



Trajectories of Physical Engagement and Expression in a Co-Creative Museum Installation

Duri Long
Georgia Institute of
Technology
Atlanta, GA, USA
duri@gatech.edu

Tom McKlin
The Findings Group, LLC
Decatur, GA, USA
tom@thefindingsgroup.org

Anna Weisling
Georgia Institute of
Technology
Atlanta, GA, USA
aweisling@gatech.edu

William Martin
Georgia Institute of
Technology
Atlanta, GA, USA
williammartin@gatech.edu

Hannah Guthrie
Georgia Institute of
Technology
Atlanta, GA, USA
hguthrie6@gatech.edu

Brian Magerko
Georgia Institute of
Technology
Atlanta, GA, USA
magerko@gatech.edu

ABSTRACT

Co-creative (i.e. collaboratively creative) activities involving physical interaction are becoming more prevalent in museums as a way of promoting opportunities for exploratory learning-through-doing (e.g. [15, 32, 18]). However, there is still a need for new techniques for understanding how physical interaction relates to engagement and creative expression in order to both evaluate exhibits and iterate on their design. This article reports on a study of how family groups physically interact in a museum environment with a specific co-creative exhibit—*TuneTable*. We relate observable markers of physical interaction with stages of engagement/expression based in the literature and identify several different trajectories of participant engagement and creative expression as they navigate the exhibit. We explore what these trajectories tell us about the types of inquiry and experimentation that *TuneTable* supports and discuss design implications. This paper's main contribution is a deep study of how physical markers reveal trajectories of creative engagement within a specific co-creative installation.

Author Keywords

informal learning; co-creativity; physical engagement; tangible user interfaces

CCS Concepts

•**Social and professional topics** → **Informal education**;
•**Human-centered computing** → *Field studies*; •**Applied computing** → *Interactive learning environments*;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
C&C '19, June 23–26, 2019, San Diego, CA, USA.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-5917-7/19/06...\$15.00
<https://doi.org/10.1145/3325480.3325504>

INTRODUCTION

There has been an increase in co-creative experiences involving technology in informal learning spaces like museums, many of which aim to promote opportunities for exploratory learning-through-doing (e.g. [15, 32, 18]). A number of research projects have explored this design space, experimenting with new approaches to tangible interaction in a variety of learning domains (e.g. [15, 21, 17]). Many of these projects, however, lack an empirical evaluation (necessary for iterative design and improvement) that assesses whether engagement that can lead to opportunities for learning and creative expression is actually taking place [22].

There is a particular need for improved understanding of physical markers of engagement [1]. Conversational analysis can be used to illuminate when participants are developing inquiry skills and engaging in “learning-talk” [1, 26]. However, physical skill development is also important (and arguably more readily observable), and while there has been some work in evaluating physical interactions and their relationship to patterns of engagement in museums [30], there is an identified need for additional work in this area [1], especially in regards to understanding physical interaction with computational systems.

We aim in this paper to improve our understanding of the relationship between physical activity and engagement/creative expression in co-creative museum installations by exploring the following research question: *How can observable markers of physical interaction inform our understanding of participants’ progression of engagement and creative expression over time?* We begin by reviewing literature that relates physical engagement to creative expression and opportunities for learning. We then discuss existing methodologies used in museums and other public spaces to better understand how participants physically engage with exhibits over time. Drawing on several guiding frameworks from the related work, we conduct an in-depth analysis of physical interactions with a specific co-creative exhibit—*TuneTable*. We explore the relationship

between physical interaction and stages of engagement and expression in *TuneTable*, identifying a number of different trajectories of engagement that occur and considering design implications.

RELATED WORK

Physical Engagement and Creative Expression

Physically acting on an object produces a change in the user's behavior, which in turn creates change in the artifact—an “interactive feedback loop” [25]. This action-reaction loop plays can play an important role in facilitating creative *flow*, or the experience of participating in activities that involve discovery and result in “an almost automatic, effortless, yet highly focused state of consciousness” [25, 10]. Several researchers have presented frameworks for understanding this relationship between physical interaction and creative expression. One of these frameworks, *creative sensemaking*, provides a way to understand how participants engage in co-creative interactions [11]. Analyzing interactions through the lens of creative sensemaking allows researchers to better understand how participants alternate between mental exploration and planning, executing plans, and engaging in participatory sensemaking by thinking-through-interaction with the world. In addition, Bilda et al. present a framework for understanding physical interaction with interactive art installations that explores the interplay between *Body* (moments when the participants' focus is on their body), *Thought* (moments when the participants' focus is on their thoughts), and *Feedback* (moments when the participants' focus is on the feedback coming from the artwork) [4].

All of these approaches to understanding the relationship between physical interaction and creativity examine the interplay between perception and physical action in the process of making sense of an interactive artwork or a creative collaborator's actions. In this paper, we more closely examine what physical action can tell us about a participants' progress from the start of an interaction to the end.

Physical Engagement and Learning

There is a strong theoretical basis suggesting that learning experiences involving physical interaction may be particularly effective (see [22] for a review), and both interaction designers and museum practitioners have recognized that physical interaction has the potential to make learning experiences more intuitive and engaging [32]. Klemmer et al. originally suggested that physical interaction could facilitate learning via “thinking through doing” in the context of interaction design. This spurred a wave of design research into tangible user interfaces (TUIs) (i.e. interfaces where participants use physical objects to interact with technology (e.g. [15, 29])). Other modes of physical interaction (e.g. motion detection, exercise bikes) have been used in museums to engage participants in a “visceral” learning process [21, 7, 32, 23, 28].

Designers have taken the initiative to bring physical interaction and TUIs into informal learning spaces; however, Marshall points out that there is a lack of empirical research assessing whether tangible interaction actually leads to learning gains and which design features best promote such gains [22]. Some

recent work is starting to explore how physical actions such as hand gestures [33, 9, 3] and movement between exhibits [30] relate to learning and creativity, but more extensive research is needed in this domain. Assessing learning gains is outside of the scope of this paper, which is focused primarily on understanding engagement. However, this paper does examine whether visitors physically engage with *TuneTable* in a manner that could provide opportunities for learning (i.e. physically engaging with the learning materials in a complex way).

