4-2022

# ASTEROIDS: Exploring swarms of mini-telepresence robots for physical skill demonstration

Jiannan LI
*Singapore Management University*, jiannanli@smu.edu.sg

Maurício SOUSA

Chu LI

Jessie LIU

Yan CHEN

*See next page for additional authors*



## Citation

LI, Jiannan; SOUSA, Maurício; LI, Chu; LIU, Jessie; CHEN, Yan; BALAKRISHNAN, Ravin; and GROSSMAN, Tovi. ASTEROIDS: Exploring swarms of mini-telepresence robots for physical skill demonstration. (2022). *CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, USA, April 22-29*. 1-14.
**Available at:** https://ink.library.smu.edu.sg/sis_research/8022

## Author

Jiannan LI, Maurício SOUSA, Chu LI, Jessie LIU, Yan CHEN, Ravin BALAKRISHNAN, and Tovi GROSSMAN

# ASTEROIDS: Exploring Swarms of Mini-Telepresence Robots for Physical Skill Demonstration

Jiannan Li
jiannanli@dgp.toronto.edu
Department of Computer Science, University of Toronto
Toronto, Ontario, Canada

Maurício Sousa
mauricio@dgp.toronto.edu
Department of Computer Science, University of Toronto
Toronto, Ontario, Canada

Chu Li
chuchu.li@mail.utoronto.ca
Department of Computer Science, University of Toronto
Toronto, Ontario, Canada

Jessie Liu
heng.liu@mail.utoronto.ca
Department of Computer Science, University of Toronto
Toronto, Ontario, Canada

Yan Chen
yanchen@dgp.toronto.edu
Department of Computer Science, University of Toronto
Toronto, Ontario, Canada

Ravin Balakrishnan
ravin@dgp.toronto.edu
Department of Computer Science, University of Toronto
Toronto, Ontario, Canada

Tovi Grossman
tovi@dgp.toronto.edu
Department of Computer Science, University of Toronto
Toronto, Ontario, Canada

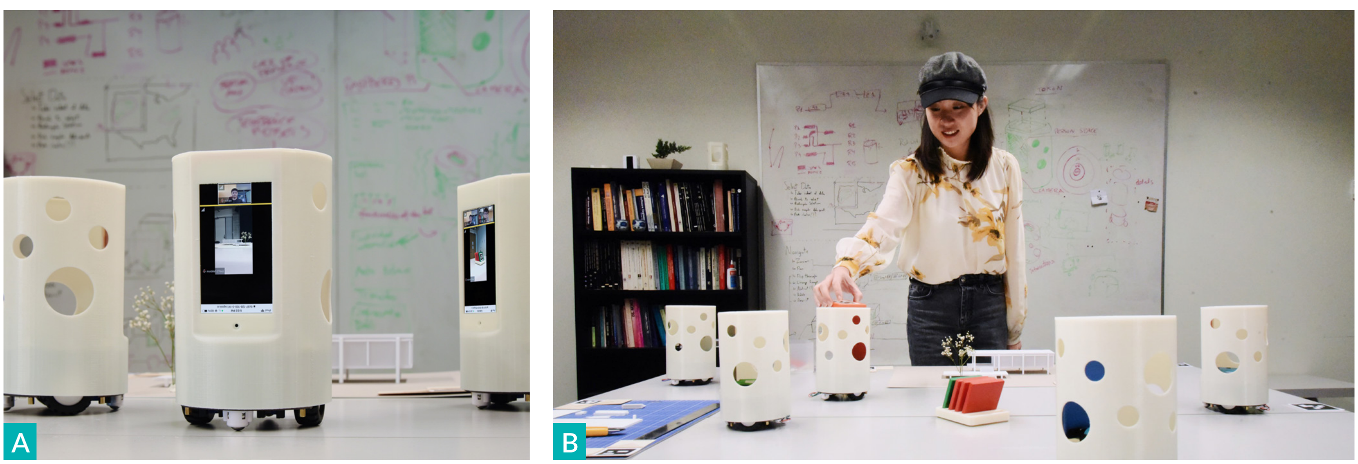


**Figure 1: Asteroids is a novel approach for remote physical demonstrations using a swarm of telepresence robots. A) With Asteroids, remote audience members can inhabit and control small robots on a workbench to follow the instructor's guidance or roam around looking at activities at various locations and scales. B) And, a demonstrator can physically interact with the remote audience and use tangible artifacts to control the flow of the demonstration.**

## ABSTRACT

Online synchronous tutoring allows for immediate engagement between instructors and audiences over distance. However, tutoring physical skills remains challenging because current telepresence approaches may not allow for adequate spatial awareness, viewpoint control of the demonstration activities scattered across an entire work area, and the instructor's sufficient awareness of the audience. We present Asteroids, a novel approach for tangible robotic telepresence, to enable workbench-scale physical embodiments of remote people and tangible interactions by the instructor. With Asteroids, the audience can actively control a swarm of mini-telepresence robots, change camera positions, and switch to other robots' viewpoints. Demonstrators can perceive the audiences' physical presence while using tangible manipulations to control the audience's viewpoints and presentation flow. We conducted an exploratory


*CHI '22, April 29-May 5, 2022, New Orleans, LA, USA*

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3501927

evaluation for Asteroids with 12 remote participants in a model-making tutorial scenario with an architectural expert demonstrator. Results suggest our unique features benefitted participants' engagement, sense of presence, and understanding.

## CCS CONCEPTS

• **Human-centered computing** → **Collaborative interaction**; *Interactive systems and tools.*

## KEYWORDS

Telepresence, Collaboration, Robots, Physical Skill

**ACM Reference Format:**
Jiannan Li, Maurício Sousa, Chu Li, Jessie Liu, Yan Chen, Ravin Balakrishnan, and Tovi Grossman. 2022. ASTEROIDS: Exploring Swarms of Mini-Telepresence Robots for Physical Skill Demonstration. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3491102.3501927

## 1 INTRODUCTION

Tutoring physical skills typically involves an expert demonstrating to a group of the audience how to interact with task-specific artifacts and tools in a sequential order [16]. Remote and synchronous physical-task tutoring enables people to interact with experts without having to be physically present [67] and leverage real-time communication for enhancing engagement [32]. These technologies have been increasingly popular for education and entertainment, encompassing activities such as electronics prototyping, woodworking, and scale-model making [7, 15, 41].

Physical skill demonstrations on a workbench often require activities at multiple locations and scales. For electronics tutorials, a demonstrator may need to produce individual electronic parts at a soldering station and assemble components in another area on the workbench. Meanwhile, it could be beneficial to show both the close-up details while soldering and wide-angle views during assembling. To accommodate such a variety of activity locations and scales, current practices for remote physical skill demonstration typically require setting up multiple fixed cameras pointing at various locations of interest. The demonstrator then manually selects one of the streams using a switcher or combines them into a composite stream [15, 20]. However, ensuring good camera placement often demands a significant amount of demonstrator effort [15]. Further, a fixed multi-camera setup does not always support the individual goals of the audience, who may be interested in seeing the demonstration from a different viewpoint. For instance, when the demonstrator uses a close-up view to show a process, an audience member already familiar with that process may prefer to explore other areas of the workspace. From the instructor's perspective, another challenge for remote demonstration is the lack of awareness of the audience. In the absence of audience's physical presence, reactions, and focuses of attention, instructors feel like they are 'talking to a void' [71]. Consequently, the audience experiences a weakened connection with the instructor. For physical skill tutorials specifically, demonstrators may miss opportunities to tailor tutorials according to the audience's needs.

Prior research in Human-Computer Interaction (HCI) has explored various approaches for enabling remote helpers to explore local physical workspaces, including guidance systems based on Augmented and Virtual Reality [12, 19, 67] and robotic camera systems [22, 58, 69]. Yet, with a few exceptions (e.g. [67]), these approaches typically focus on remote assistance scenarios where a remote expert navigates a local workspace to provide guidance, rather than supporting a group of audience exploring the remote expert's workspaces. Telepresence robots are physical surrogates that remote people can 'inhabit' to move in an environment and communicate with local people. In addition to offering an increased level of mobility to remote people [55], telepresence robots also provide them with tangible embodiments, which prior studies have found can increase their sense of being there [4, 38, 50]. With a few exceptions [2, 25], the dimensions of telepresence robots are typically in the scale of a human adult, making them less practical as embodiments for groups of remote viewers around a workbench. Swarms of tabletop mini-robots have shown substantial promise in a diverse range of use cases as dynamic tangible inputs and outputs [34, 37, 48, 65]. However, their potential in the telepresence context has yet been explored.

In this paper, we present Asteroids, a new approach for remote physical skill demonstrations using a swarm of small-scale, on-workbench telepresence robots to enhance the experience of both the demonstrator and the audience. Asteroids aim to support audience agency and their sense of presence and assist the demonstrator in controlling the demonstration experiences. For the remote audience, our approach materializes their presence directly in the physical workspace (Figure 1A), allowing them to inhabit a robot, find desirable viewpoints for watching the demonstration through automatic and manual navigation, and transfer to other robots to get a different perspective.

The additional mobility and physicality offered to the remote audience by Asteroids may influence the demonstrator's carefully planned tutorial flow and content [7, 15]. Demonstrators can use direct physical manipulations to catch hold of a robot for camera repositioning and pick up and place tangible tokens to control the flow of presentation (Figure 1B). We performed an exploratory evaluation to assess the utility of Asteroids through three architectural model-making tutorials with 12 remote participants. After each session, we asked participants to answer three task-related questions, draw a sketch of the model, and provide subjective feedback. Participants reported that Asteroids recreated an in-person workshop experience that contributed to their engagement, presence and connection to the instructor, and a new way to for remote demonstration observation.

