Machine Gaze: self-identification through play with a computer vision-based projection and robotics system

RAY LC[1,2*], Aaliyah Alcibar[2], Alejandro Baez[2,3], Stefanie Torossian[4]

[1]City University of Hong Kong, School of Creative Media, Hong Kong.

[2]New York Hall of Science, New York, NY, USA

[3]New York University, Department of Mechanical Engineering, New York, NY, USA

[4]Queens College City University of New York, Department of Art, Flushing, NY, USA

*** Correspondence:**
RAY LC
luor840@newschool.edu

Keywords: play design, computational vision, participatory learning, interactive projection, creative computation, robotic art.

Abstract

Children begin to develop self-awareness when they associate images and abilities with themselves. Such "construction of self" continues throughout adult life as we constantly cycle through different forms of self-awareness, seeking to redefine ourselves. Modern technologies like screens and artificial intelligence threatens to alter our development of self-awareness, because children and adults are exposed to machines, tele-presences, and displays that increasingly become part of human identity. We use avatars, invent digital lives, and augment ourselves with digital imprints that depart from reality, making the development of self-identification adjust to digital technology that blur the boundary between us and our devices. To empower children and adults to see themselves and artificially intelligent machines as separately aware entities, we created the persona of a salvaged supermarket security camera refurbished and enhanced with the power of computer vision to detect human faces, and project them on a large-scale 3D face sculpture. The surveillance camera system moves its head to point to human faces at times, but at other times, humans have to get its attention by moving to its vicinity, creating a dynamic where audiences attempt to see their own faces on the sculpture by gazing into the machine's eye. We found that audiences began attaining an understanding of machines that interpret our faces as separate from our identities, with their own agendas and agencies that show by the way they temperamentally interact with us. The machine-projected images of us are their own interpretation rather than our own, distancing us from our digital analogs. In the accompanying workshop, participants learn about how computer vision works by putting on disguises in order to escape from an algorithm detecting them as the same person by analyzing their faces. Participants learn that their own agency affects how machines interpret them, gaining an appreciation for the way their own identities and machines' awareness of them can be separate entities that can be manipulated for play. Together the installation and workshop empower children and adults to think beyond identification with digital technology to recognize the machine's own interpretive abilities that lie separate from human being's own self-awareness.

1 Background

Development of self-awareness

The maxim of "Know thyself" has been touted since the Greek days by Protagoras, as it indicates ultimate understanding of our own identity and action that allows us to more objectively evaluate our influence on the world. Recognition of self-awareness and self-identity fosters understanding of our relation to ourselves and our society as children and adults. Experiments show that the affirmation that comes with self-awareness leads to increased compassion for one's own actions as well as increased positive social helping behavior following surprising incidents like an accidentally collapsing shelf (Lindsay & Creswell, 2014). Self-awareness increases the attribution of causality for negative consequences to the self (Duval & Wicklund, 1973), serving to deter blaming others and deflecting criticism. Publically suggesting self-awareness using a webcam reduces the bystander effect of not helping someone in need when other people are present (van Bommel et al., 2012). Self-awareness induced by a mirror even reduces aggressive action, whereas audience presence does not (Scheier et al., 1974). Thus self-awareness and identity go hand-in-hand with socially positive behaviors that promote integration in society.

The development of self-awareness and identity in children occurs in systematic stages that are often assayed using their response to seeing themselves in a mirror. Throughout the course of 5 years after birth, children go through eras of confusion, differentiation, identification, and meta-awareness in interactions with a mirror, characterized by what they do with their own bodies and objects placed in conjunction to them, such as post-its attached to their heads (Rochat, 2003). The last awareness stage involves how they present themselves publically, as if imagining how the mirror can be projected in the mind of others (Goffman, 1959). From 6 to 10 years old, children begin to consider alternatives to their own identities and at 10 years old, can even consider that her personality remains the same when her name is taken away (Guardo & Bohan, 1971), and incorporating awareness of another viewpoint's perspective into their own self-awareness (R. W. Mitchell, 1993). This development is thought to occurs in conjunction with biofeedback from parents, who present a reflective view for the child much like a mirror does in regulating her affective states (Gergely, 1996). The child begins to understand herself by seeing the way others see her. In particular, the awareness of not being seen gives rise to an identification of the self as apart from the others' gaze.

Self-awareness adaptation doesn't end with childhood. Reflexivity in social interactions in considering one's own current and past selves allows emerging adults to construct their self-identity in the counseling setting (Guichard et al., 2012). Self-awareness is also crucial in leadership development (Hall, 2004) and promoting well-being in jobs such as mental health professionals (Richards et al., 2010). Public self-awareness of adults in a controlled interaction is found to predict variables like social anxiety, self-esteem, and perception of others (Ryan et al., 1991), indicating its importance in determining self-competence and social success. This self-identity in adults is bound up with bodily awareness. Those who lose bodily awareness due to trauma or injury are meliorated using self-awareness-based touching and performance in psychological contexts (Fogel, 2009).

Technologies for self-awareness

Getting good at theatre and dramaturgy involves comparing one's actions to their perception, as well as working together with forces outside oneself. This has led to the use of ideas from theatre in teaching strategies for self-development. Studies have used collaborative theatrical projects to empower youths in such areas as creating meaning about the self (Beare & Belliveau, 2007), learning to improvise in hypothetical situations (Lehtonen, 2012), and achieving positive mental health (Ennis & Tonkin, 2015). One approach uses puppetry to enact fear, anger, sadness, and other emotion-based stories as part of a "feelings curriculum" to teach emotional awareness and self-comprehension to children (Maurer, 1977). These traditions leverage the way theatre forces individuals to reflect back on themselves upon identifying with actors in a scene. One system engages youths to use Twitter posts to emotionally affect physical actions of a puppet theatre installation using a robotic arm in a video, allowing them to reflect on their communication for development of self-awareness (Yamaguchi, 2018). Essentially theatre serves as an immersive version of a mirror that allows young people to gaze at their own actions and consequences as compared to those of others, driving a deeper meaning of what constitutes self-identity in the context of self-presentation. In particular, youths learn that social interactions involve presenting themselves in different ways in different contexts, much as actors play their roles in dramaturgy (Goffman, 1959). The practice of this self-presentation is made possible by both understanding the consequences of our own actions, and observing how others see us through their own lenses.