One existing framework that looks at physical interaction as a component of a multifaceted process of engagement is Humphrey et al.'s Active Prolonged Engagement (APE) framework [18]. We chose to look at this particular framework because it offers an approach to studying open-ended interactions that encourages participants to freely explore concepts rather than arrive at a museum-guided objective. In addition, it examines multiple facets of the informal learning experience. Humphrey et al. explore how visitors interact with “APE exhibits”, where (A)ctive means that interaction with the exhibit is led by visitors, (P)rolonged means that visitors spend more time at these particular exhibits, and (E)ngaged means that visitors build on previous actions as they interact with the installation. The framework consists of descriptions of behavioral markers of four different types of engagement—intellectual, social, physical, and emotional. One of the strengths of APE is that it emphasizes the roles that different components of engagement play in an interaction, looking not only at the role that physical engagement plays but also how it affects and relates to intellectual, social, and emotional engagement. For this reason, we chose to build off of APE in our analysis (see Developing Codes for Video Analysis). However, the APE coding scheme for analysis simply assigns APE scores to exhibits in the four categories of engagement—it does not examine visitor growth over time. We looked to existing literature on stages of progressive engagement to inform this aspect of our analysis.

Stages of Engagement

Researchers studying interactive installations have often described participant engagement in the form of a series of stages. For example, Wouters et al. take the approach of tracking participants' individual trajectories through various user roles [38]. Participants take on the roles of *passer-by*, *bystander*, *audience member*, *participant*, or *actor*, and eventually transition out of engagement to become *dropouts* due to various triggers.

Barriault and Pearson's Visitor Engagement Framework (VEF) categorizes observable interaction behaviors into three stages—*initiation*, *transition*, and *breakthrough* behaviors [2]. The authors claim that participants who reach breakthrough behaviors are able to take full advantage of the installation's learning opportunities. The framework, however, does not provide information about *when* and *in what order* behaviors are happening, and as a result, does not speak to how participants are engaging over time.

Bilda et al. argue that investigating the temporal and sequential aspects of interaction is essential to understanding the participant's experience [5]. Bilda et al. discuss *interaction*



Figure 1. A group of participants interacting with *TuneTable*

modes (dialogues between the participant and the installation) and interaction phases (representations of the participant's cognitive processes) in their model of creative engagement. The authors base their model around the intention, expectation, and perspective of the participants using the system.

de Valk presents three phases of interaction in open-ended play environments—an *invitation* to play in which the participant is attracted to the installation; an *exploration* stage in which the participant “explores opportunities for interaction” by examining the affordances of the installation; and an *immersion* stage in which the participant actually engages in play [12]. Each of these three stages of play support different design elements relating to creativity—namely: curiosity, exploration, challenge, fellowship, and competition [12]. Elements differ in their level of importance depending on the stage of play the user is in—for example, curiosity plays a critical role in the invitation stage.

The aforementioned frameworks all contribute to our understanding of the stages of physical engagement. Physical engagement for Wouters et al. begins with a simple participant and ends with an expressive actor. The Visitor Engagement Framework shows physical interactions can be categorized by their ability to provide opportunities for learning, and de Valk's research suggests that stages of engagement in playful experiences can also be categorized by their ability to support design elements related to creativity. Finally, to truly capture the nature of the interaction, the framework should preserve when and in what order interactions happen [5].

TUNETABLE

TuneTable (Fig 1) is an interactive tangible tabletop experience where participants can use computer science coding concepts to co-create sample-based music compositions. The table is designed for use by family groups visiting museums and is targeted at middle-school age (i.e. 11-14 year old) students and their families. The main goals of *TuneTable* as a project include: a) providing participants with a high-level understanding of certain computing concepts (e.g. scope, loops, conditionals); b) improving participants' perceptions of computing by introducing these concepts within a creative context;

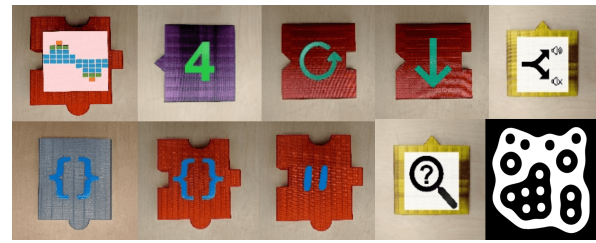


Figure 2. Block types (left to right): Sample, Value, Loop, Up/Down, Conditional, Reference (blue), Sub-reference (orange), Ditto, Inspector, Fiducial (placed on the back of each block as a unique identifier)

and c) fostering the future development of participants' interest in computing.

Participants interact with the table by placing and connecting tangible blocks on the table's surface. The table uses *reactTIVision*¹ technology to view and understand fiducial markers (Fig 2) on the bottom of each block. The fiducial markers act as unique identifiers for the blocks. This allows the *reactTIVision* computer vision library to understand which blocks are being used and how they are connected. The computer vision system can also recognize fingers on the table, which means that participants can touch graphic elements such as play buttons and on/off toggles to interact with the software.

The table utilizes a bespoke programming language called *Blockhead*, which is based off of a puzzle metaphor in which blocks can be connected together to create a subroutine. Blocks representing music samples (i.e. short clips of music, such as those you might hear in a hip-hop or electronic song) can be joined together in subroutines that consecutively play a series of sound samples, creating a musical tune (Fig 1). These tunes can be crafted to take advantage of both horizontal and vertical movement on the table: anything connected horizontally will play simultaneously before moving to the next vertically-connected component.

Further complexity can be added by connecting sample blocks with function blocks, which are designed to reflect concepts commonly used in computer programming. For example, adding a loop function block causes the connected sample block to repeat. A value block may be connected to the loop function block to direct the sample block to be played for a specified number of times, just as a loop would do in a computer program. Each block type and a description of the block's functionality is provided below (see Fig 2 for visuals).

Sample: This block spawns a playhead, which when hit will play a sound sample. Short sample blocks play shorter sounds than tall blocks. The playhead can be stopped and restarted with the touch of a finger.

Loop: When attached to a sample block, this causes the sample block to loop indefinitely or for a specified number of times (depending on the attached value block).

Up/Down: The functionality of this block is similar to a “goto” statement in many programming languages. A composition

¹<http://reactivision.sourceforge.net>

built on the table plays vertically, from top to bottom. The up/down arrow will cause the software to “jump” up/down a certain number of rows in the composition (depending on the attached value block).

Value: Value blocks can attach to loop or up/down blocks to specify the number of times the block should be applied. A special value block marked with a “?” chooses a random number between 1 and 3.

Conditional: When attached to a sample block, this block spawns a toggle, which can be turned on and off with the touch of a finger. When on, the attached sound sample plays; when off, the attached sound sample takes up time in the play sequence, but is muted.

Reference/Sub-reference: These blocks are used in pairs. A reference block in a chain will cause the composition to “jump” to its matching sub-reference block. When the sub-reference block is done playing, the composition will resume at the reference block.

Ditto: When placed next to a sample block, this causes the sample block to play twice.

Inspector: This block can be placed above any block to find out more information about it.