Therefore, in this paper we contribute: 1) Asteroids, a novel approach for remote physical skills demonstrations using swarms of telepresence robots. 2) A set of scenarios exemplifying interactions and affordances of our approach. 3) A discussion of the results from a user study using an architecture model building scenario.

# 2 RELATED WORK

Our work builds on prior work on video conferencing and livestreaming, remote physical task assistance and tutoring, robotic telepresence, and robotic tangible interfaces. In this section, we discuss previous related research.

## 2.1 Video Conferencing and Livestreaming

Real-time video-mediated communication technologies, in particular video conferencing livestreaming, have seen significant progress in connecting people across distance [42, 56]. Video conferencing uses live video and audio links to support synchronous, engaging one-to-one or group communication for a variety of purposes, from maintaining intimacy between partners [49] to delivering online lectures [26]. Similarly, livestreaming uses live video and audio as primary communication channels, but focuses on a one-to-many paradigm where the livestreamers broadcast video and audio to an online audience [24, 42]. In recent years, livestreaming platforms have grown rapidly, attracting a large number of viewers interested in topics such as gaming [62], cultural heritage [41], and learning skills such as coding [9] and physical skills [15].

Despite video-mediated communication's success, the technology is not without limitations [51]. Previous research examined challenges related to visibility [30], awareness [23], presence [4], and social norms [49]. Our work focuses on designing novel video-mediated communication experiences for one particular scenario – physical skill demonstrations.

## 2.2 Remote Physical Task Assistance and Learning

Studies have shown that working remotely with physical objects poses challenges such as visibility, perspective, and attention [47]. Early work in remote physical task assistance has experimented with video links from multiple fixed cameras [20] and body-worn cameras [35] on the local worker, but found the two methods could not provide remote experts with sufficient awareness of the activities within physical workspaces [43]. Later research explored technologies that enable remote experts to freely explore the local workspace on their own. One method is to present a 3D-reconstructed workspace to the remote expert, who can navigate it using 2D graphical interfaces [19, 63] or head-mounted displays [67]. Mini-me [54] presents remote experts as a small-sized avatar on the local user's side to enhance gaze and gesture awareness. However, the local users cannot see the experts' physical space and their interactions with physical objects. Loki [67] supports live demonstration from an instructor rendered in point clouds. Such reconstructed workspaces can be highly immersive, but they lack the fine details that are necessary for certain physical tasks. The other method is to connect the remote expert to the local workspace through video feed from a teleoperated robotic camera. For example, TeleAdvisor [22] exploits a camera for visual information and a projector for annotations, both mounted on a robotic arm. Heimdall [31] rotates the workbench so that the camera can provide the remote expert with a full rotation view of the circuit board to be debugged. RobotAR [69] allows an instructor to pilot a mini-robot on remote workbenches to provide students with in-situ guidance on hardware prototyping exercises.

Most previous systems for remote physical task support focus on remote assistance i.e. a remote expert guiding the physical job performed by a local worker, but few (such as Loki [67]) have been designed for novices to watch and learn from live physical skill demonstrations. Our approach builds on previous methods that leverage robotic cameras for physical task assistance, but focuses on interaction designs for the less-explored space of remote physical skill demonstrations.

## 2.3 Telepresence Robots

A key challenge of video-mediated communication is that participants are less aware of other people's presence, since the physicality of people and the world around them has been flattened to two-dimensional video streams [23, 38, 51]. This problem becomes exacerbated in multi-party conferences, such as business meetings [66] and online classrooms [45, 71]. To restore the lost physicality, researchers explored using robotic embodiments to provide remote people with a sense of the local environment and to engage in conversations. [52] The use of telepresence robots has been found beneficial for a wide range of scenarios, including office work [38], healthcare [46], academic conferencing [50], and shopping [70]. Specifically, research has found that the physicality and mobility of telepresence robots can enhance casual communication [38], collaborative task performance [55], and local people's awareness and attention of the remote people [11]. Research has shown that university students feel more present and expressive when using telepresence robots to attend lectures compared to standard remote learning software [13].

Most telepresence robots are designed for interpersonal communication rather than to view the detailed actions of physical skill demonstrations on workbenches. Building on Adalgeirsson and Breazeal's [2] work on mini-telepresence robots for a single remote person, our research explores the potential of having a swarm of tabletop telepresence robots as the physical presence of a group of remote audience. Moreover, we explore the possibility of switching between multiple physical embodiments and viewpoints to enhance the viewing experience.

## 2.4 Robotic Tangible Interfaces

Tangible user interfaces (TUI) materialize the inputs and outputs with digital information systems [28]. Actuated shape-changing tangible interfaces, such as TRANSFORM [27], DynaBlock [64], and Elevate [29], further increase the flexibility of TUI. Recently, self-propelled [48] and self-reconfigurable [37, 65] tangible interfaces based on swarms of tabletop mini-robots have expanded the vocabulary of tangible interfaces, and allow for more versatile and expressive physical interactions. Zooids [37] are a collection of mini-robots that can reorganize themselves as dynamic shape displays and haptic rendering devices [34]. Particularly relevant to our goal, Siu et al. [61] explored Zooids as shape displays and physical actuators controllable by both remote and local collaborators. ShapeBot [65] expands the expressiveness of swarm interfaces by adding the shape-changing capability to individual swarm members. HERMITS [48] proposes tangible robots that repurpose themselves by changing shells. These tangible robotic interfaces have

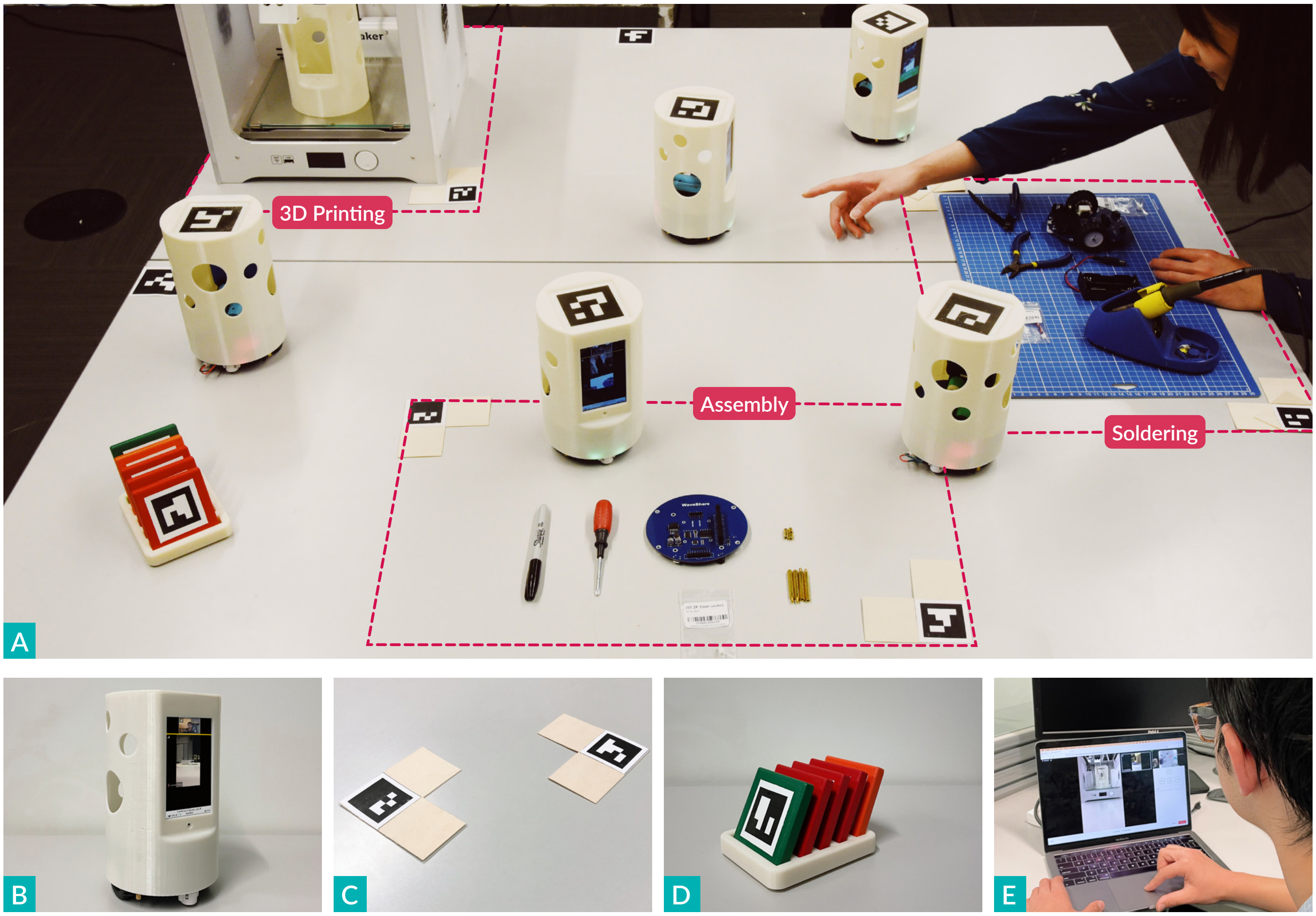


**Figure 2: Elements of a physical skill demonstration that uses the Asteroids approach. A) a workspace with multiple (e.g. soldering, assembly, and 3D Printing) zones managed by the demonstrator, B) Asteroids robots, C) Zone tokens, D) Command tokens, and E) a member of the remote audience.**

shown a range of promising applications, such as data physicalization [37, 65], mobile tactile feedback [34], and performance [48]. Through an elicitation study, Kim et al. [33] revealed that users prefer gestures, touch, verbal commands, and a combination of gestures and verbal utterances to control swarm interfaces.

Our approach expands existing robotic tangible interfaces – the mini-robots are not representations of abstract digital information but embodiments of remote humans and their viewpoints. We explore interactions with these tangible robotic representations imbued with human agency, specifically in remote physical skills demonstrations.

