Conducting Remote Design Research on Embodied, Collaborative Museum Exhibits

Duri Long
Department of Communication Studies, Northwestern University
duri@northwestern.edu

ABSTRACT

Research activities in interaction design and HCI were widely altered by the COVID-19 pandemic, with many studies shifting online as health concerns inhibited in-person research. Tangible and collaborative activities are often used in informal learning spaces and child-computer interaction, but they are neither designed for nor easily adapted to online formats. In this case study, I present findings and reflections on my experience adapting an in-situ study of embodied, collaborative museum exhibits to a remote user study during COVID-19. I identify several considerations and notes of inspiration for researchers working on similar projects, which I hope can aid in furthering iterative design research on embodied and/or collaborative activities both during the ongoing pandemic and in other current and future contexts that require remote research or interactions. The reflections I present in this case study additionally play a role in documenting the ongoing history of interaction design as researchers adapt to the rapidly changing global circumstances caused by COVID-19.

CCS CONCEPTS

• **Social and professional topics** → Professional topics; Computing education; Informal education; • **Human-centered computing** → Human computer interaction (HCI); HCI design and evaluation methods; User studies; Interaction design; Interaction design process and methods; Interface design prototyping.

KEYWORDS

Additional Keywords and Phrases AI literacy, COVID-19, pandemic, remote user study, informal learning, museum exhibits, tangible, embodied, collaborative, methods, design research, at-home learning

ACM Reference Format:
Duri Long. 2023. Conducting Remote Design Research on Embodied, Collaborative Museum Exhibits. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3544549.3573842

1 INTRODUCTION

Research activities in many fields were widely altered by the COVID-19 pandemic, and HCI is no exception. Restrictions on travel, country or state-wide lockdowns, and social distancing disrupted, delayed, or altogether halted many research agendas, affecting research productivity, scholarly collaboration, and the development of early career researchers, amongst other issues [2]. Many within the HCI community demonstrated resilience and ingenuity in response to the pandemic, exploring new research agendas to address the unique set of challenges presented by the pandemic [4]. Searching the ACM digital library for "COVID-19" already yields over 320,000 results, including research on visualizing and tracking health data, investigating online misinformation about the pandemic, improving online learning/work experiences, and addressing mental health issues resulting from the pandemic. Others in the HCI community were able to shift existing research plans to accommodate issues caused by the pandemic, using alternative methodologies such as having users test tools in virtual environments, conducting interviews and workshops via video calls, and administering surveys online [2, 12].

Despite the laudable flexibility demonstrated by HCI researchers during this time, many projects were not easily adapted to the "new normal." Human-centered research necessitates involving stakeholders in research, design, and evaluation, but conducting in-person studies in locations with widespread COVID-19 transmission posed health and safety risks to both participants and researchers. Many factors complicate transitioning to online user studies. For example, conducting remote user studies requires a study population with access to and literacy in a variety of different technologies [12]. Interfaces that require tangible or embodied interaction are not designed for virtual environments and it may be difficult to study them online. More generally, adults and children alike have been overwhelmed with concerns about online work/school, employment, childcare, social decision-making, and health/safety during the pandemic, and it is challenging to figure out how to avoid instigating user studies that add to participants' burdens.

When COVID-19 was declared a global pandemic by the World Health Organization in early March 2020, I found myself confronting many of these issues in my own research. I was investigating how to design embodied, collaborative museum exhibits to promote public AI literacy [15] in audiences without technical backgrounds (e.g., without prior experience in data or computer science). I had several user studies planned during the summer of 2020 to investigate whether the installations we were developing fostered learning and interest development on the museum floor. Museums closed, it was unsafe to travel to our partner museum's location and pivoting to in-person user studies in an alternative location (e.g., the lab) posed health and safety risks for both researchers and study participants. These events led me to explore the meta-research question that I present in this case study: How can we conduct iterative design research on embodied, collaborative

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3573842

learning interventions when we are unable to conduct in-person user studies? In this paper, I present the methods that I used to move forward with my design research during a global pandemic. I describe and reflect on various aspects of this approach, including design challenges, data collection, health and safety, diversity, and more.

In the past two years, many researchers have been able to return to conducting in-person research, thanks to breakthroughs in vaccines and the use of high-quality masks. However, the impact of the pandemic continues to ebb and flow as new variants of the virus develop, many places around the world still have limited access to vaccines, and individuals with comorbidities remain at high risk. The methods presented in this paper can most obviously be of use to design researchers aiming to continue their work building and studying embodied and/or collaborative activities during the pandemic (particularly in learning contexts). Beyond the pandemic, his paper can also be useful to researchers interested in a) studying informal learning/interaction in home environments; b) conducting design research with populations that may not be interested in or able to visit the study location (e.g. disabled communities; communities that may not ordinarily come to a museum, etc.); and c) conducting remote design research on collaborative embodied systems for reasons unrelated to the pandemic (e.g. physical distance from study population, efforts to reduce travel-related carbon emissions, other global or local crises, etc.). Finally, I hope that this case study can play a role in documenting some of the challenges faced by HCI researchers during the COVID-19 pandemic to contextualize research conducted during this time.