STUDY METHOD

Two data collection sessions were conducted at the Museum of Science and Industry Chicago in July and November 2017. In both sessions, *TuneTable* was installed in a classroom workspace, while children and their parents were recruited from the museum floor to engage in 20 minute interaction sessions. We chose to install the table in a classroom workspace (rather than directly on the museum floor) as a way to control certain variables (such as the number of people interacting with the exhibit at one time) during this early prototype stage of the work, which is a limitation of the study that is further discussed in Future Work. The researchers on the floor recruited parents with two or more roughly middle-school age children. Participant interactions with *TuneTable* were video recorded from two different perspectives (top and side view), and children participated afterwards in a brief interview with exhibit attendants while their parents filled out a survey on their demographic information.

The version of the table used in the November study was slightly different from the table used in the July study. Several new blocks (inspector and conditional, see Fig 2) were introduced in the November version of the table. The ditto block (Fig 2) was removed in the November version due to an ambiguous icon that led to participant confusion. In addition, some bugs were fixed and improvements were made to the table hardware.

Facilitators also played different roles in the interaction in the July vs. November studies. In July, participants were given a short introduction on how to interact with the table before they began the recorded interaction session. In November, participants were asked to start interacting without any introduction, but were later given a demo of more complex table compo-

nents mid-way through the interaction session. This change was made in order to explore whether the role of facilitation had any effect on participant interaction.

DEMOGRAPHIC SURVEY RESULTS

A total of 41 demographic surveys were collected across the July and November studies (this number is larger than the total number of groups who participated (32) because some groups were made up of more than one family, so more than one survey was completed per group). 73% of participants identified as White (Caucasian), while 6% identified as Mexican American or Chicano, and 8% identified as Asian American. All parent participants indicated that they had at least some high school education, with 71% having either a four-year college degree or a Master’s degree. The child participants were 51% male and 49% female. Most of the child participants were near to our target age range: 72% of child participants fell between the ages of 9 and 14 (4th-8th grade), although there was some variability due to the presence of siblings and the difficulty of assessing age when recruiting participants on-the-floor. Specifically, of 68 participants, 6% were between the ages of 3 and 6, 19% were between 7 and 9, 57% were between 10 and 14, and 18% were between 15 and 24.

DEVELOPING CODES FOR VIDEO ANALYSIS

We analyzed 32 recordings of participant group interactions ranging from 4 minutes 20 seconds to 12 minutes 10 seconds ($M = 8:51$; $SD = 1:57$). Our analysis was focused on developing an understanding of how participants’ physical interactions changed over time. In this section, we walk through the iterative process that we used to develop our coding scheme for video analysis.

Step 1: Active Prolonged Engagement (APE)

We initially began analysis by attempting to use the APE physical coding scheme to code the *TuneTable* video data. The Exploratorium APE team defines physical engagement as, “the different ways in which visitors physically interact with an exhibit. Physical engagement includes the amount of time they spend, the labels they read, where they sit or stand, and what buttons they push. It also includes the sequence of activities in which they participate” [13]. In the original APE coding scheme provided in [13], observers rank a participant group’s physical engagement on a scale of 0-5, with the following code definitions for each point on the scale: “*Level 0*: no indication of any engagement; *Level 1*: minimal engagement, very incomplete; *Level 2*: slightly more engagement but still incomplete or inadequate; *Level 3*: basic meaningful engagement. Level 3 is what we look at and say to ourselves ‘They’ve got it. This is acceptable. It is adequate.’; *Level 4*: more meaningful engagement; *Level 5*: the best we could hope for. This engagement is what we look at and say to ourselves ‘This is awesome! If only all visitors did this!’” [13].

Step 2: Adapting APE to TuneTable

We found that the APE codes as written were too general and subjective to apply to *TuneTable* and subsequently did not yield acceptable inter-rater reliability. Therefore, we developed a new set of codes that was more specific to *TuneTable*

and supplemented them with examples of actions we had observed during the study. Rather than labeling participant groups with a single number to define their entire period of engagement, we hypothesized based on the literature on stages of engagement (see Related Work) that the “ideal” participant group would progress from stage 0 to stage 5 during the course of an interaction, spending about 60 seconds in each stage.

Initial Engagement (0): The group looks at the table, approaches the table, asks what the table is about.

Pre-programming (1): The group reads label/listens to museum staff, locates blocks, puts blocks on the table, “doodles” [18] with blocks.

Partial Programming (2): The group places blocks on table, connects blocks, tests to see what sounds are made.

Simple Programming (3): The group intentionally organizes sounds, plays simple sound chains.

Complex Programming/Aesthetic Exploration (4): The group organizes blocks to make music instead of simple sounds, uses more complex function blocks like loops.

Ending Actions (5): The group puts blocks away, shows music to members of the group not already interacting.

Step 3: Thematic Analysis

When we began to apply these codes to video analysis, we realized that some of our codes were too vague to be coded reliably. For example, in level 4, we state that “The group organizes blocks to make music instead of simple sounds”. Different video coders disagreed on what qualified as “music” vs. simple sounds. We also noticed that many physical actions that participants took were not included in our coding scheme. We realized that even the version of APE that we had adapted specifically to *TuneTable* was not working in this context, and therefore we set out to develop a new model. We decided to use a bottom-up approach to analysis and defined codes based on observed physical actions. We chose to pursue a deductive thematic analysis [6], guided by the idea that participants progress through stages of engagement (as suggested by [5, 2, 38, 12]) as they interact with co-creative installations.

Nowell et al. provide a multi-step approach for conducting and reporting on thematic analysis [24]. In the first stage, researchers familiarize themselves with the data, which we had already done via our previous attempts at analysis. The next stage is generating initial codes. We began to watch through the videos to identify what we termed *atomic actions*. We defined atomic actions as physical actions taken by a single person that could not be broken down into a series of sub-actions. Examples include pressing play, pointing to the table, or connecting two sound blocks. In stage three (searching for themes), we sorted our list of atomic actions into five *composite actions* (or themes) that described higher-level action sequences involving multiple group members (e.g. “auditioning” disconnected sound blocks, “composing” a song by connecting two blocks).

We utilized reliability as a tool for refining and reviewing our themes (the fourth stage of thematic analysis). We began

an iterative process of having multiple analysts code videos using the set of themes we had developed. At team meetings, we discussed a) whether we had achieved an acceptable level of inter-rater reliability; b) what discrepancies were inhibiting reliable coding; and c) whether the themes accurately represented the data and told a cohesive story about stages of participant engagement and expression (the guiding idea behind our deductive analysis).