## 3 ASTEROIDS

Our design goals are to support the audience's agency, audience's presence, and demonstrators' control in physical skill demonstrations. Asteroids use a swarm of miniature telepresence robots for demonstrating and broadcasting physical skill performances on workbench-scale workspaces (Figure 2). We design a web-based video conferencing experience allowing participants to navigate in a remote workspace using Asteroids robots and switch between multiple viewpoints to focus on different workspace regions. We arrived at the interaction and hardware design choices for Asteroids after several rounds of iterative prototyping and testing. We ran a number of physical task tutorials through Asteroids prototypes within our research group, starting with low-fidelity foamcore robot shells and moving on to 3D-printed shells. Synthesizing feedback from these sessions, we refined the software interface, the robot form factor, and its interaction affordances.

In this section, we introduce the design of Asteroids robots, audience interfaces for moving and switching between robots, and a set of tangible interactions to facilitate demonstrators' active control over camera viewpoints and presentation flow. Finally, we explore robot autonomous behaviors that can enhance the experience of both demonstrators and remote audience.

## 3.1 Physical Workspaces with Asteroids Robots

We focus on a single workbench scenario to explore movable embodied presence for physical task demonstration. As shown in Figure 2A, physical task workflows can include subtasks spread across multiple separate zones and require different tools and equipment. For example, in robotics prototyping, demonstrators may build individual components with a soldering station in one zone, while assembling and testing them with an oscilloscope in another. The importance of such work zones in organizing workflows prompted us to develop tangible zone tokens (Figure 2C) that can be used to define and dynamically resize these zones. To promote audience agency while minimizing robot interference with ongoing demonstrations, zone tokens set constraints such that robots can move freely outside these areas to observe activities and artifacts of interest, but cannot cross the zone boundaries.

## 3.2 Asteroids Robots

Asteroids robots are self-propelled, graspable physical embodiments of the audience. Demonstrators can grab the robots, reposition them, and communicate with the remote people inhabiting the robots (Figure 2B). The robots have a cylindrical shape with a footprint similar to a coffee mug (115 mm by 115 mm by 195 mm),. We choose this size as it is large enough to accommodate a sizable screen and the necessary actuation mechanisms, while small enough for single-hand manipulation. Because of their compact form factor, demonstrators can easily integrate these robots into physical workspaces without adding significant overhead. The robots can navigate the workbench surface with their wheels and capture a live video stream of the workspace using a front-facing camera (Figure 2B). The robot display allows demonstrators to view live video streams of the remote participants inhabiting the robots as well as a local camera view for their own awareness of what the robot sees. The multiple different-sized holes on the robot's shell make the smooth surface easily graspable by the demonstrator's fingers when picking up the robots to adjust their positions. The top surface of the shell features a square dent containing a fiducial marker for positional tracking. The dent's shape serves as a holder for the demonstrator to place a physical token. Following the token+constraint paradigm for tangible interaction [68], we designed the dent as the constraint for command tokens that associate various statuses with a robot, such as making it the primary viewing camera and restricting its mobility. The tangible command tokens are color-coded based on the status they set for robots, as shown in Figure 2D. Thus, they combine status update and display into one physical object that is easily manipulated and noticed.

## 3.3 Audience Interactions

In this subsection, we present the web interface for audience members to remotely control Asteroids robots (Figure 3).

*3.3.1 Audience interface.* The audience can connect to an Asteroids robot swarm through a web interface (Figure 2E), which shows each participant the live camera stream from the robot they currently inhabit, a top-down-angle livestream of the workspace as a map, and buttons for direction control (Figure 3). The interface also provides a live audio connection with the demonstrator.

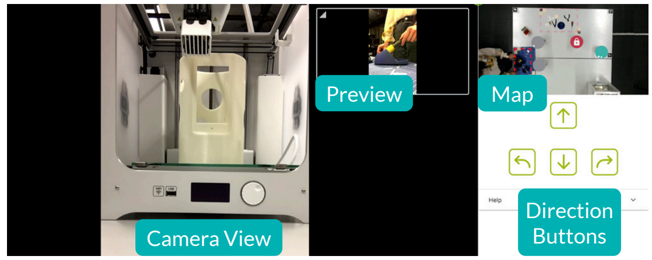


**Figure 3: The audience web interface displays the camera view of the robot that the remote audience member currently inhabits, a preview of another robot's camera view, the map, and direction buttons.**

In line with prior robot teleoperation interface designs [39, 53] for supporting operators' spatial awareness, we incorporate a map that provides an overview of the entire workspace. To facilitate viewpoint selection and control, the map is augmented with robot status visualizations. (Figure 4). A circle indicates the location of an Asteroids robot on the map. Each white dot represents one audience member inhabiting the robot, and the dots are arranged in an arc near the circle's edge. The arc's midpoint points to the forward direction of the robot.

The remote user gains control of a robot when selecting an unoccupied one, as depicted in Figure 5A. Before making a switch, the user can hover the cursor above any robot to see a picture-in-picture preview of the camera stream from that robot (Figure 3). Other audience members who later select this robot can still watch its camera livestream but cannot control the robot. Instead, they are placed in a queue for control access, and are granted control once all the previous users have left.

The robot's location indicators are color-coded according to the user's status (Figure 4). When the user inhabits a robot and has control over it, the indicator is green. When the user occupies that robot but does not have control, the indicator is teal. An orange indicator shows that the robot is the main camera, and red depicts a robot pinned in place by the demonstrator. Gray indicators show uninhabited robots. Lastly, red dashed rectangles on the map indicate work zones defined by the demonstrators using zone tokens. The demonstrator interactions will be introduced in Section 3.4.

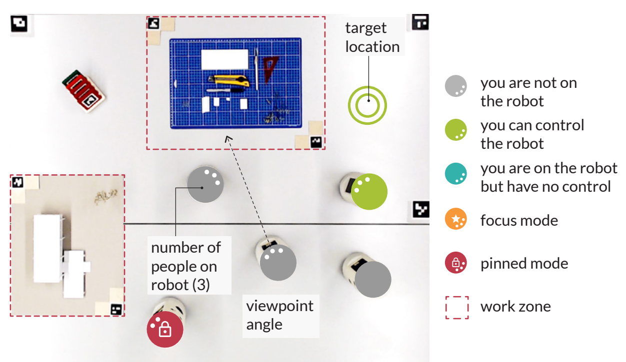


**Figure 4: Audience interface map and visualization legends**

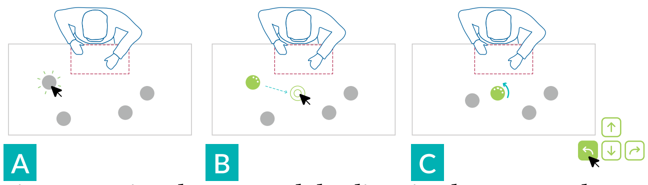


**Figure 5: Using the map and the direction buttons on the audience interface, a remote user can A) select a robot to watch its camera stream, B) select a target location to move the robot, and C) use direction buttons to move the robot.**

*3.3.2 Robot control.* Teleoperating a robot can be a demanding task [53]. Following prior robot teleoperation research [8, 14, 39], we provide both automatic and manual navigation methods for a remote user to move a robot in the workspace.

Using the automatic method, when the audience member clicks on a location on the map (Figure 3), the robot follows an obstacle-free path to reach the targeted location (Figure 5B). Upon reaching the goal, the robot orients itself towards the closest work zone in the workspace. If an unoccupied robot is closer to the target location than the robot currently controlled by the user, we perform 'robot hopping' to switch the user to the closer robot. This behavior reduces wait time for viewpoint control.

With the manual method, a remote user can use the direction buttons in the right sidebar (Figure 3) to move the robot forward or backward, or rotate it clockwise or counterclockwise (Figure 5C). We expect users to benefit from these two complementary mechanisms: automatic navigation to reach relatively distant goals in the workspace without having to explicitly plan a path; and manual navigation for fine-tuning the viewing angle.

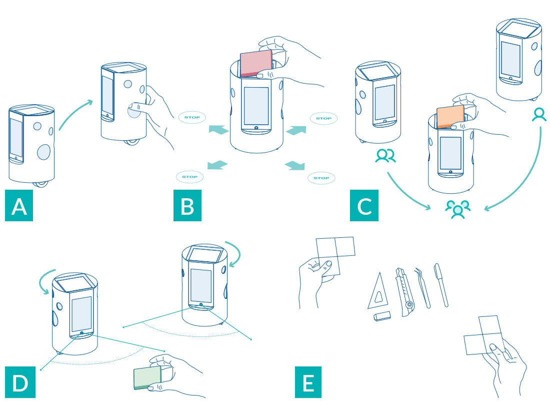


**Figure 6: Tangible interactions for demonstrators: A) Picking up and relocating Asteroids robots. B) Making a robot stationary using a red token. C) Moving every audience member to a robot using an orange token. D) Specifying a region of interest and making all robots face that spot with a green token. E) Rearranging the work zones using a pair of zone tokens.**

## 3.4 Demonstrator Interactions

The demonstrator's control over live video streams can be crucial for leveraging their domain expertise and accomplishing instructional plans [7]. However, the added agency of audience moving and switching cameras may add uncertainty to the demonstrators' everyday practices. Therefore, Asteroids incorporates interactions for actively controlling camera viewpoints and presentation flows. Our designed interactions exploit tangible input carried out via direct physical manipulation of the robots and a set of tangible tokens. We choose the tangible modality to control the robots for its naturalness and expressiveness [1, 21, 60, 65], in addition to reducing instructor-audience communication interference that could result from gestures [33] and speech [5]. This modality also helps provide a consistent mental model for both physical tasks and viewpoint control.