Interactive technologies for development of self-awareness have focused on vulnerable populations who have difficulty adjusting to societal norms due to their deficits in self-awareness, such as those suffering from communication and social disorders like autism and ADHD (Boucenna et al., 2014). Therapeutic strategies have included using touched-based devices to engage youths to foster development (Kagohara et al., 2013), applying virtual environments (such as VR cafes and buses) to allow youths to apply their social awareness skills incrementally without fear (P. Mitchell et al., 2007), creating serious games that effectively teach facial recognition in social situations (Serret, 2012), and utilizing social media platforms to enhance self-esteem by the way of profile identification (Gonzales & Hancock, 2010). Digital technologies of human-computer communication have been found to higher levels of private self-awareness compared to face-to-face communication, which heightened public self-awareness (Matheson & Zanna, 1988).

Of the various forms of communication technology, one of the most promising is robotics, for it enables physical interaction in addition to virtual enablement, bringing the private and public world of audiences to bear. Early studied focused on using robots to imitate child action, generating a sequence of motor actions that reproduces a detected human gesture (Berthouze et al., 1996). This work has modeled social interaction as observation followed by motor control, producing statistical models of motor representations that attempt to capture the human-robot interaction, exemplified by a study utilizing a game played by the robot Vince and its human interlocutor (Sadeghipour & Kopp, 2011). While simple actions can be approximated by robot movements, complex interactions that involve environmental constraints and rules require applying statistical learning theory to marginalize over the different possibilities in complex spaces for all possible movements, even in tasks as seemingly simple as putting objects into a box (Hersch et al., 2008). Recent work has modeled interactive tasks like tossing and catching arbitrary objects using both physics and computer vision to adaptively learn and generalize complex tasks (Zeng et al., 2020). One important contribution of related work is showing that using a game involving imitation with each other, human and robot become involved in feedback loops of reciprocal imitation, relying on human recognition and awareness on one hand and robot pose detection on the other (Boucenna et al., 2012). This begs the question of whether using simpler technologies like face detection is sufficient to elicit rich interactions that rely on human understanding rather than on complexity on the robotics side.

The use of robotics to elicit behaviors in human participants relies more on a rich interaction environment as opposed to a sophisticated computer vision detection model, due to the way humans are innately drawn to interpret even simple machine gestures as representing affective gestures analogous to human emotional behaviors (LC, 2019). Robots in this regard has taken such simple forms such as bubble-blowing agents (Feil-Seifer & Matarić, 2009), geospatial robots (Nugent et al., 2010), and gaze-directing toy (Keepon) (Kozima et al., 2007), all using simple interactions utilizing remote control of robot interactions to promote pro-social behavior. The effectiveness of the strategy comes not from the intricacy of the interaction, but rather the rich set of environmental cues and interpretations available to the child that makes the experience rewarding. One way to increase the interaction and immersion in the physical environment is by augmenting it with strategies like projection (Greene, 1986). Recent work has been able to projection map custom imagery onto complicated forms like faces (Bermano et al., 2017) and moving objects (Zhou et al., 2016), opening up possibilities for single-object projection experiences that respond to human interaction. It is possible to map robotic responses onto interactive objects much like an immersive form of computer based sculpture (Keskeys, 1994). The projection would then give voice to the robot via an external material, adding an additional layer of interaction capabilities as if the robot is controlling the external visual interaction based on audience feedback.

General approach

Given the considerations above, we decided to use the robot's own interpretive ability—its gaze—to show young audiences the process of self-awareness, allowing them to understand themselves by seeing the way machine sees them. We used a simple face detection interaction with a moving robot to engage young audiences to become aware of the self through looking at themselves on a responsive projection mapped face sculpture, relying on the innate human ability to interpret the interaction environment in an affective manner.

This approach leverages: (1) the way children learn of self-awareness through the way others see them, (2) the physical proxemics and performance-like interactions that robotics create to make this learning embodied in the real world, (3) the richness in self-gaze-directed interactivity provided by environmental augmentation through the mirror-like projected sculpture, and (4) the collaborative learning and play through workshops in multiple media and perspectives.

2 Materials and Methods

The experience consisted of the following main components: (1) a motorized security-camera-like robot that moves either casually on its own or in response to audiences to keep its gaze on a face in the crowd, (2) a projection system that maps the audience's own face onto a 3D face sculpture whenever the audience's face is detected by the robot, (3) a feedback screen that allows audiences to see what the machine is seeing, i.e. whether a face is detected, to interpret the machine's awareness of the audience, and (4) a workshop where audiences are asked to escape the machine's detection by putting on disguises, showing a comparison of being seen and not being seen as a way to reinforce the separation between self-awareness and lack thereof.

Exhibition

A set of four Appro and Panasonic CP414 security cameras (circa 1980) were cleaned, refurbished, and mounted on metal plates. Two of the cameras were further chosen for prototyping, with their internal fisheye cameras removed and replaced by webcams connected to an Intel NUC 7 (Windows 10) mini computer. The internal circuit was taken out, and the lens chassis was then reattached over the webcam. The body of the robot was constructed from a rotating base plate and an arm that tilts up and down at two different joints (Lewansoul kit), spray-painted silver upon completion. The three degrees of freedom (one in rotation, two in tilt) were controlled using three LDX-218 servo motors connected to a controller board, which was interfaced to an Arduino UNO board using custom routines. Figure 1 shows the look of the camera and body, which were designed to appeal to young audiences, to evoking playfulness and simplicity as opposed to traditional mechanized robots. The movements of the robot were similarly designed for serendipity, as sometimes the robot moved to fix its gaze on a face of the audience, while other times it simply moved side to side and up and down on its own. The video stream taken by the webcam was processed in Processing 3.3 using OpenCV. During the audience face tracking phase, distance from the center of the view to the center of the detected face was calculated live, and whenever the x or y distance was nonzero, a signal was sent from Processing to Arduino to move the appropriate motors in that dimension to point the camera directly at the center of the audience's face. When multiple faces were detected, the robot would direct itself at each face in succession after a one-second pause in position. At other times, a set of three predetermined movement routines had the robot scanning around the exhibition hall while occasionally moving forward or backward while maintaining similar angles of view. The narrative of the robot was that of a supermarket surveillance camera fortified with computer vision and repurposed to play and teach children about machine gaze and self-recognition.