Step 4: Final Codebook and Analysis

Our final codebook consisted of three categories that reflect stages of physical engagement—isolated manipulation, investigative manipulation, and integrated manipulation—plus an additional code, external manipulation, that was intended to capture composite actions that did not relate to the operation of the table. The code definitions are provided below:

Isolated Manipulation: Participants are physically interacting with the exhibit, but only with isolated components. They are trying out components of the system and have not yet begun to execute sequences of related actions involving multiple exhibit components. This involves placing a new sound block on table, moving it around, and (possibly) pressing play to test this block.

Investigative Manipulation: Participants are beginning to methodically explore how the exhibit works using physical means. This may consist of trying to use some components of the system together or executing short sequences of actions. This will involve making connections between sound blocks (via moving an existing sound block or adding a new one) in chains containing *only* sound blocks, and (possibly) pressing play to test these sound blocks.

Integrated Manipulation: Participants are physically manipulating aspects of the installation in a complex way. Visitors are using different physical components of the exhibit together and executing complex sequences of actions. This will involve moving or adding a sound or function block that is already part of a chain containing both sound and function blocks on the table, and (possibly) pressing play to test these blocks.

External Manipulation: This occurs when visitors are physically manipulating aspects of the exhibit in ways that are not pertinent to the primary objective of the installation, but still demonstrate engagement and/or sense-making. This involves actions with blocks that are related to the experience but do not involve the reactive surface (e.g. looking at a block while holding it, turning a block over, arranging blocks around the perimeter of the table, handing a block to a friend, etc.).

We used a one-zero sampling approach to track how participants moved through these four codes over time [31]. Using this approach, interactions are broken down into time segments (in our case, we divided each video recording into ten second segments) and each code is given a ‘1’ if it occurred during that time segment and a ‘0’ if it did not occur. One-zero sampling has been shown to be both a reliable and valid method of behavior observation, correlating significantly with measures of actual frequency and duration while avoiding the associated difficulties of defining behavior initiation and termination [31].

We calculated a final inter-rater reliability (IRR) score for physical engagement using a subset of the video data. Two coders analyzed four different approximately ten minute participant interaction sessions (two from the July study and two from the November study). A stratified random sampling of 157 units of analysis from these four videos was used to calculate an IRR score. The size of our sampling was determined using Lacy and Riffe's method for calculating minimum sample sizes needed to establish reliability [20]. The IRR score for physical engagement was .80 Cohen's Kappa, which indicates substantial agreement [35].

HYPOTHESES

We hypothesized based on the literature (see Related Work) that participants interacting with *TuneTable* would move through three stages of engagement—*initial engagement* (L1—during which participants build an initial mental model of the system), *exploratory engagement* (L2—during which participants test and revise their mental model of the system), and *expressive engagement* (L3—during which participants have a clear mental model of the system and are able to express themselves creatively). The first three of our physical codes (*isolated*, *investigative*, and *integrated* manipulation) reflect this progression in terms of observable physical markers. We do not examine our fourth code, *external* manipulation, in depth in this paper, as it did not contribute to our understanding of stages of engagement.

While all three stages are important components of engagement, we approached our analysis with the following assumptions: 1) Progression from L1 → L2 → L3 indicates that participant groups are more deeply engaging with the exhibit and the concepts embedded within (a similar assumption to that made by [2] and [5]); and 2) Participants have the most tools with which to creatively express themselves when in L3, since they are integrating multiple components of the exhibit during this stage, including both function and sound blocks. In addition, participants are only able to experiment with computational concepts when in L3, since function blocks are not used in L1 or L2. Therefore, a longer time spent in L3 indicates that participants have more opportunities to learn about computing concepts and creatively express themselves using those concepts, two of the design goals of the project.

Based on these assumptions, we hypothesized that an “ideal” trajectory of physical engagement would involve participants moving sequentially through the three stages of engagement, with a significant period of time spent at L3 towards the end of the interaction (we refer to this as the *rising-step trajectory*). We originally thought that it would take groups some time to progress from L1 to L3, and therefore we hypothesized that participants would spend roughly an equal amount of time in L1, L2, and L3.

RESULTS

Plotted Trajectories

We plotted the set of codes for each of the 32 family groups as a separate figure using R [34]. The plots (see Figures 3–5) show the three physical coding categories as three levels (1. Isolated, 2. Investigative, and 3. Integrated) along the

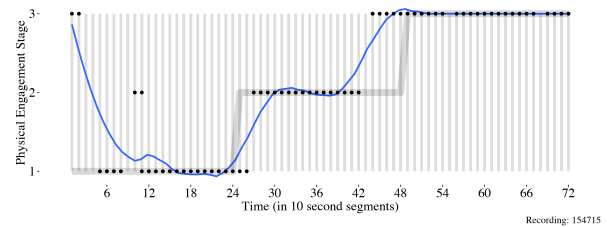


Figure 3. Plot representing the initially-hypothesized rising-step trajectory. This group had a high level of prior coding knowledge. They began their interaction by “stepping back” from computational construction, focusing instead on what they wanted their composition to sound like. The subsequent steps through L2 and L3 involved the testing and confirmation of coding concepts that the group was already familiar with.

y-axis and each time segment along the x-axis. Recall that we analyzed each video in 10-second increments, and each increment is indicated on the plot as a vertical gray bar. We added dots to the gray bar to identify the code associated with each segment. The plot is also divided into quartiles for ease of reference, shown as vertical lines that separate the plot into four equal sections. We added a LOWESS (locally-weighted scatterplot smoother) curve to the plot using a span of 0.3 [37]. We chose this span value because it made the curves responsive to the codes but not so responsive that there was a drastic change or spike in the curve for minor code deviations. Finally, we use a gray trendline to indicate our hypothesized rising-step trajectory as a benchmark for comparison.

Co-Creative Trajectories

When examined individually, each plot represents a unique group experience with *TuneTable* shaped by factors such as the duration of interaction with the exhibit, number and age of participants, and the level of parental involvement. When considered as a group, however, discernible similarities among trajectories emerged across the entire sample. While remaining cognizant of the distinctive features of each plot, we were able to identify several notable patterns that, when considered as possible interaction trends, provide a lens through which we can better appreciate the experiences of participants and their possible modes of creative expression and engagement. In the following sections, we examine the trajectory categories that emerged and what they can tell us about engagement and expression with *TuneTable*.

Rising-Step (Hypothesized)

The *rising-step* trajectory represents our initially-hypothesized “ideal” physical trajectory—a steady, stepped slope from L1 to L2 to L3—with roughly the same amount of time spent in each level and a plateau when reaching L3. Only 1 one of 32 curves fit this trajectory exactly. Six additional curves follow a stair-step trajectory, but with considerably less time spent in L1 and L2 than we had originally hypothesized. Figure 3 exemplifies the rising-step trajectory even though it begins at L3 (integrated manipulation) because a completed set of blocks was left on the table when the group began.