2 A CHANGE OF PLANS

In this section, I provide more detailed context about my original plans for the research and how they shifted due to COVID-19. I had originally set out to conduct an iterative design research process to develop a set of museum exhibits to teach family groups about *AI literacy*, which I define as a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace [15]. A key aim of this work was to broaden access to opportunities to learn about AI to both adult and child learners without technical backgrounds, to equip people with the skills and knowledge needed to critically engage with AI technologies they may encounter in their everyday lives. I planned to design *embodied*, *collaborative*, and *creative* activities to foster learning about AI. This choice was based on prior research that suggests that embodiment, collaboration and creativity have the potential to concretize an abstract topic like AI to make it more understandable and foster interest development amongst populations who may not otherwise have an interest in computing [6, 8, 19, 23, 24]. I refer the reader to our other papers on this work for more detail on the motivation behind our study topic and design choices [14, 16, 17].

Prior to COVID-19, I had completed most of the *generative* stage of the design research (i.e., ideation) [10]. I had created a *design workbook* [5] of exhibit ideas, developed several low-fidelity paper prototypes, and conducted a co-design workshop with family groups at the Museum of Science and Industry, Chicago (hereafter MSI, our museum partner in this research) to develop new ideas and gather feedback from stakeholders [14]. In March 2020, I was analyzing data from the workshop and considering which exhibits to develop into higher fidelity prototypes. My intent was to install "pop-up" exhibits on the museum floor at MSI and test them out with family visitors. I planned to use a mixed methods approach to understand visitor engagement with the exhibits, including conducting a learning talk analysis [23] of video and audio data of participant interactions to understand when participants were engaging in dialogue relevant to the learning goals of the exhibit and conducting exit interviews with participants to assess content knowledge gain and interest development.

When COVID-19 was declared a global pandemic, most museums (including MSI) closed, research was suspended, and it was unsafe to conduct in-person studies with families. At the time, it was unclear if these issues would resolve in a few weeks, a few months, or a few years. I considered numerous options for adapting the designs to ensure my research could continue, including developing web-based version of the projects that could be interacted with virtually or building the exhibits I had originally conceived and hoping I would be able to later test them in some limited capacity (e.g., by inviting family groups to individually visit our lab). Ultimately, I decided to move forward with physical–not virtual or web-based–exhibits because of the key design considerations underlying the existing exhibit prototypes. Providing opportunities for *physical interaction* was key to maintaining the focus on embodiment, and I was concerned that opportunities for in-the-moment *collaboration* would be restricted in a web-based tool where a single user would control the interaction device [7]. I did not want to take the risk of developing exhibits that we may not be able to test with users (even a lab-based study seemed uncertain, as at the time we were unable to visit or access our university lab space), so I needed to pursue a scaled-down alternative to the originally conceived exhibits. Finally, I sought to adapt the designs in such a way that I could easily scale them up later to install and evaluate them on the museum floor using similar methods. I felt that conducting isolated user studies in a lab-based environment would not be conducive to this transition, as prior research has shown that findings from lab studies often do not translate to studies on the museum floor [9].

When I was considering how to redesign the user studies in March 2020, there was almost no published research on how to conduct design research during a pandemic (except for a few blog posts discussing virtual alternatives to in-person studies [1, 18]). As a result, I took inspiration from two sources when considering how to engage families in collaborative, embodied interaction in their homes—1) cultural probes and 2) at-home learner activity boxes.

Cultural probes are "collections of evocative tasks designed to elicit responses that can provide insight into peoples' activities, concerns and values" [3]. Users are encouraged to collect their own data, using devices like disposable cameras, audio recorders, and journals. In addition to providing a technique for collecting data remotely, cultural probes also addressed several other design concerns. Most tools used in cultural probe kits are low-tech, enabling use by participants without high technical literacy or access to expensive technologies. Cultural probes are also playful by design [3], involving engaging prompts and design, which was well-suited to my goal of providing an experience that was fun for participants rather than increasing pandemic-related stress.

I also drew inspiration from educational activity boxes. While not widely discussed in published literature, various companies have developed activity kits that can be delivered to families' homes in a box. For example, TinkerCrate by Kiwi Co. allows families to purchase monthly subscription boxes with science and engineering activities inside (e.g., in one box, learners can build a walking robot using provided materials). I took inspiration from these activity boxes as a way for learners to engage with educational topics in a collaborative and tangible way from home.

I ultimately decided to pursue an approach that I call "Exhibits-in-a-Box"—I selected three exhibits and reformulated their design so that they would fit in a box. I recruited families living in the Atlanta area (my home at the time) and delivered the boxes to their doorsteps. Families interacted with the exhibits in an embodied, collaborative way within their homes, and I evaluated them using similar methods to those I had originally proposed. This way, I was able to preserve key design considerations, utilize similar evaluation methodologies to those that would be used on the museum floor, and ensure families would be able to engage with exhibits in a comfortable, informal leisure environment. The ability for families to engage informally with the exhibits without supervision in their own homes created a study environment more akin to the museum floor than an in-lab study. The following sections detail how I adapted each exhibit design for this context.