A set of prototypes for the 3D face sculpture were made using different media: clay, paper mache, PLA (3D print), a mushroom-based polymer, and foamular (CNC). Figure 2 shows two experiments in sculpture construction. We decided ultimately to work with foam due to the ability to scale up in size, the lighter weight of the material, the ability to precisely craft the 3D look of the sculpture using CNC, and its ability to reflect projection imagery properly upon being painted. A 3D face model was constructed in Cinema4D, and one half of the face was transformed using the poly effect to look pixelated with large polygons. Thus the two sides of the face looked slightly different under projection of a face, with one side appearing more digitally manipulated than the other. The models were converted to stl format and printed on a 48x32x8 inch foam. The face was painted white to allow projection image to reflect, while the rest of the foam was painted black and mounted on a dark-colored podium (Figure 3). Canon LV8320 (3000 lumens) projectors were used to project face images from a ~40 degree angle above the setup (Figure 4). The image was projection mapped onto the face sculpture and controlled from the NUC 7 computer using the Kantan Mapper module from Touch Designer v099.

Completed views of the main interaction area are shown in Figure 3. The camera-mounted robot sat at the left of the projected sculpture. To its left was placed a live-view screen that showed the audience what the camera saw. When no faces were detected, the projection looped through a set of faces from the Chicago Face Database (chicagofaces.org) while the robot scanned the room. When a face was detected, Processing scaled the subject's face to the size of the projected image on the sculpture and used Spout to send the live video stream to Touch Designer to project onto the sculpture. The robot could follow the audience face by rotating or tilting during this interval so the image displayed was always dynamic. The size of the face projected on the sculpture was always the same regardless of the audience walking forward or away due to the scaling done in Processing. The image resolution is thus lower when the audience is farther away from the robot. When a face was found, a yellow square was also shown on the screen to the left superimposed on the camera's view. The complete system is diagrammed in Figure 4, and shown in audience view in Figure 5, both in prototype and final exhibition forms. Ambient lighting in the exhibition hall was turned down so that the projected image can be seen. Unfortunately this reduces the reliability of the computer vision. Thus two lamps were mounted, one for illuminating the side of the robot, the other for lighting the audience's face for proficiency of computer vision through the robot's webcam camera. The lighting was calibrated at the beginning of each day of exhibition (from May to September of 2019) to ensure optimal audience experience each day.

Workshop

A workshop opened to participants of all ages was created and presented 5 times at New York Hall of Science (NYSCI) by members of the museum's Explainers Program. At least half of the participants at each workshop were under the age of 18. Each workshop had 7-9 laptops with the capacity for 10-15 participants. The workshop began by asking subjects to draw what their own ideas of a face while focusing on features like eyes, nose, lips, and glasses. For the next 5 minutes, everyone showed their drawings to the crowd, and the workshop staff showed a computer-generated face from thisfacedoesnotexist.com, highlighting uniquely human features and discussing briefly how computers see human faces differently from us. We also outlined the main goal of the workshop to understand and play with the way machines see us. The next 5 minutes were spent getting a laptop setup and navigating a webpage that shows how poses can be detected by the computer vision on the webcam on the laptop. In this phase, participants could get out of their chair and move around to see how it affects the pose determination.

For the main part of the workshop (the remaining 25 minutes), we introduced how machines learn to recognize specific faces and how we can escape their detection, a fun activity for younger audiences. We showed audiences a custom script based on an existing p5 sketch we used to train a face classifier (https://editor.p5js.org/AndreasRef/sketches/BJkaHBMYm). First, the audience clicked a button repeatedly to take pictures of their faces with multiple samples. After training the program, we let the partcipants come in and out of the view of the webcam to verify that the machine learning algorithm has learned a representation of their faces. Workshop staff were available to fix any issues children had, but overall we were surprised by the amount of computer literacy displayed by the children.

Next we provided props like fake ears, hats, garments, mustaches, and jewelry to allow partcipants to dress up to escape the detection of the program despite being seen by the webcam (Figure 6). In this stage we showed how audiences can exist independently of the awareness of the machine. We let participants pick one outfit and train the program on the same person's face but as model for a different face. At this point, audiences could put on and take off their disguises and see the program recognizing different faces as different individuals (Figure 7). For example, one participant would train the program with his own face until it outputs "Danny" whenever his face is in front of the webcam. Then Danny would dress up as a football player and train the program to recognize the disguise as "Eli" (name of a well-known football player in New York). Then Danny would escape the program's detection of "Danny" by dressing up as Eli and vice versa. Throughout the process the workshop staff informed participants details about computer vision and machine learning. For example, we showed how taking many pictures (samples) were necessary to good recognition, the way different angles and conditions of a face for a given training made the algorithm more successful, and how these technologies were implemented in our own devices, etc.