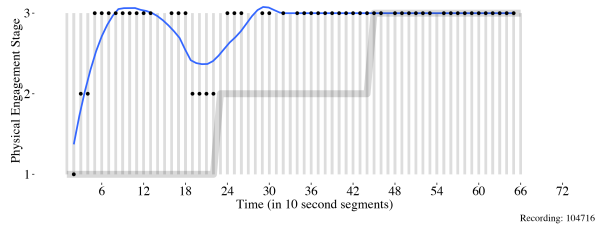


Figure 4. Plot representing the wavering-step trajectory. This group shifted from L3 to L2 when they found that they were unhappy with the particular sound which was repeating—an aesthetic choice. When they returned to L3, they continued to experiment with different function blocks.

Wavering-Step

Wavering-step trajectories often share a general path with hypothesized ones—that is, from L1 or L2 to a plateaued L3—however, they exhibit faster initial slopes from L1 to L3, and contain between one and three identifiable “dips” in their curves (Fig 4). The majority of the dips we see in the wavering-step category are minor and most often from L3 to L2; however, we see variations both in depth (one group moved from L3 to L1 and back over the course of quartile 3) and frequency (another group had three slight dips over the course of quartiles 2 and 3). In all cases, the trajectories of wavering-step curves end at or near L3 (the level at which groups have the most opportunities to experiment with computational concepts), which suggests that the wavering-step trajectory might promote our design goals equally well, if not better, than the rising-step trajectory.

Unsteady

Unsteady trajectories deviate significantly from both rising-step and wavering-step trajectories in several major ways. The strongest characteristic that defines unsteady curves is the high frequency of movement between levels—if rising-step and wavering-step trajectories are sloping hills, unsteady trajectories are jagged mountains. There is often both upward- and downward- movement between levels, and, unlike wavering-step curves, there is a recurring dip-and-swell nature to the curves (Fig 5). Another interesting feature of unsteady plots is the variance between starting-points. 7 of our 9 unsteady plots begin at or below L2. Despite their unpredictable and unsteady general trajectories, it is interesting to note that unsteady curves often finish at strong levels of integration in the fourth quartile, possibly indicating in-depth engagement with the exhibit even with unsteady progression during the first three quartiles.

No Fit

There were three interactions that did not fit in any of our categories. In two of these interactions, participants were at integration (L3) for the entire duration of their experience. In both of these videos, the participants had a reference block as part of their chain for the entire interaction, but never successfully paired it with a sub-reference block. If the reference block had not been present, the trajectory of both videos likely would have more closely resembled a single stair step from L2 to L3 about 4 minutes into the interaction (a closer fit with

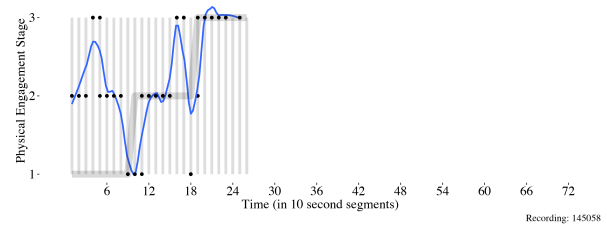


Figure 5. Plot representing the unsteady trajectory. This group experienced confusion due to the appearance of the play button in unexpected locations when they were connecting sound blocks. In response to this, the group decided to test individual sound blocks, resulting in a drop to L1. One of the participants gradually introduced some function blocks, resulting in the rise back to L3.

the rising-step trajectory). This discrepancy likely indicates that the reference block was difficult to understand due to the un-intuitive icon on the block (a set of brackets).

Video Review

We originally hypothesized that groups that followed a steady rising-step growth trajectory would spend a significant amount of time at L3 and demonstrate complex engagement with computational concepts while interacting with *TuneTable*. A review of the social and intellectual activities of participants in the rising-step trajectory category points to complex problem-solving, hypothesis testing, and development across investigative and integrated modes of interaction. For example, during L3, the participants in one rising-step video investigated whether there was a difference in lag time when using the up block vs. the reference/sub-reference blocks to jump to different code segments.

We also found that many participant groups progressed through their interaction in a very different way—in the case of the wavering-step curves, jumping quickly to L3 and then experiencing some small “dips”; in the case of the unsteady curves, bouncing between all three levels for the entirety of the interaction. We decided to look into why so many curves diverged from our hypothesized rising-step trajectory. In particular, we focused on analyzing the moments when participants had reached L2 or L3 but then experienced a “setback”, or a “dip” in their trajectory curve. After reviewing the “setback” moments that occurred in the wavering-step and unsteady trajectories, we identified several main causes of setbacks, which we describe below.

Creative Exploration: Participants who have “figured out” how to use the table want to go back and choose sounds and sound sequences that fit their creative goals.

Debugging: Participants are confused by or have a question about some aspect of the table and they remove function blocks from the table to figure out how simpler chains work. This happened often with groups who put a lot of function blocks on the table at the beginning of an interaction without understanding what the blocks did.

Setback Type	Number of occurrences in unsteady trajectories	Number of occurrences in wavering-step trajectories
Creative Exploration	5/16 (31%)	2/24 (8%)
Debugging	5/16 (31%)	8/24 (33%)
Technical Issue	2/16 (13%)	5/24 (21%)
Thrashing	3/16 (19%)	3/24 (13%)
Boredom	0/16 (0%)	3/24 (13%)
Resolved	7/16 (43%)	9/24 (36%)
Partially Resolved	4/16 (25%)	7/24 (29%)
Unresolved	4/16 (25%)	5/24 (21%)

Table 1. Number of setbacks categorized by trajectory (unsteady or wavering). Setbacks are also categorized by the type or cause of the setback as seen in the video data.

Technical Issue: Some aspect of the table malfunctioned and the participants responded by removing blocks and testing out simpler sound chains.

Thrashing: The participant group is confused about what is happening with the table and is “thrashing” (i.e. connecting random blocks to see what happens) or using the inspector block to read about what blocks do.

Boredom: The participants seem to have reached the end of a creative goal and take all blocks off the table to start over or reformat their chains into new creations. It is likely that in a free-choice environment (rather than a workshop study) they might have left the table at this point.

We additionally categorized relevant setbacks based on whether they were resolved (i.e. participants accomplished what they “went back” for), partially resolved (i.e. participants appeared to gain some knowledge about the table, but did not fully resolve their original problem), or unresolved (i.e. participants did not solve their original problem or demonstrate any apparent knowledge gains).