3 THE EXHIBITS

3.1 Selection Process

I had developed numerous ideas for museum exhibits related to AI literacy. I developed three into box-sized exhibits, selecting activities that could be adapted to a smaller size without reducing the experience. In addition, I carefully considered the necessary technology. Could it be easily set up by someone unfamiliar with it? Was there enough space to set it up and interact with it? Was it portable? How expensive was it? For example, I did not move forward with an idea proposed by workshop participants of an exhibit where learners could train an AI to recognize different smells, since smell sensors ("electronic noses") are quite expensive. Similarly, I did not pursue an exhibit idea I had developed in which visitors could write a poem with an AI that used natural language processing and a robotic arm to write, as I felt a key component of the embodied experience was the act of creating a story on the same sheet of paper as a robot (which would not be easy to package into a box). The following sub-sections describe in detail the three exhibits that I chose to adapt.

3.2 Knowledge Net

The *Knowledge Net* activity as originally conceived was intended to engage visitors in building semantic networks that describe concepts of interest to them (e.g., their family, animals) using strings, pins, and a corkboard (Figure 1). Different string colors indicate different relationships (e.g., is, has, likes, dislikes). Visitors can take a picture of their creation and it will be parsed by a computer. The visitor can then ask the computer questions about the semantic network that they created (e.g., Q: "What does a cat have?" A: "A cat has fur"). By building a network and asking the AI questions about it, participants can explore ideas including the limitations of knowledge representations, understanding how agents make decisions, the differences between human and computer intelligence, and what it means to "know" or "understand" a concept.

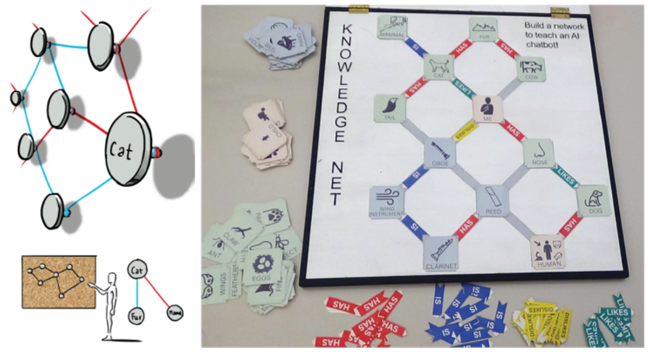

Figure 1: Knowledge Net. Original design sketch (left) and final adapted "box-sized" implementation (right)

I adapted the large-scale tangible interaction with the strings and corkboard into a more board-game-like activity, in which users engaged with tiles and arrows to build a semantic network on a playmat or board. I experimented across iterations with materials, settling on paper tiles and a wooden board to facilitate better image recognition and easy disinfecting/recycling of exhibit components between user groups. I preserved the interaction of having users photograph their networks. In an earlier iteration of the project, I had learners photograph their networks using their personal smartphones. In a later iteration, I shifted to using a provided iPad and Osmo device (which allows the iPad camera to look down at a game board) to facilitate a more collaborative viewing experience and more accurate image recognition.

3.3 Creature Features

The original *Creature Features* exhibit was conceived as an interactive table consisting of a network of nodes connected by elastic strings. Each node in the network represents a feature and putting physical weight on that feature corresponds to training the network on more data with that feature, thereby placing an emphasis on that feature in the network. In the design sketch shown in Figure 2, the neural network is a bird classifier. Placing more weight on a feature like "ability to fly" would mean that the network is trained on mostly examples of flying birds, causing birds like kiwis to be misclassified as "not a bird." Visitors can explore different inputs and weights to see what outputs result. This was intended to help learners draw on knowledge of their body and weight to make sense of abstract concepts like dataset bias.

I adapted the table-sized interaction into a card-based activity. First, based on user feedback, I altered the weight metaphor from the original design. I was asking participants to place weight on features when really the weight occurred as a by-product of biased training data (e.g., more pictures of birds with wings in the training dataset yields more metaphorical weight on wings as a feature). To resolve this point of confusion and scale down the size, I redesigned the activity to focus on curating a dataset for a feature-based machine learning algorithm. In the final activity, learners can