Our interaction design focuses on four basic viewpoint control operations, and a method to set work zones. Following the token+constraints framework [68] for tangible interaction, we devise block-shaped *command tokens* that fit the physical constraints on the robot shells to give them instructions. Our L-shaped *zone tokens* for defining work areas in the workspace are inspired by prior tangible token designs for area selection [60].

*3.4.1 Moving the robots.* During demonstrations, it can be desirable to reposition a robot so that a demonstrator can show a particular action or object from a specific perspective, such as highlighting a particular detail or showing an overall view. To do so, the demonstrator can grab the robot's shell and relocate it to the desired location (Figure 6A). When repositioning a robot, the demonstrator can refer to the local camera view on the robot display to find a desired framing of the scene.

*3.4.2 Pinning a robot in place.* When a robot is pinned in place, its locomotion ability is disabled. This feature is useful when demonstrators select particular views that they think are important. In such scenarios, audience members can either choose a view set up by the demonstrator or pick a different robot. For example, the demonstrator may choose to pin a robot at an angle that covers an entire work zone. They can do so by placing a red command token on the top of the robot (Figure 6B). Demonstrators can revoke this pinned status at any time by removing the token.

*3.4.3 Making a robot the focus.* In current physical skill demonstration practices, instructors often choose a single camera angle to present their demonstration around it. Despite lacking flexibility, this method can be very effective for the instructor to enforce a particular presentation flow. To direct audience's attention using Asteroids, the demonstrator can place an orange command token on a robot to make it the focus, transferring all audience members to that robot (Figure 6C). Additionally,the demonstrator can arrange a sequence of robots pointing at objects of interest, then transfer the focus token from one robot to another, creating a more structured instructional narrative. Audience members may choose to leave the focus robot to inhabit other robots. The demonstrator can revoke the focus status by removing the token.

*3.4.4 Spotlighting a location in the workspace.* In addition to transferring all audience members to the focus robot, the demonstrator can use a green command token to direct all robots to look at the location marked by the token (Figure 6D). Using the spotlight token is a subtler way to draw attention than the focus token, as the audience would not lose control of their current robot. The demonstrator can also move the spotlight token along a path on the workbench to simulate a camera-pan effect, guiding the audience in observing a scene wider than the robot's field of view.

*3.4.5 Setting work zones.* Setting work zones allows demonstrators to delineate the boundaries of the areas reserved for demonstration activities. Our program prevents the robots from entering these areas to avoid interference with the ongoing physical task. The demonstrator can use a pair of L-shaped zone tokens to mark the two diagonally opposite corners of a rectangular work zone (Figure 6E).

### 3.5 Autonomous Behaviors

In addition to following local and remote commands, Asteroids robots can act autonomously to enhance the experience of both demonstrators and remote audience. In this section, we describe the autonomous robot behaviors for adapting to the demonstrator's location, adapting to work zone changes, and using robots as automatic cameras that track points of interest. Future work could extend this repository of behaviors by adapting the behavior designs from autonomous robot [59] and camera [10] systems.

*3.5.1 Robots adapting to the demonstrator's location.* To reduce the physical effort required for demonstrators to reach an uninhabited robot each time a new viewpoint needs to be added, we automatically move an uninhabited robot into the demonstrator's reach if there are currently no other available robots within the range.

*3.5.2 Robots adapting to work zone resizing.* When the demonstrator expands a work zone, some robots may end up in the resized zone. These robots are then automatically moved to the edge of the zone to minimize interference with the forthcoming demonstration activities. Similarly, when a work zone is reduced in size, the robots facing it move closer to the edge of the zone to ensure a consistent framing of the scene.

*3.5.3 Robots as automatic cameras.* Depending on the demonstrator's configuration, some Asteroids robots are reserved by the system to serve as automatic camera persons that track the demonstrator or other points of interest. For example, the camera of a robot can always be pointed at a demonstrator from opposite ends of the workbench, capturing an overview of the demonstration.

## 4 IMPLEMENTATION

We developed a prototype to study our vision of mini-telepresence robots for remote physical skill demonstrations. The prototype consists of four components: the Asteroids Server that controls the robots and handles user input, the audience web interface for remote operation of the robots and real-time video and audio communication, the Asteroids robots, and the tangible tokens.

### 4.1 Prototype Overview

**The Asteroids robots and tangible tokens** operate on a workbench, which is monitored by two downwards facing cameras, one

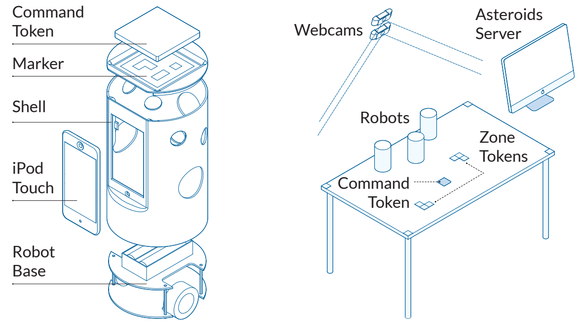


**Figure 7: Telepresence robot explosion diagram (left), and Asteroids prototype overview (right).**

for position tracking and the other for the top-down map view in the remote audience interface (Figure 7 right). The Asteroids robot carries an iPod device that streams live video of the workbench from its front-facing camera to **the audience interface** web application running on remote audience's computers. The iPod also receives the live webcam feeds from the audience and shows them on its display. **The Asteroids Server** performs visual tracking of robot and token positions, computes robot trajectories, and sends motion control signals to the robots. In addition, it communicates with each audience interface application to receive audience input and send robot status updates. The Asteroids software code and hardware design are available online[1].

### 4.2 Asteroids Server

The Asteroids Server is a Python program that tracks robots and tokens, manages audience interfaces, and controls robot motion. It analyzes the top-down video stream of the workspace to calculate the positions and orientations of the ArUco markers [18] on the robots and the tokens using OpenCV at a rate of 30 frames per second. The server maintains two-way WebSocket connections over the Internet with every audience interface application for receiving user commands such as moving a robot and for sending robot status updates such as their positions. Based on audience input and the positions of robots and tokens, the server computes robot motion control signals and sends them to the on-board computer of the robots via a local WiFi network.

For robot trajectory planning and motion control, the server program first locates the fiducial markers at the three corners of the workbench to define the workspace coordinate system, which the robot and token positions are transformed into for all the following computation. The server uses an A* global planner for autonomous robot navigation to compute an obstacle-free path towards the user-specified goal. A local planner based on Artificial Potential Field then computes the target linear velocity and orientation for the robot to travel to the waypoints on the global path. The robot's on-board control program tracks target velocity and orientation using a proportional–integral–derivative (PID) controller. We decided to let the robots track target orientation values instead of angular velocity, as angular velocity measurements through optical tracking

[1] Asteroids Github Repository: https://github.com/jchrisli/asteroids-chi22

alone can be extremely noisy. The global planner, local planner, and the PID tracker on the robot run at a frequency of 0.2Hz, 5Hz, and 30Hz, respectively.

## 4.3 Asteroids Robots and Tangible Tokens

The Asteroids robots consist of a robot base, a shell, and an iPod Touch (Figure 7, left). The robot is cylindrical, with a radius of 115 mm and a height of 195 mm. We use an AlphaBot 2 differential-drive robot as the robot base. A Raspberry Pi 3B on-board computer runs a Python program to receive motion control commands through the local WiFi network and set speed values for the two motors. The robot shell is 3D-printed and made of ABS plastic. It has ten holes (three large ones and seven smaller ones) distributed on the surface to facilitate grasping. The top of the shell has a square-shaped dent (70mm by 70mm by 5mm), where a tangible command token can be installed. The command tokens are 3D-printed with ABS plastic and later painted. They have a dimension of 70mm by 70mm by 10mm. The AruCo markers on the robots and the tangible tokens are 60mm by 60mm and 40mm by 40mm, respectively. To set robot status with the tangible tokens, we detect which robot has been covered by a token and alter its status based on the token type. The zone tokens are pairs of L-shaped cutouts from veneer sheets, each with a 30mm-by-30mm AruCo marker at the corner.

There is a slot in the robot shell for sliding in an iPod Touch (7th Generation). The iPod is positioned upside-down, so the camera is close to the workbench surface for improved workspace visibility. It runs a custom web application that sends its front-facing camera streams to the audience interface over a WiFi Internet connection. The application also receives the webcam video streams from the audience members and displays them on the iPod screen.

## 4.4 Audience Interface and Video/Audio Streaming

The audience interface is a web application that can run in web browsers with WebRTC and WebSockets support. We run a special instance of the audience interface on the Asteroids Server to grab the top-down video feed and send it to other audience's browsers as the map view. This instance also keeps two-way audio channels with other instances to support live conversation between the demonstrator and the audience. We built the application using the React Framework[2]. All live video and audio streaming is performed through WebRTC. We use the Twilio Video API[3] to create a group conferencing room that includes all the iPods, the Asteroids Server, and the audience members. Video and audio streams are then selectively turned on for individual participants based on the participant's type and status.

# 5 SCENARIOS

In this section, we use four example scenarios to illustrate how audience and demonstrators can utilize Asteroids' unique features. The scenarios revolve around the experience of two remote physical skill tutorial attendees and their instructors. The two attendees are interested in learning hands-on skills for robot making and architectural modelling, respectively, for academic or leisure purposes.

[2] React Framework: https://reactjs.org
[3] Twilio Video API: https://www.twilio.com/video

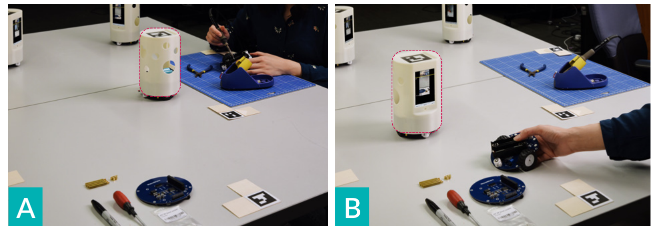


**Figure 8: An audience member follow the instructor's work in multiple areas. A) Observing soldering in the zone for making individual parts. B) Observing robot assembly in the assembling zone.**

While they typically practice these skills at their home workshops, they seek to further their techniques by watching demonstrations from remote experts, who run such tutorials on Asteroids-equipped workbenches.