There were 16 total setbacks in the unsteady trajectories and 24 in the wavering-step trajectories. We categorized 15/16 and 21/24 trajectories, respectively—the four uncategorized setbacks did not fit in any particular category, although we did mark them for resolution. We found three main reasons for these outliers: 1) multiple group participants were working on different levels in parallel; 2) there was a facilitator demonstration that interrupted the interaction; or 3) the participants were exhibiting minimal physical engagement. The “boredom” setbacks were not tagged as being resolved/partially resolved/unresolved, as we did not feel there was a clear resolution to boredom. Table 1 summarizes the number of occurrences of each setback type and resolution status, broken down by unsteady vs. wavering-step trajectory type.

This analysis revealed several key insights. First, *setbacks can be productive and used by participants as a tool for creative exploration and debugging*. The majority (28/40) of

Type	n	Isolated	Investigative	Integrated
All	32	1:12 (13%)	1:58 (24%)	6:13 (70%)
Rising-Step	7	1:34 (18%)	2:19 (29%)	4:49 (56%)
Wavering-Step	13	1:07 (12%)	1:18 (14%)	7:23 (79%)
Unsteady	9	1:22 (16%)	2:53 (36%)	5:27 (63%)
No Fit	3	0:07 (1%)	1:20 (14%)	6:43 (85%)

Table 2. Average time spent in each stage of engagement, organized by the four types of engagement trajectories. Note: Time in each category does not sum to 100% because multiple codes could be associated with each time segment.

the setbacks were resolved or partially resolved in the sense that participants accomplished what they “went back” for. Our analysis also revealed that *the range of concepts being explored by users is perhaps broader than we originally intended*. Some visitors engaged with computing concepts by interacting with function blocks in L3, which was our original design goal. However, other groups engaged with concepts that extended beyond our original intentions. For example, several groups discovered properties of the sound blocks (e.g. thick blocks made longer sounds than thin blocks) by experimenting in L1, and a number of groups investigated how the table hardware worked as a result of technical glitches they had to resolve. We did not notice a significant difference between the types of setbacks/setback resolutions that occurred in unsteady vs. wavering-step trajectories.

Descriptive Statistics

We conducted an analysis of the relationship between trajectory type and a number of different factors (detailed below) to complement our qualitative analysis. These preliminary results do not have sufficient group sizes to make statistical comparisons ($n=32$), but are intended as initial guidance into future research. *Most of the 32 interactions that we analyzed moved along an overall upward trajectory from start to finish (25 of 32)*. Table 2 shows the average amount of time spent at each level of engagement (isolated, investigative, integrated) in minutes:seconds along with the average percentage of time spent at each level.

We hope that participants spend more time in L3 (since participants are able to experiment with the widest range of concepts at this stage, see Hypotheses), and we see that *the wavering-step trajectory yields the greatest amount and percentage of overall time in L3*, and less time in both L1 and L2. This conflicts with our original hypothesis that the rising-step trajectory would yield the most amount of time in L3. We decided to delve deeper and investigate whether participants with wavering-step trajectories reach sustained integration (i.e. remain at the integration level for 1 minute or longer) more quickly than the other groups and whether wavering-step participants who leave the exhibit at the integration level remain in L3 longer than the other groups.

Type	n	Average Time to Sustained Integration
All	30	1:55 (22%)
Rising-Step	7	3:16 (37%)
Wavering-Step	13	1:13 (14%)
Unsteady	9	1:56 (23%)
No Fit	3	1:43 (20%)

Table 3. Average time spent in sustained integration for each trajectory type that was analyzed. *Sustained integration* means that the group spent one or more minute in the integration (L3) level of interaction.

Type	n	Average Time Completing at Integration
All	20 of 32	2:20 (26%)
Rising-Step	4 of 7	1:20 (16%)
Wavering-Step	9 of 13	2:47 (29%)
Unsteady	6 of 9	7:10 (94%)
No Fit	1 of 3	3:45 (49%)

Table 4. Average amount of time for each trajectory type where the participating group ended their interaction in the integration (L3) level.

Overall, **30 of 32 groups reached sustained integration**. On average, participant groups reached sustained integration (1 minute or more spent at the integration level) approximately 2 minutes into their interaction with the exhibit which represents 22% of the overall time spent at the exhibit. Interestingly, ***the wavering-step group reaches sustained integration more quickly than the other groups*** and in approximately one third of the time of the rising-step group (see Table 3).

Out of 32 participant groups, 20 finished the exhibit at L3 (integrated). On average, participants in these 20 groups spent two minutes 20 seconds at this level at the end of their experience, and this time comprises 26% of their overall time with the exhibit on average. In other words, ***almost two-thirds of the participants ended at L3 and spent over a quarter of their overall time in L3***. Table 4 shows that, on average, ***the wavering-step groups were more likely than the rising-step groups to finish at the integrated level and also remained at this level twice as long***.

We analyzed several additional factors about the groups and their interactions with *TuneTable* in order to better understand their trajectories. Specifically, we looked at the number of participants in each group (ranging from 2 to 5 participants), whether blocks from previous groups were on the table at the outset, whether they used the inspector block to learn about the functionality of the other blocks, and the gender distribution of the groups.

The number of participants in each group may have influenced the trajectory of the group. The rising-step group was more likely to have two users (5 of 7) compared to the set-back groups which typically had three or more users (10 of 13). Both the unsteady (5 of 9) and no fit groups (3 of 3)

trended toward three or more users. The rising-step groups were also less likely to use the inspector block (4 of 7) whereas almost all wavering-step, unsteady, and no fit groups used the inspector block. The gender distribution for the rising-step and unsteady participant groups were predominantly male (71% and 64%, respectively) compared to the wavering-step, and no fit groups which were evenly split (male = 49% and 53%, respectively). ***The rising-step group is distinct from the other groups in that they moved linearly through levels one and two to plateau at level 3. This trajectory may be related to their relatively small group size, their low use of the inspector block, and their gender distribution.***

DISCUSSION AND DESIGN IMPLICATIONS

TuneTable Evaluation

Our results generally indicate that the *TuneTable* exhibit supported experimentation with computing concepts as well as co-creative activity. 30/32 participant groups reached sustained integration and 20 groups spent over a quarter of their time in L3 (integration)—indicating that most groups were able to integrate multiple components of the exhibit quickly and therefore had ample time to experiment with function blocks. Our analysis of “setbacks” that occurred during interactions revealed that even when groups were not at L3, the majority were engaging in productive activities such as debugging an issue they were facing or creatively experimenting with different sounds. Different behavior (in particular, shorter interactions) might be observed in-the-wild, but our study has shown that the table as an installation supports prolonged interactions involving inquiry and creative experimentation.