After the workshop, we escorted the participants to the "Machine Gaze" exhibit (Figure 8), where they interacted with the robot and projected face sculpture freely before given a questionnaire that asked the following questions: "Where do you think the security camera comes from?", "What do you think the robot's purpose is?", "What do you think computer vision is?", "How do you think computers see us?". For a selected group of audiences, we followed the questionnaire with a qualitative interview to learn about their experiences in depth, asking them to elaborate on their reaction upon seeing their own image on the sculpture, how they managed to catch up with the robot's gaze when it stopped following their faces, how they interpreted their own image on the sculpture vs. what the machine sees (as shown on the screen), how they reacted to the machine moving between multiple faces being detected, where they allocated their attention when the displayed face switched from their own to that of another and vice versa, etc. The questionnaire answers were qualitatively coded into categories, tabulated and plotted in R 3.6.0. Finally, we passively observed audiences as they interacted with the exhibit, taking note of their tendencies, moments of joy, moments of confusion, and issues that arose. The interview questionnaire, and observation data were used to further refine the exhibit after the workshop ended and the main exhibition timeline began at NYSCI.

3 Results

Production and prototyping of the exhibition is seen here: https://youtu.be/V42towEXruk. Note the discretized movements of the robot tracking movement in 0:28. We decided to keep the discretized movements after audiences indicated in the first item in the questionnaire that it made they feel like the camera was made long time ago in "factory," and "corners in rooms." The prototyping also showed that due to the OpenCV xml template used, even animal and cartoon faces were detectable (1:05), further allowing audiences to identify the machine's particular method of perception as something separate from human faculties. The initial face images we projected were also not uniform enough to suggest a set of possible machine perceptions, so we replaced them with the photos from the Chicago Face Database. Finally, we realized from preliminary interactions that the camera tended to move between multiple faces in practice, so we set a timer of one second before it can move again during face tracking periods. Other materials/processes refined throughout the process included the material used to make the face sculpture, the lighting in the exhibition hall, the color of the podiums used, the speed of the robot movements, the number of projectors used, and size of the safety area around the robot, etc.

The full exhibition took place from May to September, 2019, with workshops kicking off the schedule in May. Documentation of audience interactions is here: https://youtu.be/kVoqkzZT4IQ. Our observation of the audience yielded three types of participants: (1) those curious about the device but refraining from making excessive contact with the machines (0:40), (2) those who take an active role to make expressive faces in engaging with the system (1:15), (3) those who bring others to the interaction by inviting them to the exhibit or enabling them to be in the view of the robot, creating a multi-face interaction (1:00). From our five days of observation, type (2) were the most numerous, with type (3) close behind, and perhaps exceeding type (2) on Sundays (the only weekend day we were observing). Interestingly, we found that group (1) audiences tended to come back to the exhibit at multiple points during their visits, as if they took the machine's guardian role quite seriously. Group (2) audiences tended also tended to make interesting discoveries in their interactions, such as using their hands to cover their faces so that the machine cannot see them (but they can see the machine move), and other pictures, people, and instruments in the environment as bait for the machine to focus its gaze on. Group (3) audiences included many parents who took their children in their arms while exploring the interaction together. They tended to initially guide the child's discovery, but very frequently ended up competing with them for the machine's attention.

The audience survey revealed interesting perceptions that we were initially unaware of (Figure 9). While most participants equated computer vision with some sort of camera seeing process (Yellow), some were associating it with recording or human-augmentation, topics that computer vision is associated in the popular culture with. Interestingly, audiences tended to assign machine intelligence to the robot system beyond simple mechanical processes. In answer to how the robot sees, most participants attributed its ability to some recognition capability beyond simply sensor reading or photography. We were also surprised to see that 3 of the 10 audiences surveyed also attributed the purpose of the machine to its curiosity or need for discovery, an inherently non-mechanical goal that assigns a human-like emotional content to the machine. One artistic audience member even draw some prospective logos for us. Remarkably, her drawings equated the shutter of the camera to the human eye, and its hardware with the human brain, again assigning anthropomorphic qualities to the machine. We believe this reaction is due to the ability of the machine to move in space, indicate emotions like curiosity, aversion, boredom, intelligence, and attention through movement and changes in projected content. This may drive a sense of the audience feeling perceived by a being aware of the audience's persistence. It also validates the use of robotics as a performance experience in evoking audience reaction.

Interactions from the workshop are shown here: https://youtu.be/pIRETXKZngg. For the face training phase, we saw that audiences liked to work as teams, usually with one member of the team (such as the parent) driving the others. Participants became creative with their interactions, such as turning around, glancing from beneath the table, and moving their face from side to side (0:48) as many ways to test the limits of escaping machine detection. We also observed parents teaching children about what it means to see their own image and how the machine interprets the face image (0:53). During the disguises section, we saw that the most popular items were hats (1:02). Frequently the participants helped each other put on the costumes and props and showed a feedback loop of asking for an opinion, then rearranging the props, and asking for opinion again, as if the questioner was using the opinion as a proxy for a mirror. Outlandish costumes were observed as well (1:11), because some faces did not easily escape the face detection algorithm, necessitating extreme measures. Interestingly, family members would sometimes wear matching outfits (1:18). This may be an indication of in-group affinity, but it could also indicate one member of the family teaching the other which disguises appear to be working. Generally the workshop was highly collaborative, with families working together and learning together. Finally, children tended to keep part of their disguises while visiting the exhibit (1:32). There was usually great excitement when seeing their own (disguised) faces appear on the 3D face sculpture, indicating their own shift in identity was registered by the perceiving system as well.

4 Discussion

Children's perception of being seen or not seen by external entities like mirrors and other people helps define their self-awareness. This identity is associated with their own self-presentation, which forms a performative behavior in public that in turn reinforces who they should be (Goffman, 1959). In this artistic intervention, we created a mirror-projection system that shows audiences their own faces, but only when interaction requirements are met, so that their perception of themselves are framed by what a machine sees, a form of performance in spatial interaction. We leveraged the prior demonstration of effectiveness in using robotics to help socialize children with communication disorders like autism (Boucenna et al., 2014) to create embodied physical actions that transform simply passive viewing to interactive behaviors that capture the subtleties of a self-perception-dependent form of performance. As audience interactions and experience shows, the exhibit leaves participants more aware of how machine perception works, how their own actions interacts with these perceptions, and how their own performance with the machines engender cooperative awareness of the limitations of each.