## 5.1 Audience Following Work in Multiple Areas

The following scenario illustrates how a remote audience member, Ada, can follow physical tasks distributed across multiple work zones by leveraging the mobility of Asteroids robots. Ada is attending a tutorial on robot assembly with Asteroids robots. The tutorial workspace has two work zones, one for making individual parts and another for assembling. The first step of the tutorial is to mount an external battery holder to the robot base. The instructor starts the demonstration by soldering a connector to the robot base at the zone for making individual parts, where there is a soldering station. Ada moves the robot closer to the soldering iron to examine the soldering technique (Figure 8A). Next, the instructor takes the robot base to the assembly zone to start assembling. Ada follows the instructor, moving to the assembly zone to watch how other parts are added to the robot base (Figure 8B).

## 5.2 Demonstrator Adapting to Audience Interest

This scenario depicts how a demonstrator can tell when audience's interests change through observing the movements of robots, and

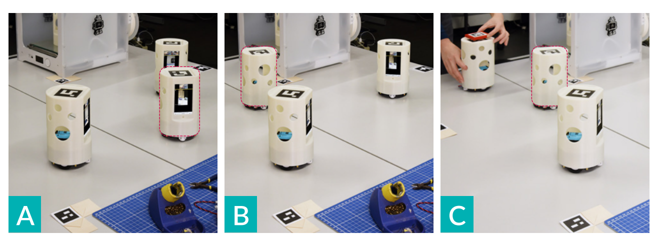


**Figure 9: A demonstrator notices a change in audience interest and adapts the demonstration flow. A) and B) An Asteroids robot moves away from the soldering iron and moves towards the 3D printer. C) Noticing the movement of this robot, the instructor adds a robot and uses a red command token to pin this viewpoint in place.**

adjust the demonstration flow if needed. Building up on the previous robot assembly scenario, the instructor Alice introduces a third work zone for printing the shell of the robot with a 3D printer. She briefly explains and starts the 3D printer before moving to assemble the robot. While working in the assembly zone, Alice notices an audience member moving a robot towards the 3D printer to study the printing process (Figure 9A and B). She realizes that the audience is interested in learning more about how the 3D printer works. So she moves back to the printer zone, does more explanation, and uses the red token to leave a pinned robot there so that the audience can always go back to check the printing in progress (Figure 9C).

## 5.3 Demonstrator Composing a Narrative Using Tangible Interactions

This scenario shows that the demonstrator can use tangible interactions with Asteroids robots to direct the audience's attention and compose a narrative to support the instructional goal. Alice is running another robot assembly tutorial. This time she plans to start with a tour of the three work zones so that the audience will be familiar with all the necessary tools and equipment. She also wants to have the audience's full attention during the tour, as it will prepare them for the upcoming tutorial. To do so, she first places three unoccupied robots pointing at each of the three zones. She then places an orange token on the robot capturing a close-up view of the soldering zone (Figure 10 A1 and A2) and transfers all the audience to that robot. With all the audience's attention on the soldering zone, she introduces the tools and equipment used there. She repeats a similar process at the assembly zone (Figure 10 B1 and B2) and the 3D printer zone (Figure 10 C1 and C2) to guide all the audience through the workspace.

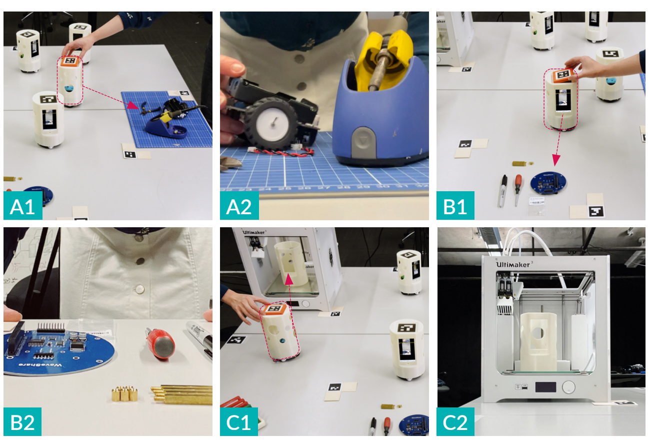


**Figure 10: The demonstrator constructs a workspace tour using the orange command token. A1) The demonstrator places the orange token on a robot pointing at the soldering zone of the workspace. A2) As a result, all audience members transfer to this robot and focus on the soldering zone. The demonstrator later directs the audience's attention to the assembly zone (B1 and B2) and then the 3D printer (C1 and C2) with the orange token.**

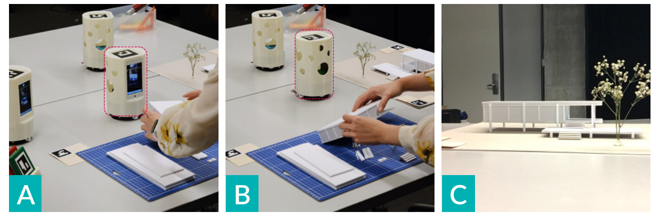


**Figure 11: An audience member freely explores the workspace. A) The audience member realizes that she already understand the current step as it has been repeated several times, B) so she drives the robot to explore other areas of the workspace and C) studies the finished model in another zone.**

## 5.4 Audience Freely Exploring Points of Interest

This scenario shows that an audience member can exercise their agency through Asteroids robots and freely explore the workspace to learn more aspects about the demonstrated activity. Anna is attending a remote tutorial on building architectural models through Asteroids robots. There is a repetitive step during the model building process where the instructor needs to install eight columns following the same procedure (Figure 11A). After watching the installation process for two of them, Anna decides to explore the workbench (Figure 11B) and study the finished model displayed in another zone. (Figure 11C).

# 6 EXPLORATORY EVALUATION

We conducted an initial observation user study in three instructional sessions with groups of four participants to assess the utility of our approach and usability of the designed features. Furthermore, we looked for the impact of Asteroids on user engagement and sense of presence. The main goal was to gain an initial understanding of the audience's experience when they use Asteroids to control camera viewpoints, communicate with the instructor, and observe the demonstrated activities. We choose to focus on the audience's experience in this study, as they are the end consumer of remote physical skill tutorials.

## 6.1 Design

Our study design aimed at exploring the impact of the proposed approach on the audience with a demonstration process representative of some common physical skill tutorials. To this end, we chose architectural model building as the physical task. It is a skill commonly acquired through observing others' practices and involves the appropriate level of complexity regarding tools, procedures, and the use of workspaces.

An architect with instructional experience developed an online tutorial to demonstrate how to build a scale model of the Farnsworth House, an icon of Modernist architecture. Three groups of four remote participants attended the tutorials using Asteroids robots. Figure 12 shows the demonstrator instructing the participants on how to install the roof piece of the model. Following each tutorial session, participants answered quiz questions regarding techniques,

**Figure 12: Participants focusing on the demonstrator during a study session.**

tools, and artifacts discussed during the tutorial. The quiz questions were not disclosed to the demonstrator. The participants also submitted a sketch based on their recollection of the model.

## 6.2 Task and Procedure

The study began with a training session for the participants on the Asteroids audience interface. We reminded participants to pay attention to the demonstration, as they would be asked to answer questions about the task and to provide a sketch of the model after the session. The demonstrator performed the demonstrations across two work zones on the workbench, one for preparing the materials (A) and another for assembling the model (B). The demonstrator showcased how to prepare and assemble individual parts, the tools used in the process, and the final product of the architecture model. Specifically, the demonstrator started the session by explaining how to prepare for materials at Zone A, moved to Zone B for assembly of the model's main body, returned to Zone A to demonstrate two specific techniques, and concluded the tutorial at Zone B with a finished model. In each session, the demonstrator used the focus token three times, the spotlight token once, and the pin token once. The demonstrators generally followed the plan for every session with some improvisation to meet audience needs. After the tutorial, the participants completed a post-study questionnaire, submitted a sketch of the model, and provided additional feedback through semi-structured interviews. The tutorial lasted 25 to 31 minutes, depending on how much communication occurred between the participants and instructor. Including participant training and debriefing, each full session lasted 60 minutes. A 15-minute interview was conducted with each participant after the session.

## 6.3 Setup and Apparatus

We conducted the study on a 1.52m by 1.00m workbench in a research lab. The participants and the demonstrator interacted with an Asteroids prototype that included five Asteroids robots. We did not enable autonomous robot behaviors introduced in Section 3.5 during the study. The participants joined the study through the web application described in Section 3.3.1 using their personal computers.

## 6.4 Participants

We recruited 12 people (seven females, four males, and one participant who chose not to disclose) interested in learning architectural model building from the local community. None of the participants had architectural design training. Their ages ranged from 24 to 36 years ($M = 28.7$, $SD = 3.8$). All reported having previous experience in video-conferencing software. Six participants reported using video conferencing software to watch physical skill demonstrations, and four stated that having watched physical skill livestreams.

## 6.5 Results

For each evaluation session, we employed questionnaires to assess participants' perceptions of engagement, sense of presence, understanding, and the utility and usability of the designed features. As an initial assessment of the remote audience's understanding of the demonstrated task, we asked them three questions regarding (1) how to install a slice of tracing paper on the model base (a *technique*); (2) the number of stairs in the model (an *artifact*); and (3) the tools used to put up a tree in front of the model (a *tool*). In addition, participants were asked to sketch a model from memory after each session.