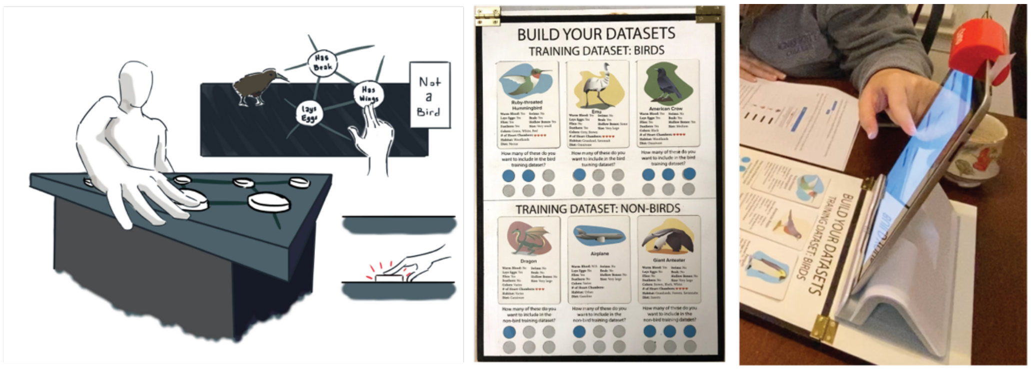

Figure 2: Creature Features. Original design sketch (left) and final adapted "box-sized" implementation (right)

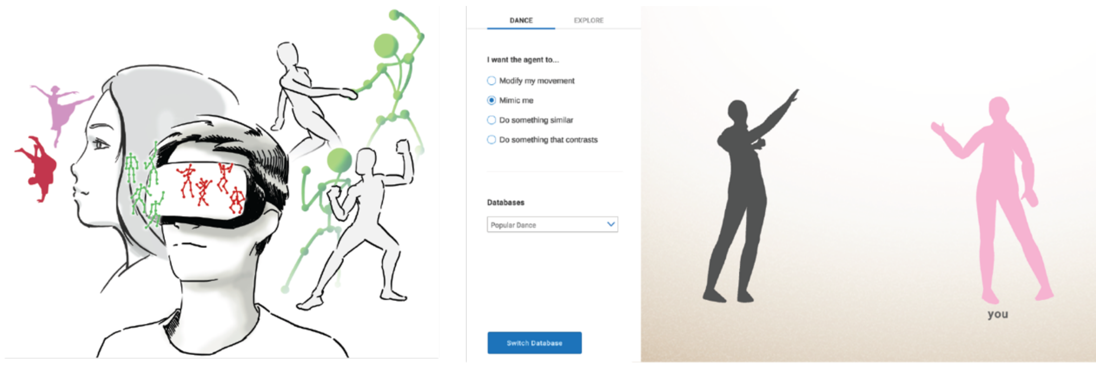

Figure 3: LuminAI. Original design sketch (left), adapted interactive menu/dance interface (right)

use a card deck and "weight tokens" to provide training data to a feature-based ML algorithm that classifies animals as birds. Each card depicts a creature (e.g., bluebird, flying fish) and includes a list of features describing that creature (e.g., swims, has feathers). Learners are encouraged to look at the features for each creature on the playmat and consider how to place their weights to train an algorithm that can correctly recognize many different types of birds. The more weight tokens placed on a card, the more examples of that bird will be included in the dataset. Learners can take a picture of their playmat (using a smartphone or an iPad + Osmo device) and upload it to a website. The results of an algorithm trained using their dataset are then shown to the learners, who are encouraged to iterate on their dataset.

3.4 LuminAI

The *LuminAI* exhibit (Figure 3) builds on an existing AI installation in which participants can improvise movement together with an AI dance partner that is projected onto a screen. A "shadow" of the user's body (generated from a motion capture sensor) is projected on the screen next to an AI dance partner, which senses participant movements via a motion capture sensor and responds with movements in its memory that it deems to be similar. I originally proposed expanding *LuminAI* into an educational exhibit in which participants can move between multiple interaction stations to create a personally customized AI dance partner and learn about the AI dancer's gesture memory. Participants would be able to switch between datasets of different genres of dance moves (e.g., hip-hop, ballet) to change the AI agent's knowledge of dance. Participants could put on a VR headset and use a tool to explore a visualization of the gestures in a chosen dataset, which are clustered based on their similarity (e.g., a right-hand wave and left-hand wave would be clustered together) [13]. Visitors could then dance with an AI dancer (projected onto a screen or wall) that is trained on the dataset they selected. For the scaled-down activity, I developed a user interface that could be displayed on a single laptop screen. Learners could toggle between different views and use interactive menus to explore the agent's gesture memory, turn on/off different response modes, switch out databases, and see how that affected the dance interaction. The boxed activity still used a Microsoft Kinect motion sensor, which was the most complicated piece of technology that I sent participants. However, I provided detailed instructions and participants were able to set up the Kinect in their homes without issue.

4 THE BOXES

Each family was given a set of two exhibit boxes to try out, in addition to a box labeled "Open this box first!" containing data collection instruments and instructions (Figure 4). Per IRB requirements at the time, I either disinfected and/or recycled all box components between user groups. I wanted the boxes and enclosed materials to

Figure 4: Boxes (from left to right). Open this box first!, Knowledge Net, Creature Features, LuminAI

be easy to understand, inviting, and fun. I chose to separate each exhibit into its own box and the introductory instructions and data collection materials into a separate box. I labeled all boxes and was able to easily refer to them in the instructions. This minimized confusion, kept components organized, and aided in making the activities easy to follow. The enclosed instructions were intended to provide about as much data as a participant would find when walking up to an exhibit in a museum. I provided detailed instructions on how to set up and use the exhibit (something that visitors might gather visually from observing others in a museum) but kept explanatory content-related text to a minimum (i.e., the amount that you might find on a sign next to the exhibit).