Multiple “Ideal” Trajectories

We hypothesized that participant groups demonstrating ideal engagement trajectories would move along a stair-step trajectory from L1 to L3. However, we found that only seven participant groups actually followed this trajectory. In contrast, most participant group trajectories were in our wavering-step and unsteady categories. We found that the wavering-step groups actually reached L3 the fastest and spent the most overall time in L3. From this, we surmise that a wavering-step trajectory may actually lead to more opportunities for creative expression and experimentation with computational concepts than our original hypothesized rising-step trajectory.

Our video review revealed that many of the setbacks that groups experienced in the wavering-step and unsteady trajectories actually led to fairly complex inquiry and experimentation—using questions generated at L3 to motivate a “step back” to L2 or L1 to debug issues with the system or select sounds that fit their creative goals. While inquiry also occurred in the rising-step trajectory videos, it seems clear that there is more than one “ideal” trajectory that visitors can follow, and our results indicate that factors such as group size or gender distribution might influence group trajectories. This suggests that ***designers may want to consider adopting strategies for accommodating and supporting a variety of family group engagement trajectories, such as allowing for multiple levels and types of engagement [8].***

Setbacks as Opportunities for Deeper Exploration

The wavering-step and unsteady trajectories highlight a design opportunity that has been unrealized in exhibits that are solely designed for apprehendability [1, 15]. A quick ramp up to L3 (as seen in all of the wavering-step and 4/7 of the rising-step trajectories) may indicate that participants were quickly able to integrate multiple components of the exhibit in an effective way. However, the wavering-step and unsteady trajectories indicate that many learners experience challenges they must resolve during the course of their time in L3—either creative challenges that spur them to return to experimenting with sound blocks or technical challenges that cause them to rebuild their mental model of the system.

These challenges may not be a bad thing, especially when they are resolved and allow the group to return to L3 with a refined mental model (as we saw in the case of 42% wavering-step setbacks and 43% unsteady setbacks). Challenge plays an important role in both learning [36] and the creative process [10]—the key is that the challenge has to be at a level that is just beyond the learner’s skill level, but not so far that they get discouraged. Bilda et al. refer to this as *unintended mode*, or the period when participant expectations of the system are unmet and they need to revise their mental model [5]. They suggest that the most learning occurs in this stage, but it is also the stage in which participants are most likely to grow frustrated and disengage [5]. Davis et al. similarly note that during co-creative interactions, individuals alternate between clamped cognition (fluid creative expression) and unclamped cognition (periods of reflection or physical experimentation with the environment) [11]. Unclamped cognition on a group level actually corresponds to *mutual building*, or joint contribution to the generation of new structures in the creative space [11]. This all suggests that *designers of co-creative exhibits may want to consider explicitly designing opportunities for reflection and mental-model revision—rather than just focusing on apprehendability—in order to promote more in-depth creative expression and opportunities for learning.*

It seems that the inspector block that we included in the November 2017 version of *TuneTable* served as an invitation to reflect and engage more deeply with the exhibit—participant groups with unsteady or wavering-step trajectories were more likely to have used the inspector block. We saw in our analysis of setbacks that discovering the inspector block sometimes led to setbacks—most often (eight times), it spurred participants to ask new questions about the blocks, although it should be noted that in four videos, setbacks were caused by participants confusing the inspector block with a function block. In addition, participants who were debugging often used the inspector block to answer questions. This latter point suggests that *designers may also want to consider coupling opportunities for reflection with scaffolding that can help visitors to overcome challenges that may be slightly beyond their skill level* [36].

A number of the “setbacks” we observed were caused by malfunctions with the *TuneTable* prototype software and hardware (18%). While system errors can be frustrating, it is interesting that they can incidentally spur opportunities for reflection,

debugging, and inquiry (also noted by [16]). This relates to the notion of “learning through breakdown”, or the idea that breakdown or failure can serve “as a means of revealing the nature of the world around us” [19]. *Designers seeking to foster deeper engagement could consider designing for “breakdown” in their work as a way of spurring reflection that may be able to lead to opportunities for learning as well as creative contributions.* Of course, it is important to ensure that such breakdowns are more intriguing than frustrating—otherwise, participants interacting in-the-wild may abandon the exhibit.

FUTURE WORK

In this paper, we have provided insight into physical interaction and stages of engagement/expression with one particular exhibit, *TuneTable*. These findings have not yet been generalized to other exhibits. In addition, our study was conducted in a controlled workshop environment within the museum and therefore cannot speak to what physical interaction would look like in a more in-the-wild interaction [27, 14]. We plan to iteratively revise our coding scheme for physical engagement as we apply it to understanding *TuneTable* and other co-creative exhibits in-the-wild. We strive to eventually create a generalizable framework for understanding physical interaction with co-creative museum installations. We are also currently exploring the interplay between physical engagement and other modes of engagement (i.e. intellectual, social, and emotional [18]) that relate to exploration and creative expression. Investigating how social, emotional, and intellectual engagement relate to physical engagement will likely provide a different perspective on the trajectories of individuals interacting with *TuneTable*. Additional evaluation metrics, such as pre/post tests, could also aid in understanding whether learning is occurring.

CONCLUSION

In this paper, we have conducted a deep study of physical interaction with *TuneTable*, looking at how observable physical actions relate to stages of engagement. Our study revealed that *TuneTable* as an exhibit is able to support fairly complex inquiry and creative experimentation with computing concepts. However, our analysis of participant trajectories revealed that some of our initial assumptions about the nature of engagement with co-creative exhibits were wrong. Our analysis also revealed several potential design implications for co-creative exhibits, including the idea that designing to facilitate reflection and mental-model revision can encourage deeper engagement and creative expression. In our future work, we aim to continue to iterate on our coding schema and refine our understanding of the relationship between physical interaction and engagement as we apply our coding scheme to other co-creative installations.

ACKNOWLEDGEMENTS

We would like to thank the following people/institutions, without whom this research would not have been possible: Astrid Bin, Emily Bryans, Mirina Garoufalidis, Katlyn Voravong, Nylah Boone, Anna Xambó, Gerard Roma, Aaron Price, the Museum of Science and Industry Chicago. This project is funded by the National Science Foundation (DRL #1612644).