Furthermore, we conducted a post-study interview with each participant to gather feedback and understand usage patterns. Specifically, we asked them to compare the experience of using Asteroids with conventional video conferencing software or livestreaming platforms for watching physical skill demonstrations. We present the results from the questionnaire (Table 1) and interviews in detail below.

*6.5.1 Audience Engagement.* As the questionnaire results indicate, participants enjoyed the experience (Q1.1), and most of them were able to follow along during the demonstration(Q1.2). P5 described

| Statements | Median (IQR) |
|---|---|
| **1.Engagement** | |
| 1.1 I enjoyed the experience. | 7(1) |
| 1.2 I was able to focus on the demonstration. | 6(1.25) |
| **2. Utility and Usability** | |
| 2.1 It was helpful to change robots using the map. | 7(1) |
| 2.2 It was helpful to move a robot using the map. | 7(3) |
| 2.3 It was helpful to use direction buttons to control the robot. | 7(1) |
| 2.4 It was helpful that the demonstrator could move me to a robot. | 7(0.25) |
| 2.5 It was helpful that the demonstrator could spotlight a location. | 7(0.5) |
| 2.6 It was helpful that the demonstrator could pin a robot. | 7(1) |
| **3. Presence** | |
| 3.1 The demonstrator noticed me. | 7(2) |
| 3.2 The demonstrator paid attention to me. | 6.5(2) |
| 3.3 The demonstrator reacted to me. | 6.5(2) |
| 3.4 The demonstrator adjusted for my position or orientation. | 6.5(2) |
| **4. Understanding** | |
| 4.1 Overall, the task was easy. | 6(1.25) |
| 4.2 The mental effort required to watch the demonstration was low. | 5(2) |
| 4.3 I was able to see the demonstrator's actions I was interested in. | 7(0.25) |
| 4.4 The demonstration was easy to understand. | 7(1) |
| 4.5 I could see the details that I was interested in. | 6(1.25) |
| 4.6 I have a good understanding of the overall layout of the model. | 7(1) |
| 4.7 Other robots did not block what I wanted to see. | 5(2) |

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

**Table 1: Results for the user preference questionnaires (Median, Inter-quartile Range).**

the experience as *'This is a unique experience...It's like we are injecting our souls into these things (robots) and making them alive and be our eyes.'* The additional mobility and user agency afforded by Asteroids contributed to the participants' engagement. P6 commented *'Sometimes when watching a Twitch livestream, I just leave it there and look at it every few minutes because you cannot do much if the current (content) is not that interesting. This (Asteroids) is different. There is much more to explore here.'*

*6.5.2 Utility and Usability.* Asteroids features contributed to the audience experience in that participants enjoyed controlling camera positions through direction buttons (Q2.3) and switching between viewpoints by inhabiting different robots (Q2.1). Most participants found it helpful that they could move the robot to a specific location by clicking on the map (Q2.2).

Overall, participants found that the demonstrator's viewpoint control actions – making a robot the focus, spotlighting a location, and pinning a robot – contributed to their positive experience (Q2.4-Q2.6). Participants especially appreciated that the demonstrator could set a robot as the focus and transfer their viewpoints to that robot.

While Asteroids' features were generally found to be valuable and practical, the study has revealed several usability and functionality issues. P1 and P11 would like more flexibility in controlling the robot's final orientation after reaching its goal through autonomous navigation. Future work could explore using recent computer-vision algorithms to detect task-related objects, such as the demonstrator's hand [58], essential tools, or the artifact being created, as focuses of the robot. P12 expressed difficulties with the 'robot hopping' behavior and found the change of viewpoints abrupt. We believe transition effects would smooth viewpoint changes.

*6.5.3 Sense of Presence.* Most participants reported that the demonstrator had noticed and paid attention to their presence (Q3.1 and Q3.2). In particular, 83% of the participants felt that the demonstrator had reacted to them (Q3.3) and actively adjusted the demonstration for their viewpoint(Q3.4), suggesting a sense of connection with the demonstrator.

Participants' feedback suggests a clear sense of being in the workspace and being together with the demonstrator. Echoing the findings of earlier telepresence research using physical embodiment [4, 50], Asteroids robots allowed the demonstrator to distinguish individual participants from the amorphous and undifferentiated crowd. Moreover, demonstrators were able to give participants more personalized consideration, which previous research suggests has a positive impact on learning [17]. P3 called it *'a superb experience'* when the demonstrator addressed him by name.

When comparing their Asteroids experiences with using conventional video conferencing software and livestreaming interfaces, participants highlighted that Asteroids recreated an in-person workshop or lab experiences together with instructors and peers. P5 described it as *'I like this experience of being there, being right in front of the instructors, being with other students. You cannot do this with a single camera.'*. P9 mentioned that *'Checking others' views is just like walking around the lab table...'*. P3 further commented that *'I feel I was there as those tools were so close to me. The next thing I hope is that I can move those tools and push them to the mentor.'*

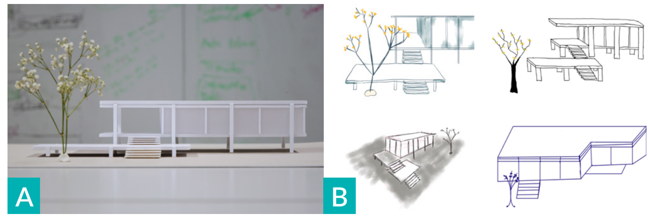


**Figure 13: A) For each evaluation session, the demonstrator built an architectural model , and B) we requested the participants to submit later sketches of what they perceived.**

This experience of 'being in the workshop' could further encourage participation, contributing to the sense of presence and connection to the instructor.

*6.5.4 Perception and Understanding.* Overall, participants reported low task difficulty (Q4.1) and mental effort (Q4.2). They found the demonstration easy to understand (Q4.4) and they could see detailed actions and artifacts of their interest (Q4.3 and 4.5), although some experienced other robots blocking their views (Q4.7). Most participants felt that they had a good understanding of the model layout (Q4.6).

We also examined participants' responses to the quiz questions and compared their sketches with the finished model to get an initial understanding of their observation during the demonstration. Overall, participants correctly answered 92% of the questions. For the question on *technique*, 75% of the participants responded correctly. For questions on *artifact* and *tool*, all participants responded correctly. 75% of the sketches were free of layout errors, suggesting that most participants had an accurate understanding of the overall layout. Figure 13 shows the finished model and examples of sketches. Out of the four example sketches, the bottom right one has a layout error.

In post-study interviews, participants described leveraging robots' mobility to find relevant details. The instructor's active control of the viewpoints also helped them follow along. Additionally, participants made use of the viewpoints of robots that they did not control. P2 mentioned that *'...later on (in the session), I liked to explore the views from other robots as they are put there by other people and might already be a good angle. It's like resource sharing.'* Participants also appreciated the freedom to wander away from the main demonstration activity and pursue their own learning goals. P7 described that *'I already know how to install the wall, so when the demonstrator was doing that, I moved to a robot close to the finished model so that I can remember its look.'*

*6.5.5 Audience Behaviors.* We analyzed the video recordings and the event logs of the study to further uncover participants' usage patterns of the robots. We paid special attention to the two scenarios described in Section 5, audience following demonstration in multiple areas and freely exploring points-of-interest. In both scenarios, audience could leverage viewpoint mobility through Asteroids robots, albeit for somewhat contradicting purposes. We found that most participants followed the instructor to the current zone, with two exceptions where one participant chose to remain in the previous zone, possibly for further examination of the previous step. Indeed, we observed in total 14 instances where a participant deviated from the current demonstrated activities and explored other areas of the workbench. These observations suggest that Asteroids could support rich audience behaviors in which they could exercise their agency.

Our initial exploration study suggests that the features of Asteroids could support audience engagement, their sense of presence and connection with the demonstrator, and their efforts to follow along the demonstration, by recreating an experience similar to an in-person workshop.

# 7 DISCUSSION

In this section, we reflect on some high-level observations from our initial evaluation, discuss limitations of our approach, and present possible directions for future research.

## 7.1 Recreating Workshop Experiences: Agency, Ownership, and Presence

Echoing previous research on telepresence, both the agency [55] and physicality [4, 38] of Asteroids robots contributed to the reported experiences of 'being in a workshop'. Participants appreciated that the instructor noticed their physical presence and that they can actively explore the workspace as opposed to being passive observers. In addition, our observations and interviews revealed interesting differences between using regular telepresence robots (e.g. [6, 13]) and using a swarm of mini-robots for recreating classroom experiences. Participants recognized the value of switching between robots for instant travel within the workspace, yet some wished to have exclusive control over the robot they were on. It is possible to coordinate shared ownership over robotic embodiments by leveraging methods from group robot control [36, 40] and preference aggregation mechanisms [3].

The fact that a remote person can be on any robot at a time hints at an exciting perspective for reconceptualizing presence. As P7 commented *'...initially I thought this robot is me and (to move around) I just drive this robot. But then I realized it was more efficient to switch between robots so I would adjust the positions of multiple robots from time to time.'* As with previous research on AI agents that can transfer between multiple physical embodiments [44], future telepresence research could investigate the implication of a distributed physical embodiment as a 'swarm presence' for remote experiences.

## 7.2 Limitations

Our current exploration and evaluation of Asteroids are subject to limitations. While Asteroids includes interactions for both the audience and the demonstrator, our current evaluation study focused solely on the audience experience. Despite the audience's high appreciation of the demonstrator's control of viewpoints, it is uncertain how well these new interactions could be incorporated into the demonstrator's current practices. Since our current study is limited to only one expert demonstrator from a single domain, future studies would aim to involve demonstrators from various domains, such as electronics and handcrafting.