5 DATA COLLECTION

In this section, I reflect on the methods that I used to collect data in this remote context and successes/challenges faced. I recruited families using a variety of methods, including posting on NextDoor (and having friends/colleagues share the post in their neighborhoods), social media, and coordinating with Georgia Tech's education outreach program. I received the most responses from NextDoor posts and posts shared in social media "parents' groups." However, most of the responses I received for the first iteration of studies were from white families living in more affluent neighborhoods in the Atlanta metro area. To reach a more diverse audience in the second iteration of studies, I reached out to local organizations including the public library, the YMCA, Boys and Girls Clubs, and local Girls Who Code (GWC) club leaders. Local GWC leaders shared the study information, resulting in a more diverse population (in terms of race and socioeconomic status) for the second round of studies. I am still exploring how to mitigate the issue of participant self-selection due to a prior interest in science and technology, which was a challenge in both my in-museum and at-home studies.

Participants contacted us with information on participating family members. I asked about several requirements for study completion (i.e., a stable internet connection, a 4x4' open space to dance for LuminAI, a smartphone that could upload a photo to a website) and asked for scheduling details regarding a good time for box delivery/pickup and study completion. I also asked adults participating in the study to complete a consent form online prior to delivering the boxes. Parents had a clear idea of what was involved in the study prior to delivery. Families were compensated with $40 for their time. The compensation was included in the delivered boxes. I increased participant compensation from the original $20 I was planning for the in-museum study to account for the increased amount of interaction time.

I dropped off the boxes on the participants' doorstep at the scheduled time, then called via either video or phone call (participant preference) at the designated study time. I briefly explained the study, pointing out key materials, and asked participating children for assent (verbal for children under 11; written for children 11 and up). Then, participants were given the option to have me stay on the call to answer questions or for me to hang up and be readily available should the participants need to ask a question. For calls that I stayed on, I took on the role of an observer, watching quietly and only answering questions when asked so as not to unduly influence the interaction.

I collected audio and video recordings of participant interactions as well as survey data from all participants ages 7 and up. Since I was not present in the participant's homes, I had to ask participants to record their own data. I wanted to preserve the recording of both audio and video data, since research indicates that the visual and auditory context of a group interaction is key to understanding the nature of embodied learning at the exhibit [23]. However, I wanted to avoid making assumptions about study participants' access to technology, so I sought to avoid having participants collect data on their personal devices (e.g., smartphones, computers). This choice was also informed by the length of the study—I intended for participants to interact with exhibits for approximately two hours in total, and a two-hour long video/audio recording would take up a lot of storage space on a personal smartphone device, not to mention causing lengthy upload times/a hassle for participants. As a result, I included devices for video and audio recording in the boxes that I delivered to participants.

I anticipated running into some technical difficulties when asking participants to record their own data, so I introduced redundancy. I had participants record their interactions using both a video and an audio recorder, so that if one failed, I would at least have audio of the group's interactions from the other device. When present, I additionally recorded the video call using the video conferencing software. Similarly, the written instructions prompted participants to verbally discuss several questions intended to assess their learning in addition to completing the surveys and aimed to qualitatively analyze learning talk from the audio/video data if the survey data was insufficient or incomplete.

I included an easy-to-use audio recorder with extra batteries. Finding a suitable video recorder was more challenging. I did not want to assume a high level of technical literacy amongst participants, or to have participants waste much time on learning how to use a video recorder, so I wanted to find an easy-to-use camcorder where participants could simply turn on the camera, adjust its position to record their playspace, and hit record. However, due to the commercial popularity and camera quality on smartphones, most easy to use personal camcorder companies have gone out of business.

I considered using GoPro like cameras, which are easy to use, but these cameras typically do not include a screen and therefore provide no visual feedback to participants to help them position the camera appropriately. I also considered using document cameras, but these typically need to be hooked up to a computer to work properly. I ended up purchasing several used Flip cameras, which are simple personal camcorders with an extremely easy-to-use interface that went out of business in 2011. I included a camera stand that allowed the camera to record from either a top-down view or a straightforward view. I included a detailed instruction packets telling participants how to set up and record data using the recording devices.

However, two of the cameras consistently died after 10-15 minutes of recording. This caused difficulties for the participants and resulted in some missing video data. I later replaced the two faulty Flip cameras with GoPro-like cameras. These worked more consistently but, as anticipated, the lack of screen feedback meant that some groups thought their camera was recording when in fact it was not. Ultimately, I did not find an adequate solution for the camera. In the future, I may consider using or creating a custom tool specifically designed for research purposes (e.g., [3]).

