
- ☐ Decorating
- ☐ Joking
- ☐ Showing
- ☐ Singing
- ☐ Close to adults
- ☐ Close to peers
- ☐ Watching materials
- ☐ Watching participants
- ☐ Watching projects / copying

Materials Used

Appendix B: Semi-structured interview protocol



EXHIBIT:

Girl: Condition: N NN Interview: Y N
Date: Crowdedness: H M L Exhibition: __ ppl. __ groups. __ girls.
Exhibition time: Facilitation: H M L Group: __ ppl. __ girls.
Group Composition (start): ☐ Changed:

Project

- Can I take a picture of what you made?
- What did you make?

Empathy

- For whom or for what did you make your design?
- How do you think anyone will use your design?
- Why did you make it?

Narrative

- Is there a backstory to this?

Engineering

- What did you want this to do?
- How did you change it?
- What school activity does this experience remind you of?

Other

- What did you learn while doing this?
- What did you like most about this activity?