Our evaluation focused on studying participants' experience with Asteroids as a full suite. Previous research has proposed telepresence approaches that employed multiple stationary cameras [4] or a single mobile camera [2] for remote communication at a workbench scale. While our qualitative results suggest that the mobility of Asteroids robots contributed to engagement and sense of presence and that our multi-robot approach benefitted understanding, we did not specifically investigate the effects of viewpoint mobility or number of cameras. By comparing tutorial experiences with Asteroids under different feature configurations, such as one robot versus multiple robots, or stationary robots versus mobile robots, our future work could study how particular design factors in telepresence systems impact the delivery and understanding of remote physical skill demonstrations.

While the four scenarios in Section 5 illustrate some potential interactions enabled by Asteroids, it is to be determined whether they are the most representative of real-world usage with Asteroids in remote demonstrations. Long-term deployment studies are needed to uncover the actual audience and demonstrator usage patterns.

## 7.3 Future Work

We plan to install pan-tilt-zoom cameras on the robots to expand their view scope. Inspired by participants' desire to move objects remotely, we plan to explore adding a mini-robotic arm to the robot for audience-demonstrator co-creation, as well as richer communication using hand-raising and other body language signals.

Additionally, we would like to examine people's social behaviors when using Astroid robots. Future studies could investigate the social dynamics among audience members, such as their sense of co-presence, and ownership of the space and individual robots.

Asteroids currently supports a small number of remote audience to control an approximately equivalent number of robots on a single workbench. Scaling up this setup in any dimension opens up exciting research possibilities. If participants significantly outnumber the available robots, sharing mechanisms [36, 40, 57] could help coordinate audience robot control. If the number of robots is so vast that they fill the space, the individuality of each robot would be less relevant and people would be able to move from one viewpoint to another seamlessly without noticing the transition. Finally, small aerial vehicles could accommodate larger workspaces consisting of multiple unconnected work locations.

# 8 CONCLUSIONS

This paper proposed Asteroids, a new approach for remote physical task demonstrations using swarms of mini-telepresence robots. We presented the interaction design for both the remote audience and the local demonstrator, and showed the affordances of Asteroids through usage scenarios. Our approach enables remote people to experience instructional sessions led by an expert demonstrator at a workbench scale and allows them to move around and control camera viewpoints to focus on points of interest. Furthermore, demonstrators can directly interact with physical embodiments of the audiences and use tangible interactions to control the demonstration flow.

We performed an exploratory user evaluation to study audience experience with Asteroids using an architectural model-building instructional task with groups of four participants in three sessions. The interactions enabled by our approach contributed to audience's engagement, and the tangibility of the participants' embodiments added to their sense of presence. Finally, participants were able to follow and observe the demonstrator's activities and recall relevant details presented during the demonstration.

## ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their feedback. This research is supported by Natural Sciences and Engineering Research Council of Canada (NSERC) grants RGPIN-2017-06415, DGDND-2017-00093, SMFSU 549149-2020, and IRCPJ 545100-18.

## REFERENCES

[1] Parastoo Abtahi, David Y. Zhao, Jane L. E., and James A. Landay. 2017. Drone Near Me: Exploring Touch-Based Human-Drone Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 34 (Sept. 2017), 8 pages. https://doi.org/10.1145/3130899

[2] Sigurdur Orn Adalgeirsson and Cynthia Breazeal. 2010. MeBot: A robotic platform for socially embodied telepresence. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 15–22. https://doi.org/10.1109/HRI.2010.5453272

[3] Felix Brandt, Vincent Conitzer, and Ulle Endriss. 2012. Computational social choice. *Multiagent systems* (2012), 213–283.

[4] William A. S. Buxton. 1992. Telepresence: Integrating Shared Task and Person Spaces. In *Proceedings of the Conference on Graphics Interface '92* (Vancouver, British Columbia, Canada). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 123–129.

[5] Jessica R. Cauchard, Jane L. E, Kevin Y. Zhai, and James A. Landay. 2015. Drone & Me: An Exploration into Natural Human-Drone Interaction. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 361–365. https://doi.org/10.1145/2750858.2805823

[6] Elizabeth Cha, Samantha Chen, and Maja J Mataric. 2017. Designing telepresence robots for K-12 education. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 683–688.

[7] Alexandra Gendreau Chakarov, Jeffrey Bush, Quentin L Biddy, Jennifer Jacobs, Colin Hennessy Elliott, and Tamara Sumner. 2021. Challenges and Unexpected Affordances of Physical Computing Going Remote. In *Interaction Design and Children*. Association for Computing Machinery, New York, NY, USA, 276–282. https://doi.org/10.1145/3459990.3460711

[8] Linfeng Chen, Kazuki Takashima, Kazuyuki Fujita, and Yoshifumi Kitamura. 2021. PinpointFly: An Egocentric Position-Control Drone Interface Using Mobile AR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 150, 13 pages. https://doi.org/10.1145/3411764.3445110

[9] Yan Chen, Walter S Lasecki, and Tao Dong. 2021. Towards Supporting Programming Education at Scale via Live Streaming. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–19.

[10] Marc Christie, Patrick Olivier, and Jean-Marie Normand. 2008. Camera Control in Computer Graphics. *Computer Graphics Forum* 27, 8 (2008), 2197–2218. https://doi.org/10.1111/j.1467-8659.2008.01181.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2008.01181.x

[11] Silvia Coradeschi, Amy Loutfi, Annica Kristoffersson, Gabriella Cortellessa, and Kerstin Severinson Eklundh. 2011. Social robotic telepresence. In *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 5–6.

[12] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies* 131 (2019), 81–98. https://doi.org/10.1016/j.ijhcs.2019.05.011 50 years of the International Journal of Human-Computer Studies. Reflections on the past, present and future of human-centred technologies.

[13] Naomi T Fitter, Nisha Raghunath, Elizabeth Cha, Christopher A Sanchez, Leila Takayama, and Maja J Matarić. 2020. Are we there yet? Comparing remote learning technologies in the university classroom. *IEEE Robotics and Automation Letters* 5, 2 (2020), 2706–2713.

[14] Terrence Fong, Charles Thorpe, and Betty Glass. 2003. Pdadriver: A handheld system for remote driving. In *IEEE international conference on advanced robotics*.

[15] C. Ailie Fraser, Joy O. Kim, Alison Thornsberry, Scott Klemmer, and Mira Dontcheva. 2019. Sharing the Studio: How Creative Livestreaming Can Inspire, Educate, and Engage. In *Proceedings of the 2019 on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19)*. Association for Computing Machinery, New York, NY, USA, 144–155. https://doi.org/10.1145/3325480.3325485

[16] Robert M Gagne and Walter Dick. 1983. Instructional psychology. *Annual review of psychology* 34, 1 (1983), 261–295.

[17] Tara Gallien and Jody Oomen-Early. 2008. Personalized Versus Collective Instructor Feedback in the Online Courseroom: Does Type of Feedback Affect Student Satisfaction, Academic Performance and Perceived Connectedness With the Instructor? *International Journal on E-Learning* 7, 3 (July 2008), 463–476. https://www.learntechlib.org/p/23582

[18] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47, 6 (2014), 2280–2292. https://doi.org/10.1016/j.patcog.2014.01.005

[19] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. In Touch with the Remote World: Remote Collaboration with Augmented Reality Drawings and Virtual Navigation. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology* (Edinburgh, Scotland) *(VRST '14)*. Association for Computing Machinery, New York, NY, USA, 197–205. https://doi.org/10.1145/2671015.2671016

[20] William W. Gaver, Abigail Sellen, Christian Heath, and Paul Luff. 1993. One is Not Enough: Multiple Views in a Media Space. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands) *(CHI '93)*. Association for Computing Machinery, New York, NY, USA, 335–341. https://doi.org/10.1145/169059.169268

[21] Cheng Guo, James Everett Young, and Ehud Sharlin. 2009. Touch and Toys: New Techniques for Interaction with a Remote Group of Robots. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 491–500. https://doi.org/10.1145/1518701.1518780

[22] Pavel Gurevich, Joel Lanir, Benjamin Cohen, and Ran Stone. 2012. TeleAdvisor: A Versatile Augmented Reality Tool for Remote Assistance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 619–622. https://doi.org/10.1145/2207676.2207763

[23] Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)* 11, 3 (2002), 411–446.

[24] Zorah Hilvert-Bruce, James T Neill, Max Sjöblom, and Juho Hamari. 2018. Social motivations of live-streaming viewer engagement on Twitch. *Computers in Human Behavior* 84 (2018), 58–67.

[25] Guy Hoffman and Wendy Ju. 2014. Designing Robots with Movement in Mind. *J. Hum.-Robot Interact.* 3, 1 (Feb. 2014), 91–122. https://doi.org/10.5898/JHRI.3.1.Hoffman

[26] Kari Hortos, Donald Sefcik, Suzanne G Wilson, John T McDaniel, and Eric Zemper. 2013. Synchronous videoconferencing: impact on achievement of medical students. *Teaching and learning in medicine* 25, 3 (2013), 211–215.

[27] Hiroshi Ishii, Daniel Leithinger, Sean Follmer, Amit Zoran, Philipp Schoessler, and Jared Counts. 2015. TRANSFORM: Embodiment of" Radical Atoms" at Milano Design Week. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. 687–694.

[28] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '97)*. Association for Computing Machinery, New York, NY, USA, 234–241. https://doi.org/10.1145/258549.258715

[29] Seungwoo Je, Hyunseung Lim, Kongpyung Moon, Shan-Yuan Teng, Jas Brooks, Pedro Lopes, and Andrea Bianchi. 2021. Elevate: A Walkable Pin-Array for Large Shape-Changing Terrains. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 127, 11 pages. https://doi.org/10.1145/3411764.3445454

[30] Brennan Jones, Anna Witcraft, Scott Bateman, Carman Neustaedter, and Anthony Tang. 2015. Mechanics of camera work in mobile video collaboration. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 957–966.