While interviews have potential to capture more in-depth data and are well-suited for young children, I shifted the assessment method to surveys for several reasons. First, I wanted to provide participants with the option to not join a synchronous live video call during their interaction to ensure this was a fun activity and did not feel like work/school. This made it challenging to conduct an interview after each activity. Second, I was concerned that even if participant joined a call after each interaction, I would not be able to assess individual learning outcomes since all participants would be answering as a group and would potentially overhear or answer for each other. As a result, I included paper surveys (again, avoiding unnecessary technical barriers) in each activity box for completion after the activity for all participants ages 7 and up. I developed surveys for different age groups with age-appropriate questions based on prior research [22]. There were drawbacks to the choice of using surveys instead of interviews. Some of the kids with shorter surveys got distracted or bored while their parents completed surveys. Some participants left items incomplete. Many young kids (under 8) did not enjoy completing the survey, and some struggled with literacy issues. I tried to mitigate this by encouraging parents to help their kids complete the survey. However, in some cases the survey still felt like a test.

6 INSPIRATIONS AND CONSIDERATIONS

I conclude this case study with a set of considerations and notes of inspiration for other researchers seeking to adapt designs and user studies to remote contexts. These notes are not intended to be a prescriptive list of "must-dos," but rather a collection of prompts to spur creative and critical thought about how to adapt research/designs to a new context. Results from the studies are not discussed in this paper, as the scope of this case study is focused on how the designs and data collection process were adapted during the pandemic. However, I refer the interested reader to [16, 17] for our findings.

Examine the Artifact: A significant part of the process I went through to adapt the designs involved revisiting possible exhibit designs, considering their affordances and requirements, and selecting activities that I thought I could adapt for this new context while preserving the core interaction mechanisms. *Consider:* Consider the affordances of your activity/artifact you have designed or are planning to design. Can you scale it down, using smaller portable components and inexpensive, easy to use technology while still preserving the key elements of the interactive experience?

Return to Early Prototypes: When I was considering how to adapt the design sketches of full-scale museum exhibits to a box-sized exhibit, I returned to early paper and low-fidelity prototypes that I had developed for the exhibits. Returning to these prototypes helped me identify the core interaction mechanisms I wanted to preserve and helped me to consider how to deliver these to participants without using expensive technology. *Consider:* Did you develop low fidelity or paper prototypes of your designs? Consider returning to early prototypes as inspiration for designing "scaled down" activities. If you did not develop low-fidelity prototypes, consider how you might create one (or—even better—actually make one!) for your design to explore how to capture the core interaction with low-cost materials and minimal technology.

Make it Easy: One of my priorities was to create activities with easy-to-use technology. I wanted to ensure that participants did not spend their time figuring out how to set up the activity but were instead able to focus on the content. This principle also influenced my choice of data collection instruments and the detailed instructions I provided. One participant commented, "the fact that this experiment could be done so easily without direct researcher help in setup is pretty amazing." *Consider:* Can you find technology (for your activities and data collection instruments) with easy-to-use interfaces and simple setup? Are there any aspects of the study you can have participants complete in advance (e.g., consent forms, preparing the space)? Consider writing your instructions as a list of numbered steps with only one action item per step to make the process easy to follow. Consider testing out your kits with people who are unfamiliar with the activities (and who are not technology experts) to see if the instructions are easy to follow and to anticipate possible breakdowns.

Make it Portable: A key affordance of the Exhibits-in-a-Box is that they were portable. This allowed me to easily deliver and pick them up from families. Portable boxes also allow for activities to be mailed if that makes more sense for the research team. *Consider:* How can you scale down the size of your designs to make them portable? Consider the weight of activity components, the size of boxes needed, and how they might be transported and stored. Also consider the packaging—how can you ensure it is easy to transport, adequately protects the equipment within, and is easy to open?

Prioritize Fun, Excitement, and Mystery: I prioritized making the designs and data collection experience fun and engaging, especially since I was asking families to engage in a user study during an already challenging and stressful time. Several families commented that the anticipation of receiving the boxes and waiting to see what was inside was exciting for the kids, and that coming home from school (or a day of online school) to a set of surprise activities was fun. *Consider*: How can you make your user study fun? Can you use study methods that do not feel like tests or extensions of the school/workday? How can packaging and presentation foster excitement and mystery surrounding the activities?

What if. . .we didn't use Zoom? Zoom (and other video conferencing platforms) has been invaluable in helping people adapt to remote work, schooling, and gatherings during the pandemic. However, "Zoom fatigue" is a widespread phenomenon [20], and because video conferencing is used so widely in workplaces and school environments, many people are eager to step away from the screen during their leisure time. Asking families to engage in a video call can introduce yet another piece of technology into the set-up process and video calls in homes can potentially make families feel like they are being surveilled. More natural interactions (including inquiry-based learning in which participants engage in dialogue with each other rather than asking the researcher) might be observed if the researcher were not watching over a video call throughout the interaction. *Consider:* In the spirit of creating user studies for leisure spaces, consider how you might conduct your study without or with minimal use of video conferencing. Do you need to be virtually present during the entire study, or just for part of it? Could you support participants over a phone call rather than a video call? Can you provide participants with video conferencing options depending on their preferences? What portion of the activities might be more fun or more naturalistic if the researcher were not watching live over video?