A first hint of these developments comes with the games that children invent while they interact with the robot. As detailed in Results, participants spontaneously perform games like covering their faces with their hands, making funny faces, seeing which of two faces the robot turns towards, etc. All these actions have a manifestation in the projected image on the 3D sculpture, some changing the detection of their face (covering with hands), some not changing the detection interaction (funny faces). The spontaneous development of these performative behaviors suggests an underlying learning process whereby children (and adults) acquire knowledge about whether they'll be perceived by the robot system based on the different performances they make. Their reaction to whether they are detected or not suggests an understanding of what the machine sees and how that relates to their concept of self. This understanding also seems to develop over the course of the interaction, with lack of understanding at first, followed by recognition of the machine gaze, then understanding of how they are perceived, and finally what they can perform to modulate this perception.

A second hint comes from the consistent attribution of human-like emotion, agenda, and behaviors to machines by audiences despite observing merely simple gestures, as previously studied (LC, 2019). The post-visit questionnaire results and exhibition audience observations both show the assigning of human-like characteristics to the machine. For example, the machine is deemed to be curious by a large contingent of observers, and subsequent drawings of the machine endow it with human characteristics like eye-sight. Audiences often treat the machine like human-like creatures both while it tracks their faces and when it ignores their faces. In the former they play movement games with it; in the latter they try to get its attention by moving towards the machine's eye voluntarily. This demonstrates that not only can machines track the human face, the human can track the machine face as well while trying to get its attention. This then creates a bi-directional interaction: if the audiences can see their own faces when the machine follows them, does the machine see its own face when they follow it? These internal models about how each entity observes and is aware of itself can provide educational moments for the participants themselves.

A third hint comes from workshop interactions, where participants specifically escape detection of the machine's gaze by dressing up as another. The dressing-up serves as a narrative approach to differentiating who one is and is not (Bamberg, 2011), showing the actors who they are by letting them experiment with situation where they are not perceived. This escape of detection may be critical in the audience's self-concept, for she is able to recognize that sometimes she won't be perceived by others if she only performed a certain way. It's as if she is playing a game of public performance akin to self-presentation that hides her own true identity in the context of robots and environments that are not sophisticated enough to understand this form of deception. More interventions will be necessary to show how these mini-deceptions and playful performances affect what participants think of themselves in the context of environmental modulations.

The use of environmentally enriched robotic interactions is promising in artistic and social design realms, both for treating those with communication issues and for creating interactive experiences for the general public. This exhibition showed one possible intervention in provoking audiences to examine what their self is by using physical embodied interactions with a computer vision-enabled camera that detects their face. These technologies provide possible future scenarios of more intimate interactions that takes into account more affective types of human data beyond face detection.

5 Figures

Figure 1: Robot head and body. (Left) The camera (head) was an APPRO model with lens and circuit replaced by a PC-connected webcam, mounted on steel plates. (Right) The body consisted of a steel frame joined by servo motors exhibiting three degrees of freedom, two of tilt and one of rotation, allowing the camera to face any direction in space.

Figure 2: Prototypes of the 3D face sculpture. (Left) A clay model with right side sculpted to be human face and left side a polygonal surface. The size required turned out to be prohibitively heavy. (Right) A reduced-size foamular model cut by CNC from an stl model and painted white to properly reflect projected image. The final exhibition model was approximately twice times the width and twice the height.

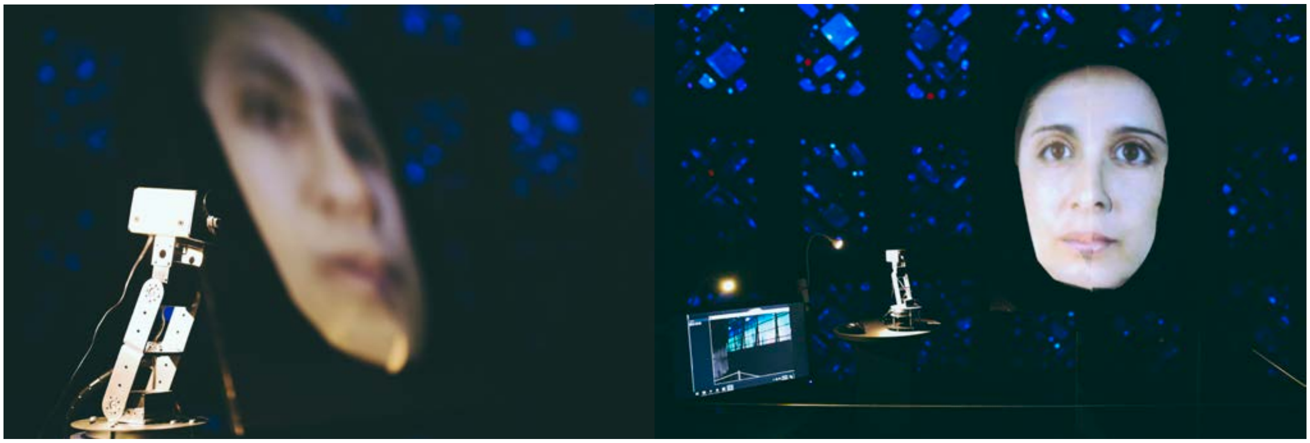

Figure 3: The exhibition setup. (Left) The camera-mounted robot sat on a dark-colored podium to the left of the face sculpture with the image of a face projected on it from approximately a 40 degree angle. (Right) The setup as viewed from an approaching audience, with a screen on the left showing the camera view from the perspective of the robot, and giving feedback to participants for when their faces were detected. One lamp lit the robot while the other lamp provided ambient lighting on the audience's face. The projected video on the face sculpture cycled between faces from the Chicago Face Database when no audience faces were detected, and a scaled version of the audience's face when it is detected by the webcam on the robot.