REFERENCES

- [1] Sue Allen. 2004. Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education* 88, 1 (2004), S17.
- [2] Chantal Barriault and David Pearson. 2010. Assessing exhibits for learning in science centers: A practical tool. *Visitor Studies* 13, 1 (2010), 90–106.
- [3] Mathilde M Bekker, Judith S Olson, and Gary M Olson. 1995. Analysis of gestures in face-to-face design teams provides guidance for how to use groupware in design. In *Symposium on Designing Interactive Systems*. Citeseer, 157–166.
- [4] Zafer Bilda, Linda Candy, and Ernest Edmonds. 2007. An embodied cognition framework for interactive experience. *CoDesign* 3, 2 (2007), 123–137.
- [5] Zafer Bilda, Ernest Edmonds, and Linda Candy. 2008. Designing for creative engagement. *Design Studies* 29, 6 (2008), 525–540.
- [6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [7] Francesco Cafaro, Alessandro Panella, Leilah Lyons, Jessica Roberts, and Josh Radinsky. 2013. I see you there!: developing identity-preserving embodied interaction for museum exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1911–1920.
- [8] Luigina Ciolfi and Liam Bannon. 2002. Designing Interactive Museum Exhibits: Enhancing visitor curiosity through augmented artefacts. In *Eleventh European Conference on Cognitive Ergonomics*. Citeseer.
- [9] Elaine M Crowder. 1996. Gestures at work in sense-making science talk. *The Journal of the Learning Sciences* 5, 3 (1996), 173–208.
- [10] Mihaly Csikszentmihalyi. 2008. *Flow : the psychology of optimal experience* (1st harper perennial modern classics ed. ed.). Harper Perennial, New York.
- [11] Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Brenda Lin, and Brian Magerko. 2017. Creative sense-making: Quantifying interaction dynamics in co-creation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. ACM, 356–366.
- [12] Linda De Valk, Pepijn Rijnbout, Tilde Bekker, Berry Eggen, Mark De Graaf, and Ben Schouten. 2012. Designing for playful experiences in open-ended intelligent play environments. In *IADIS international conference games and entertainment technologies*, Vol. 310.
- [13] Exploratorium. *Coding Scheme for APE Multi/Single Station Research Project*. Technical Report. San Francisco, CA.
- [14] Christian Heath and Dirk Vom Lehn. 2008. Configuring 'Interactivity' Enhancing Engagement in Science Centres and Museums. *Social Studies of Science* 38, 1 (2008), 63–91.
- [15] Michael S. Horn, Erin Treacy Solovey, and Robert JK Jacob. 2008. Tangible programming and informal science learning: making TUIs work for museums. In *Proceedings of the 7th international conference on Interaction design and children*. ACM, 194–201. <http://dl.acm.org/citation.cfm?id=1463756>
- [16] Eva Hornecker. 2005. A design theme for tangible interaction: embodied facilitation. In *ECSCW 2005*. Springer, 23–43.
- [17] Eva Hornecker and Jacob Buur. 2006. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 437–446.
- [18] Thomas Humphrey, Joshua Gutwill, and The Exploratorium APE Team. 2005. *Fostering Active Prolonged Engagement: The Art of Creating APE Exhibits*. Routledge, Abingdon, UK.
- [19] Timothy Koschmann, Kari Kuutti, and Larry Hickman. 1998. The concept of breakdown in Heidegger, Leont'ev, and Dewey and its implications for education. *Mind, Culture, and Activity* 5, 1 (1998), 25–41.
- [20] Stephen Lacy and Daniel Riffe. 1996. Sampling error and selecting intercoder reliability samples for nominal content categories. *Journalism & Mass Communication Quarterly* 73, 4 (1996), 963–973.
- [21] Leilah Lyons, Brian Slattery, Priscilla Jimenez, Brenda Lopez, and Tom Moher. 2012. Don't forget about the sweat: effortful embodied interaction in support of learning. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*. ACM, 77–84.
- [22] Paul Marshall. 2007. Do tangible interfaces enhance learning?. In *Proceedings of the 1st international conference on Tangible and embedded interaction*. ACM, 163–170.
- [23] Joan Mora-Guiard and Narcis Pares. 2014. Child as the measure of all things: the body as a referent in designing a museum exhibit to understand the nanoscale. In *Proceedings of the 2014 conference on Interaction design and children*. ACM, 27–36.
- [24] Lorelli S Nowell, Jill M Norris, Deborah E White, and Nancy J Moules. 2017. Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods* 16, 1 (2017), 1609406917733847.
- [25] Andrew Polaine. 2005. The Flow Principle in Interactivity. In *Proceedings of the Second Australasian Conference on Interactive Entertainment (IE '05)*. Creativity & Cognition Studios Press, Sydney, Australia, Australia, 151–158. <http://dl.acm.org/citation.cfm?id=1109180.1109204>

- [26] Jessica Roberts and Leilah Lyons. 2017. The value of learning talk: applying a novel dialogue scoring method to inform interaction design in an open-ended, embodied museum exhibit. *International Journal of Computer-Supported Collaborative Learning* 12, 4 (2017), 343–376.
- [27] Yvonne Rogers. 2012. Interaction design gone wild: striving for wild theory. *Interactions* 18, 4 (2012), 58–62. <http://dl.acm.org/citation.cfm?id=1978834>
- [28] Marie-Monique Schaper, Maria Santos, Laura Malinverni, and Narcis Pares. 2017. Towards the design of a virtual heritage experience based on the world-as-support interaction paradigm. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2034–2041.
- [29] Orit Shaer and Eva Hornecker. 2010. Tangible user interfaces: past, present, and future directions. *Foundations and Trends in Human-Computer Interaction* 3, 1-2 (2010), 1–137. <http://dl.acm.org/citation.cfm?id=1755234>
- [30] Ben Rydal Shapiro. 2018. *Interaction Geography & the Learning Sciences*. PhD Thesis. Vanderbilt University.
- [31] Peter K Smith. 1985. The Reliability and Validity of One-zero Sampling: misconceived criticisms and unacknowledged assumptions. *British Educational Research Journal* 11, 3 (1985), 215–220.
- [32] Scott S Snibbe and Hayes S Raffle. 2009. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1447–1456.
- [33] Amber Solomon, Mark Guzdial, Betsy DiSalvo, and Ben Rydal Shapiro. 2018. Applying a Gesture Taxonomy to Introductory Computing Concepts. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*. ACM, 250–257.
- [34] R Core Team and others. 2016. R: A language and environment for statistical computing. (2016).
- [35] Anthony J Viera, Joanne M Garrett, and others. 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med* 37, 5 (2005), 360–363.
- [36] Lev Semenovich Vygotsky. 1980. *Mind in society: The development of higher psychological processes*. Harvard university press.
- [37] Hadley Wickham. 2016. *ggplot2: elegant graphics for data analysis*. Springer.
- [38] Niels Wouters, John Downs, Mitchell Harrop, Travis Cox, Eduardo Oliveira, Sarah Webber, Frank Vetere, and Andrew Vande Moere. 2016. Uncovering the honeypot effect: How audiences engage with public interactive systems. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, 5–16.