Redundancy, Redundancy, Redundancy: As discussed in Data Collection, I anticipated running into technical difficulties when asking participants to record their own data, so I introduced redundancy to ensure I collected useful data. The policy of redundancy ended up being very helpful, as some users experienced technical difficulties with the cameras. Having multiple streams of data ensured I was able to capture full audio recordings for almost all participant groups. *Consider:* How can you introduce redundancy in your data collection process (without overburdening participants)? Can you anticipate which aspects of data collection users might struggle with? Consider testing out your kits with friends, families, and colleagues before starting your official user study to see if they understand the instructions or if they encounter unforeseen technical difficulties. Choose data collection technology with easy-to-use interfaces.

Parallel Play: I created three identical sets of boxes for each exhibit, which allowed me to conduct three user studies in a week. This was invaluable in keeping the study time to a manageable length, especially given the long turnaround time involved in delivering a box to a family, waiting for them to engage with it, and scheduling a time to pick it up. *Consider:* Can you make copies of your designs to allow several studies to take place in parallel? How long do you think it will take for one family to complete your study? How long will it take to deliver and pickup boxes (either yourself or via the postal service)? Plan to ensure you have enough time since this method takes longer than a traditional user study.

Diversify Recruitment Methods: As discussed in Data Collection, I had to alter the methods I was using to recruit users to reach a user study population more representative of Atlanta. Recruiting diverse study populations is not a problem unique to the pandemic [21], but COVID-19 has made recruitment more challenging. Researchers may need to explore a wider range of recruitment methods than usual to reach diverse user groups. *Consider:* Who is the audience using the platforms you are recruiting on? What reach (geographically or community-wise) do your recruitment methods have? What alternative methods could you use to ensure your user population is representative?

Find a Silver Lining: Although I was disappointed that I was unable to realize the exhibits I had originally envisioned, adapting them into Exhibits-in-a-Box had unexpected positive outcomes. Using a range of recruitment methods, I was able to recruit a diverse set of families to engage with the activities (in terms of race and socioeconomic status). This is not always easy to do in museums, where transportation costs and entrance fees can limit access (although targeted outreach programs and field trips can mitigate these issues). Adapting the exhibits for at-home use also forced me to utilize inexpensive, easy-to-use technology and materials. Often in museum spaces, exhibits are designed to inspire awe as visitors interact with technology they have never seen or experienced before [11]. However, "awe-inspiring" hardware (e.g., tangible tabletops, high-definition projection and motion sensing, VR headsets) can be expensive and not easily accessible outside of the museum. I was challenged to rethink how to provide engaging learning experiences without expensive technology. Making the exhibits portable and inexpensive allowed us to reach a wider audience, with the potential for broader outreach in the future. I see this as a "silver lining" and as a call to consider developing more portable, inexpensive versions of exhibits in the future to broaden access to learning opportunities. *Consider:* What are some possible positive outcomes of scaling down your designs or making your study remote? How can you ensure that you take advantage of/maximize these positive outcomes? What lessons from this experience can you translate to your research more broadly?

7 CONCLUSION

In this paper, I present a case study of how I adapted a design research project involving embodied, collaborative museum exhibits to be conducted in a remote context due to the COVID-19 pandemic. I reflect on the successes and challenges of this process and identify a set of considerations/notes of inspiration for others seeking to adapt their projects to remote contexts, whether due to COVID-19 or other concerns such as reaching users with disabilities, international audiences, reducing carbon footprints, or dealing with other global and local crises.

ACKNOWLEDGMENTS

I would like to thank Dr. Brian Magerko for advising this work; Anthony Teachey for helping with instrument development and data analysis; Aadarsh Padiyath, Jonathan Moon, Lucas Liu, Cassandra Naoimi, and Rhea Laroya for their contribution to prototype development; all of our study participants for participating in this study during a challenging time; and our collaborators at the Museum of Science and Industry, Chicago for their ongoing support of this research. This work was funded by the National Science Foundation (DRL #1612644).