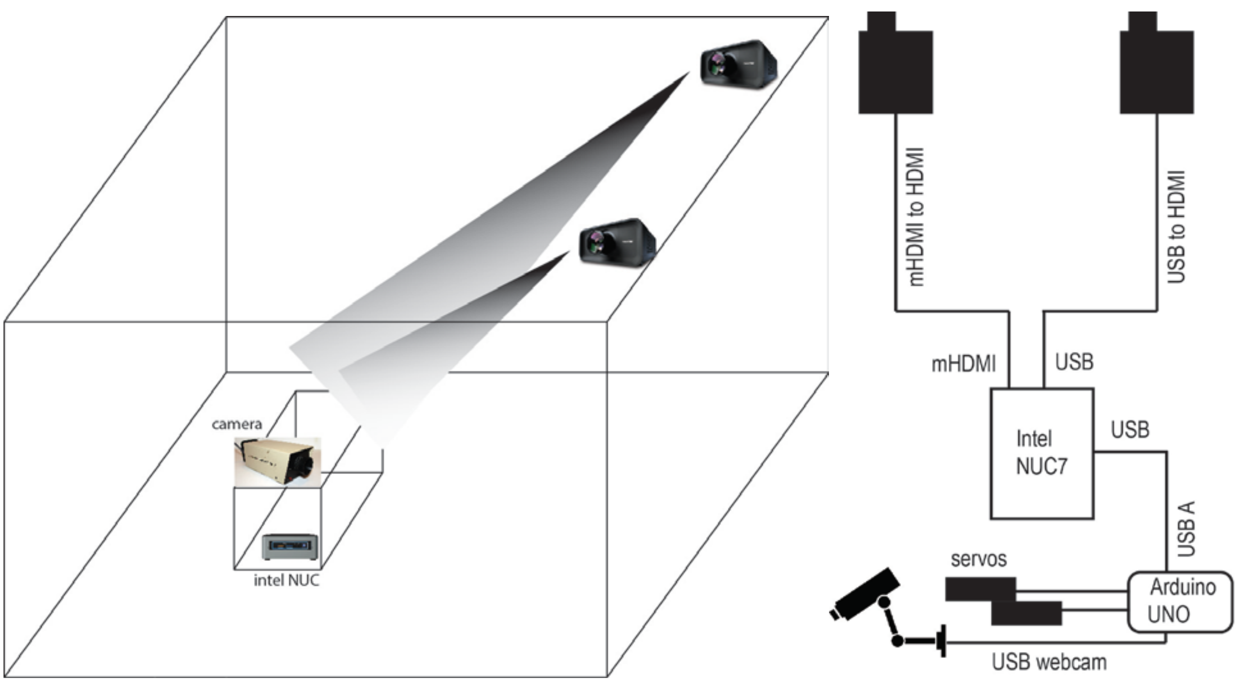

Figure 4: Exhibition plan. (Left) Projectors on railings were used to illuminate the face sculpture in the setup, while the NUC computer and motor board components were hidden in the inside of the cabinet. (Right) The connection diagram shows the NUC PC as the controller that integrated webcam input to decide whether to project a database face or a real face, and to direct the servo motors via arduino UNO how to move to keep the audience's face in the center. In other situations, it directed the robot to pan and tilt in a preprogrammed manner.

Figure 5: Audience interaction with the exhibit in prototype and finished form. (Left) Prototype stage interaction using a smaller face sculpture and brighter lamp to facilitate computer vision processes. (Right) A time during the final exhibition where the audience's face was detected, scaled, and projected onto the face sculpture. The projection mapping ensured the audience's face would be imaged on the face section of the sculpture. The audience's face, as seen from the robot's position, was shown on the screen to the left. At this stage, the robot followed the audience's face as it moved in space, as long as it was detected. When faces were no longer detected, the projection changed to flipping through the Chicago Face Database.

Figure 6: Workshop dress-up phase. (Left) Children selecting props, hats, decorations, and garments to wear that would allow them to escape the detection of a face classifier previously trained on their undecorated faces. (Right) A parent putting a fake mustache on her child after he put on football shoulder pads in an attempt to escape the computer vision's detection.

Figure 7: Workshop face-detection phase. (Left, Right) Children wearing disguises observing whether the p5 face classifier script running on the computer was able to distinguish between their real faces and their new disguises. Participants were able to vary the amount of disguises and how they were put on until the classifier detected them as unique faces.

Figure 8: Workshop exhibition phase. (Left, Right) Children were ushered to exhibition after the workshop and allowed to explore interactions with "Machine Gaze." They are currently looking into the robot's camera eye while also glancing to see if their face was detected by seeing whether their own faces appeared on the 3D face sculpture. Note that one child attempted to cover his face while looking through the slits between his fingers. The mustaches were left on by the children's choice.