[31] Mitchell Karchemsky, JD Zamfirescu-Pereira, Kuan-Ju Wu, François Guimbretière, and Bjoern Hartmann. 2019. Heimdall: A Remotely Controlled Inspection Workbench For Debugging Microcontroller Projects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[32] Karen Kear, Frances Chetwynd, Judith Williams, and Helen Donelan. 2012. Web conferencing for synchronous online tutorials: Perspectives of tutors using a new medium. *Computers & Education* 58, 3 (2012), 953–963. https://doi.org/10.1016/j.compedu.2011.10.015

[33] Lawrence H. Kim, Daniel S. Drew, Veronika Domova, and Sean Follmer. 2020. User-Defined Swarm Robot Control. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376814

[34] Lawrence H. Kim and Sean Follmer. 2019. SwarmHaptics: Haptic Display with Swarm Robots. In *Proceedings of the 2019 CHI Conference on Human Factors in*

*Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300918

[35] Robert E Kraut, Susan R Fussell, and Jane Siegel. 2003. Visual information as a conversational resource in collaborative physical tasks. *Human–computer interaction* 18, 1-2 (2003), 13–49.

[36] Walter S Lasecki, Kyle I Murray, Samuel White, Robert C Miller, and Jeffrey P Bigham. 2011. Real-time crowd control of existing interfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. 23–32.

[37] Mathieu Le Goc, Lawrence H. Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, and Sean Follmer. 2016. Zooids: Building Blocks for Swarm User Interfaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 97–109. https://doi.org/10.1145/2984511.2984547

[38] Min Kyung Lee and Leila Takayama. 2011. "Now, i Have a Body": Uses and Social Norms for Mobile Remote Presence in the Workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 33–42. https://doi.org/10.1145/1978942.1978950

[39] Jiannan Li, Ravin Balakrishnan, and Tovi Grossman. 2020. StarHopper: A Touch Interface for Remote Object-Centric Drone Navigation. In *Proceedings of Graphics Interface 2020* (University of Toronto) *(GI 2020)*. Canadian Human-Computer Communications Society / Socieíteí canadienne du dialogue humain-machine, 317 – 326. https://doi.org/10.20380/GI2020.32

[40] Jiannan Li, Maurício Sousa, Ravin Balakrishnan, and Tovi Grossman. 2021. Constellation: A Multi-User Interface for Remote Drone Tours. In *Proceedings of the 9th International Conference on Human-Agent Interaction* (Virtual Event, Japan) *(HAI '21)*. Association for Computing Machinery, New York, NY, USA, 277–282. https://doi.org/10.1145/3472307.3484685

[41] Zhicong Lu, Michelle Annett, Mingming Fan, and Daniel Wigdor. 2019. "I Feel It is My Responsibility to Stream": Streaming and Engaging with Intangible Cultural Heritage through Livestreaming. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300459

[42] Zhicong Lu, Haijun Xia, Seongkook Heo, and Daniel Wigdor. 2018. You watch, you give, and you engage: a study of live streaming practices in China. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.

[43] Paul Luff, Christian Heath, Hideaki Kuzuoka, Jon Hindmarsh, Keiichi Yamazaki, and Shinya Oyama. 2003. Fractured ecologies: creating environments for collaboration. *Human–Computer Interaction* 18, 1-2 (2003), 51–84.

[44] Michal Luria, Samantha Reig, Xiang Zhi Tan, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2019. Re-Embodiment and Co-Embodiment: Exploration of Social Presence for Robots and Conversational Agents. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 633–644. https://doi.org/10.1145/3322276.3322340

[45] Julia M Markel and Philip J Guo. 2020. Designing the Future of Experiential Learning Environments for a Post-COVID World: A Preliminary Case Study. (2020).

[46] Francois Michaud, Patrick Boissy, Daniel Labonte, Helene Corriveau, Andrew Grant, Michel Lauria, Richard Cloutier, Marc-André Roux, Daniel Iannuzzi, and Marie-Pier Royer. 2007. Telepresence Robot for Home Care Assistance.. In *AAAI spring symposium: multidisciplinary collaboration for socially assistive robotics*. California, USA, 50–55.

[47] Terrance Mok and Lora Oehlberg. 2017. Critiquing physical prototypes for a remote audience. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 1295–1307.

[48] Ken Nakagaki, Joanne Leong, Jordan L. Tappa, João Wilbert, and Hiroshi Ishii. 2020. HERMITS: Dynamically Reconfiguring the Interactivity of Self-Propelled TUIs with Mechanical Shell Add-Ons. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 882–896. https://doi.org/10.1145/3379337.3415831

[49] Carman Neustaedter and Saul Greenberg. 2012. Intimacy in long-distance relationships over video chat. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 753–762.

[50] Carman Neustaedter, Gina Venolia, Jason Procyk, and Daniel Hawkins. 2016. To Beam or Not to Beam: A Study of Remote Telepresence Attendance at an Academic Conference. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) *(CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 418–431. https://doi.org/10.1145/2818048.2819922

[51] Gary M Olson and Judith S Olson. 2000. Distance matters. *Human–computer interaction* 15, 2-3 (2000), 139–178.

[52] Eric Paulos and John Canny. 2001. Social tele-embodiment: Understanding presence. *Autonomous Robots* 11, 1 (2001), 87–95.

[53] David Pitman and Mary L Cummings. 2012. Collaborative exploration with a micro aerial vehicle: a novel interaction method for controlling a mav with a hand-held device. *Advances in Human-Computer Interaction* 2012 (2012).

[54] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billinghurst. 2018. *Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173620

[55] Irene Rae, Bilge Mutlu, and Leila Takayama. 2014. Bodies in Motion: Mobility, Presence, and Task Awareness in Telepresence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2153–2162. https://doi.org/10.1145/2556288.2557047

[56] Irene Rae, Gina Venolia, John C Tang, and David Molnar. 2015. A framework for understanding and designing telepresence. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 1552–1566.

[57] Dennis Ramirez, Jenny Saucerman, and Jeremy Dietmeier. 2014. Twitch plays pokemon: a case study in big g games. In *Proceedings of DiGRA*. 3–6.

[58] Abhishek Ranjan, Jeremy P. Birnholtz, and Ravin Balakrishnan. 2007. Dynamic Shared Visual Spaces: Experimenting with Automatic Camera Control in a Remote Repair Task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1177–1186. https://doi.org/10.1145/1240624.1240802

[59] Wilko Schwarting, Javier Alonso-Mora, and Daniela Rus. 2018. Planning and Decision-Making for Autonomous Vehicles. *Annual Review of Control, Robotics, and Autonomous Systems* 1, 1 (2018), 187–210. https://doi.org/10.1146/annurev-control-060117-105157

[60] Yasaman S. Sefidgar, Prerna Agarwal, and Maya Cakmak. 2017. Situated Tangible Robot Programming. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. 473–482.

[61] Alexa F. Siu, Shenli Yuan, Hieu Pham, Eric Gonzalez, Lawrence H. Kim, Mathieu Le Goc, and Sean Follmer. 2018. *Investigating Tangible Collaboration for Design Towards Augmented Physical Telepresence*. Springer International Publishing, Cham, 131–145. https://doi.org/10.1007/978-3-319-60967-6_7

[62] Thomas Smith, Marianna Obrist, and Peter Wright. 2013. Live-streaming changes the (video) game. In *Proceedings of the 11th european conference on Interactive TV and video*. 131–138.

[63] Rajinder S Sodhi, Brett R Jones, David Forsyth, Brian P Bailey, and Giuliano Maciocci. 2013. BeThere: 3D mobile collaboration with spatial input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 179–188.

[64] Ryo Suzuki, Junichi Yamaoka, Daniel Leithinger, Tom Yeh, Mark D. Gross, Yoshihiro Kawahara, and Yasuaki Kakehi. 2018. Dynablock: Dynamic 3D Printing for Instant and Reconstructable Shape Formation. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 99–111. https://doi.org/10.1145/3242587.3242659

[65] Ryo Suzuki, Clement Zheng, Yasuaki Kakehi, Tom Yeh, Ellen Yi-Luen Do, Mark D. Gross, and Daniel Leithinger. 2019. ShapeBots: Shape-Changing Swarm Robots. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 493–505. https://doi.org/10.1145/3332165.3347911

[66] Anthony Tang, Michel Pahud, Kori Inkpen, Hrvoje Benko, John C Tang, and Bill Buxton. 2010. Three's company: understanding communication channels in three-way distributed collaboration. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 271–280.

[67] Balasaravanan Thoravi Kumaravel, Fraser Anderson, George Fitzmaurice, Bjoern Hartmann, and Tovi Grossman. 2019. Loki: Facilitating Remote Instruction of Physical Tasks Using Bi-Directional Mixed-Reality Telepresence. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 161–174. https://doi.org/10.1145/3332165.3347872

[68] Brygg Ullmer, Hiroshi Ishii, and Robert J. K. Jacob. 2005. Token+constraint Systems for Tangible Interaction with Digital Information. *ACM Trans. Comput.-Hum. Interact.* 12, 1 (mar 2005), 81–118. https://doi.org/10.1145/1057237.1057242

[69] Ana M Villanueva, Ziyi Liu, Zhengzhe Zhu, Xin Du, Joey Huang, Kylie A Peppler, and Karthik Ramani. 2021. *RobotAR: An Augmented Reality Compatible Teleconsulting Robotics Toolkit for Augmented Makerspace Experiences*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445726

[70] Lillian Yang, Brennan Jones, Carman Neustaedter, and Samarth Singhal. 2018. Shopping over distance through a telepresence robot. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–18.

[71] Matin Yarmand, Jaemarie Solyst, Scott Klemmer, and Nadir Weibel. 2021. "It Feels Like I Am Talking into a Void": Understanding Interaction Gaps in Synchronous Online Classrooms. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 351, 9 pages. https://doi.org/10.1145/3411764.3445240