REFERENCES

[1] Schmidt Albrecht and Alt Florian. 2020. Evaluation in Human-Computer Interaction - Beyond Lab Studies (Working Paper). Retrieved from https://amp.ubicomp.net/users-off-limits/
[2] Priscilla Balestrucci, Katrin Angerbauer, Cristina Morariu, Robin Welsch, Lewis L Chuang, Daniel Weiskopf, Marc O Ernst, and Michael Sedlmair. 2020. Pipelines Bent, Pipelines Broken: Interdisciplinary Self-Reflection on the Impact of COVID-19 on Current and Future Research (Position Paper). In 2020 IEEE Workshop on Evaluation and Beyond-Methodological Approaches to Visualization (BELIV), IEEE, 11–18.
[3] Andy Boucher, Dean Brown, Liliana Ovalle, Andy Sheen, Mike Vanis, William Odom, Doenja Oogjes, and William Gaver. 2018. TaskCam: Designing and testing an open tool for cultural probes studies. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1–12.
[4] Peter Dalsgaard. 2020. HCI and interaction design versus Covid-19. Interactions 27, 4 (2020), 59–59.
[5] William Gaver. 2011. Making spaces: how design workbooks work. In Proceedings of the SIGCHI conference on human factors in computing systems, ACM, 1551–1560.
[6] Mark Guzdial. 2013. Exploring Hypotheses about Media Computation. In Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research, ACM, San Diego, CA, USA, 19–26.
[7] 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.
[8] Michael S Horn. 2018. Tangible interaction and cultural forms: Supporting learning in informal environments. Journal of the Learning Sciences 27, 4 (2018), 632–665.
[9] Eva Hornecker and Luigina Ciolfi. 2019. Human-computer interactions in museums. Synthesis Lectures on Human-Centered Informatics 12, 2 (2019), i–171.
[10] Jon Kolko. 2011. Exposing the magic of design: A practitioner's guide to the methods and theory of synthesis. Oxford University Press.
[11] Sheila Krogh-Jespersen, Kimberly A Quinn, William Leo Donald Krenzer, Christine Thi Nguyen, Jana Greenslit, and Aaron Price. 2020. Exploring the Awe-some: Mobile eye-tracking insights into awe in a science museum. (2020).
[12] Kucirkova, Cecilie Evertsen-Stanghelle, Ingunn Studsrød, Ida Bruheim Jensen, and Ingunn Størksen. 2020. Lessons for child–computer interaction studies following the research challenges during the Covid-19 pandemic. International journal of child-computer interaction 26, (2020), 100203.
[13] Lucas Liu, Duri Long, and Brian Magerko. 2020. MoViz: A Visualization Tool for Comparing Motion Capture Data Clustering Algorithms. In Proceedings of the 7th International Conference on Movement and Computing, 1–8.
[14] Duri Long, Takeria Blunt, and Brian Magerko. 2021. Co-Designing AI Literacy Exhibits for Informal Learning Spaces. In Accepted to Proceedings of The 24th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW).
[15] Duri Long and Brian Magerko. 2020. What is AI Literacy? Competencies and Design Considerations. In Proceedings of the 2020 ACM Conference on Human Factors in Computing Systems (CHI 2020), ACM, Honolulu, Hawaii. DOI:https://doi.org/10.1145/3313831.3376727
[16] Duri Long, Aadarsh Padiyath, Anthony Teachey, and Brian Magerko. 2021. The Role of Collaboration, Creativity, and Embodiment in AI Learning Experiences. In Proceedings of the 13th ACM Conference on Creativity and Cognition, 1–10.
[17] Duri Long, Anthony Teachey, and Brian Magerko. 2022. Family Learning Talk in AI Literacy Learning Activities. In Proceedings of the 2022 ACM Conference on Human Factors in Computing Systems.
[18] Lupton (Ed.). 2020. Doing fieldwork in a pandemic (crowd-sourced document). Retrieved from https://docs.google.com/document/d/1clGjGABB2h2qbduTgfqribHmog9B6P0NvMgVuiHZCl8/edit?ts=5e88ae0a#
[19] Brian Magerko, Jason Freeman, Tom McKlin, Mike Reilly, Elise Livingston, Scott McCoid, and Andrea Crews-Brown. 2016. EarSketch: A STEAM-Based Approach for Underrepresented Populations in High School Computer Science Education. ACM Transactions on Computing Education (TOCE) 16, 4 (2016), 14. DOI:https://doi.org/10.1145/2886418
[20] Hadar Nesher Shoshan and Wilken Wehrt. 2022. Understanding "Zoom fatigue": a mixed-method approach. Applied Psychology 71, 3 (2022), 827–852.
[21] Ihudiya Finda Ogbonnaya-Ogburu, Angela DR Smith, Alexandra To, and Kentaro Toyama. 2020. Critical race theory for HCI. In Proceedings of the 2020 CHI conference on human factors in computing systems, 1–16.
[22] C Read and Stuart MacFarlane. 2006. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In Proceedings of the 2006 conference on Interaction design and children, 81–88.
[23] Jessica Roberts and Leilah Lyons. 2017. Scoring Qualitative Informal Learning Dialogue: The SQuILD Method for Measuring Museum Learning Talk. . Philadelphia, PA: International Society of the Learning Sciences.
[24] Linda L. Werner, Brian Hanks, and Charlie McDowell. 2004. Pair-programming helps female computer science students. Journal on Educational Resources in Computing (JERIC) 4, 1 (2004), 4.