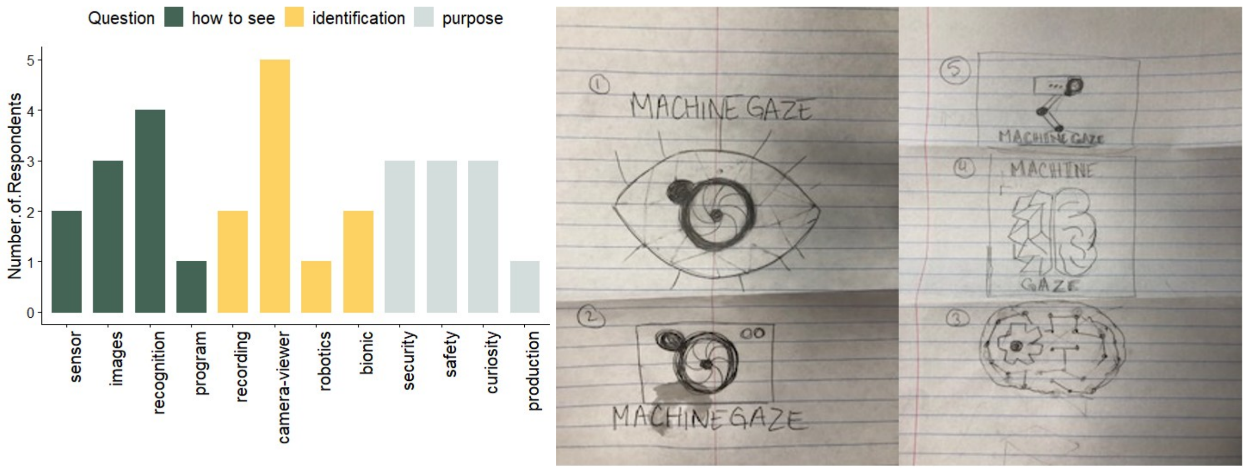

Figure 9: Audience experience during the exhibition. (Left) Distribution of coded answers to each pertinent question in the survey given after exhibition experience (n=10). (Green) Answers to "How do you think computers see us?" ranged from mentioning the camera's sensor abilities, by taking images, by recognizing people, and by using a computer program. (Yellow) Answers to "What do you think computer vision is?" ranged from computer as a recording device, to machine vision as a camera that views its environment, to robotics, to computer vision as a bionic device. (Grey) Answers to "What do you think the robot's purpose is?" included a role to protect security, a way to promote safety, as a curious machine, and for production of resource. (Right) Drawing by an young audience member that served as her interpretation of what the "Machine Gaze" exhibit meant to her.

6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

RLC, AA, AB, and ST created the exhibition. RLC and ST produced the figures. AA, AB, and ST ran the workshops and collected the data. RLC wrote the manuscript.

8 Funding

Funded by the New York Hall of Science (NYSCI) Designer-in-Residence 2019 program: https://nysci.org/event/machine-gaze/

9 Acknowledgments

Materials and facilities contributed by Parsons School of Design and New York University. The authors would like to thank Elizabeth Slagus and Erin Thelen (NYSCI Programs); Michael Cosaboom, Nolan Quinn, Sean Walsh, and Karl Szilagi (NYSCI Exhibits); Jeffrey Michael Geiringer and Philipp Schmitt (Parsons School of Design); and Truck McDonald (Explainers Program) for their support and assistance.

10 Data Availability Statement

The data generated from this study can be found at: http://www.raylc.org/machinegaze/. Video of audience interaction can be seen at https://youtu.be/kVoqkzZT4IQ. Video of production and prototyping is here: https://youtu.be/V42towEXruk. Video of the exhibition series is found here: https://vimeo.com/363395482. Material and footage collected from the workshop can be seen at: https://youtu.be/pIRETXKZngg.

11 References

Bamberg, M. (2011). Who am I? Narration and its contribution to self and identity. *Theory & Psychology*, *21*(1), 3–24. https://doi.org/10.1177/0959354309355852

Beare, D., & Belliveau, G. (2007). *Theatre for Positive Youth Development: A Development Model for Collaborative Play-creating*.

Bermano, A. H., Billeter, M., Iwai, D., & Grundhöfer, A. (2017). Makeup Lamps: Live Augmentation of Human Faces via Projection. *Computer Graphics Forum*, *36*(2), 311–323. https://doi.org/10.1111/cgf.13128

Berthouze, L., Bakker, P., & Kuniyoshi, Y. (1996). Learning of oculo-motor control: A prelude to robotic imitation. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS '96*, *1*, 376–381 vol.1. https://doi.org/10.1109/IROS.1996.570702

Boucenna, S., Delaherche, E., Chetouani, M., & Gaussier, P. (2012). *Learning postures through an imitation game between a human and a robot*. 1–2. https://doi.org/10.1109/DevLrn.2012.6400880

Boucenna, S., Narzisi, A., Tilmont, E., Muratori, F., Pioggia, G., Cohen, D., & Chetouani, M. (2014). Interactive Technologies for Autistic Children: A Review. *Cognitive Computation*, *6*(4), 722–740. https://doi.org/10.1007/s12559-014-9276-x

Duval, S., & Wicklund, R. A. (1973). Effects of objective self-awareness on attribution of causality. *Journal of Experimental Social Psychology*, *9*(1), 17–31. https://doi.org/10.1016/0022-1031(73)90059-0

Ennis, G., & Tonkin, J. (2015, May 25). *Youth arts and mental health: Exploring connections in the Top End*.

Feil-Seifer, D., & Matarić, M. J. (2009). Toward Socially Assistive Robotics for Augmenting Interventions for Children with Autism Spectrum Disorders. In O. Khatib, V. Kumar, & G. J. Pappas (Eds.), *Experimental Robotics* (pp. 201–210). Springer. https://doi.org/10.1007/978-3-642-00196-3_24

Fogel, A. (2009). *The psychophysiology of self-awareness: Rediscovering the lost art of body sense*. W W Norton & Co.

Gergely, G. W. (1996). The Social Biofeedback Theory Of Parental Affect-Mirroring: The Development Of Emotional Self-Awareness And Self-Control In Infancy. *Int. J. Psycho-Anal.*, *77*, 1181–1212.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday.

Gonzales, A. L., & Hancock, J. T. (2010). Mirror, Mirror on my Facebook Wall: Effects of Exposure to Facebook on Self-Esteem. *Cyberpsychology, Behavior, and Social Networking*, *14*(1–2), 79–83. https://doi.org/10.1089/cyber.2009.0411

Greene, N. (1986). Environment Mapping and Other Applications of World Projections. *IEEE Computer Graphics and Applications*, *6*(11), 21–29. https://doi.org/10.1109/MCG.1986.276658

Guardo, C. J., & Bohan, J. B. (1971). Development of a Sense of Self-Identity in Children. *Child Development*, *42*(6), 1909–1921. JSTOR. https://doi.org/10.2307/1127594

Guichard, J., Pouyaud, J., de Calan, C., & Dumora, B. (2012). Identity construction and career development interventions with emerging adults. *Journal of Vocational Behavior*, *81*(1), 52–58. https://doi.org/10.1016/j.jvb.2012.04.004

Hall, D. T. (2004, April 12). *Self-Awareness, Identity, and Leader Development*. Leader Development for Transforming Organizations; Psychology Press. https://doi.org/10.4324/9781410610102-18

Hersch, M., Guenter, F., Calinon, S., & Billard, A. (2008). Dynamical System Modulation for Robot Learning via Kinesthetic Demonstrations. *IEEE Transactions on Robotics*, *24*(6), 1463–1467. https://doi.org/10.1109/TRO.2008.2006703

Kagohara, D. M., van der Meer, L., Ramdoss, S., O'Reilly, M. F., Lancioni, G. E., Davis, T. N., Rispoli, M., Lang, R., Marschik, P. B., Sutherland, D., Green, V. A., & Sigafoos, J. (2013). Using iPods(®) and iPads(®) in teaching programs for individuals with developmental disabilities: A systematic review. *Research in Developmental Disabilities*, *34*(1), 147–156. https://doi.org/10.1016/j.ridd.2012.07.027

Keskeys, D. J. (1994). Computer sculpture—New horizons. *ACM SIGGRAPH Computer Graphics*, *28*(4), 255–258. https://doi.org/10.1145/193234.193240

Kozima, H., Nakagawa, C., & Yasuda, Y. (2007). Children-robot interaction: A pilot study in autism therapy. *Progress in Brain Research*, *164*, 385–400. https://doi.org/10.1016/S0079-6123(07)64021-7

LC, R. (2019). Secret Lives of Machines. *Proceedings of the IEEE ICRA-X Robotic Art Program*, *1*, 23–25.

Lehtonen, A. (2012). Future Thinking and Learning in Improvisation and a Collaborative Devised Theatre Project within Primary School Students. *Procedia - Social and Behavioral Sciences*, *45*, 104–113. https://doi.org/10.1016/j.sbspro.2012.06.547

Lindsay, E. K., & Creswell, J. D. (2014). Helping the self help others: Self-affirmation increases self-compassion and pro-social behaviors. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00421

Matheson, K., & Zanna, M. P. (1988). The impact of computer-mediated communication on self-awareness. *Computers in Human Behavior*, *4*(3), 221–233. https://doi.org/10.1016/0747-5632(88)90015-5

Maurer, C. G. (1977). Of puppets, feelings, and children. *Elementary School Guidance & Counseling*, *12*(1), 26–32. JSTOR.

Mitchell, P., Parsons, S., & Leonard, A. (2007). Using virtual environments for teaching social understanding to 6 adolescents with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(3), 589–600. https://doi.org/10.1007/s10803-006-0189-8

Mitchell, R. W. (1993). Mental models of mirror-self-recognition: Two theories. *New Ideas in Psychology*, *11*(3), 295–325. https://doi.org/10.1016/0732-118X(93)90002-U

Nugent, G., Barker, B., Grandgenett, N., & Adamchuk, V. I. (2010). Impact of Robotics and Geospatial Technology Interventions on Youth STEM Learning and Attitudes. *Journal of Research on Technology in Education*, *42*(4), 391–408. https://doi.org/10.1080/15391523.2010.10782557

Richards, K., Campenni, C., & Muse-Burke, J. (2010). Self-care and Well-being in Mental Health Professionals: The Mediating Effects of Self-awareness and Mindfulness. *Journal of Mental Health Counseling*, *32*(3), 247–264. https://doi.org/10.17744/mehc.32.3.0n31v88304423806

Rochat, P. (2003). Five levels of self-awareness as they unfold early in life. *Consciousness and Cognition*, *12*(4), 717–731. https://doi.org/10.1016/s1053-8100(03)00081-3

Ryan, R. M., Plant, R. W., & Kuczkowski, R. J. (1991). Relation of Self-Projection Processes to Performance, Emotion, and Memory in a Controlled Interaction Setting. *Personality and Social Psychology Bulletin*, *17*(4), 427–434. https://doi.org/10.1177/0146167291174011

Sadeghipour, A., & Kopp, S. (2011). Embodied Gesture Processing: Motor-Based Integration of Perception and Action in Social Artificial Agents. *Cognitive Computation*, *3*(3), 419–435. https://doi.org/10.1007/s12559-010-9082-z

Scheier, M. F., Fenigstein, A., & Buss, A. H. (1974). Self-awareness and physical aggression. *Journal of Experimental Social Psychology*, *10*(3), 264–273. https://doi.org/10.1016/0022-1031(74)90072-9

Serret, S. (2012). *Jestimule, a serious game for autism spectrum disorders*. https://www.semanticscholar.org/paper/Jestimule%2C-a-serious-game-for-autism-spectrum-Serret/83a5678d00bc0612348b1439b99c6f3bad3b3a4e

van Bommel, M., van Prooijen, J.-W., Elffers, H., & Van Lange, P. A. M. (2012). Be aware to care: Public self-awareness leads to a reversal of the bystander effect. *Journal of Experimental Social Psychology*, *48*(4), 926–930. https://doi.org/10.1016/j.jesp.2012.02.011

Yamaguchi, S. (2018). *Cybernetic systems for self-awareness @heyhexx: Interactive social media puppetry theatre for grasping emotions* [UCL Interactive Architecture Lab]. http://www.interactivearchitecture.org/cybernetic-systems-for-self-awareness-heyhexx-interactive-social-media-puppetry-theatre-for-grasping-emotions.html

Zeng, A., Song, S., Lee, J., Rodriguez, A., & Funkhouser, T. (2020). TossingBot: Learning to Throw Arbitrary Objects with Residual Physics. *ArXiv:1903.11239 [Cs, Stat]*. http://arxiv.org/abs/1903.11239

Zhou, Y., Xiao, S., Tang, N., Wei, Z., & Chen, X. (2016). Pmomo: Projection Mapping on Movable 3D Object. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 781–790. https://doi.org/10.1145/2858036.2858329