



Designing Robots with Movement in Mind

Guy Hoffman

Media Innovation Lab, IDC

Herzliya and

Wendy Ju

Center for Design Research, Stanford University

This paper makes the case for designing interactive robots with their expressive movement in mind. As people are highly sensitive to physical movement and spatiotemporal affordances, well-designed robot motion can communicate, engage, and offer dynamic possibilities beyond the machines' surface appearance or pragmatic motion paths. We present techniques for movement centric design, including character animation sketches, video prototyping, interactive movement explorations, Wizard of Oz studies, and skeletal prototypes. To illustrate our design approach, we discuss four case studies: a social head for a robotic musician, a robotic speaker dock listening companion, a desktop telepresence robot, and a service robot performing assistive and communicative tasks. We then relate our approach to the design of non-anthropomorphic robots and robotic objects, a design strategy that could facilitate the feasibility of real-world human-robot interaction.

Keywords: human-robot interaction, design, non-humanoid, non-anthropomorphic, gestures, movement, case studies, expressive movement

1. Introduction

As robots enter applications where they operate around, before, and with humans, it is important for robot designers to consider the way the robot's physical actions are interpreted by people around them. In the past, robots' actions were mainly observed by trained operators. However, if we are to deploy robots in more lay contexts—in homes and offices, in schools, on streets, or on stages—the quality of these robots' motion is crucial.

It is therefore time to consider more seriously the expressive power of a robot's movement, beyond the crude, choppy, and awkward mannerisms they are famous for. Instead we want to design robots with movements accurately expressing the robot's purpose, intent, state, mood, personality, attention, responsiveness, intelligence, and capabilities.

Any robot inherently displays in an interplay between its surface appearance and its physical motion. The way a robot looks sets the context for the interaction, framing expectations, triggering emotional responses, and evoking interaction affordances; however, movement is critical to conveying more dynamic information about the robot. The robot's movement in space can support

Authors retain copyright and grant the Journal of Human-Robot Interaction right of first publication with the work simultaneously licensed under a Creative Commons Attribution License that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.

action coordination, communicate internal states, and also has its own emotional impact. While recent years have seen an increasing respect for the importance and influence of robot appearance, far less attention has been paid to date on the design and effects of accurate expressive motion.

Designing robots with expressive movement in mind presents a number of interesting challenges for HRI designers and researchers. Significant iteration is required to understand how the robot's physical motion relates to its surface appearance. People's understanding of non-verbal factors is often tacit and intuitive, and new methods for evaluating designs and eliciting guidelines are needed. Finally, prioritizing movement and expressivity will create new engineering challenges that will likely require technological innovations to solve.

In this paper, we outline an expressive movement centric design approach for interactive robots. We delve more deeply into the rationale for well-designed expressive robot motion and discuss the **implications of prioritizing the communicative aspects of movement** in the design of robotic systems. We list specific techniques we employ to design robots with expressive movement in mind. To illustrate these points, we present **case studies** illustrating different methods used to design interactive robots. We conclude with **key challenges** of our approach, and why we believe that this approach can lead to simple non-anthropomorphic robot designs that might lower the barrier of entry for real-world human-robot interaction.

1.1 Contrast to Other Robot Design Approaches

Our approach differs from common design approaches in the robotics and HRI community, in particular, two differences that we denote are the **pragmatic** and the **visual** approach.

We use the word *pragmatic* after Kirsh and Maglio (1994), who distinguish between *pragmatic* and *epistemic* actions in human behavior and define *pragmatic actions* as “actions performed to bring one physically closer to a goal.” Similarly, we will distinguish between the **pragmatic movements** of a robot, those that are aimed at **achieving a physical goal**, and its **expressive movements**, those aimed at communicating the robot's traits, states, and intents to human interaction partners.

A *pragmatic* design approach sets out from specifications required of the **robot's spatial activity towards physical goals**, as defined by users of the system. Mechanical engineers design the robot's parts and relationships to fulfill these requirements as efficiently as possible, both in terms of energy and of cost. The resulting design is usually an assembly of limbs with more or less exposed links, actuators, and cables. For reasons of mechanical optimization, these limbs are often structured as chains of cantilevers from a rotation point, and follow principles of symmetry, orthogonality, and concentric relations. In some cases, a shell is designed post-hoc to cover internal parts and achieve a certain “look” for the robot. The shape and structure of the shell is highly constrained by the existing core of the robot, and usually follows its lines and proportions closely.

A *visual* approach is to **design robots with appearance in mind**. This is common for robots intended specifically for expressive interaction, as well as for entertainment robots. Industrial design practices are used to develop the look of the robot through a variety of **sketching and modeling techniques**. The robot's parts are **specified by the robot's users** to support the intended interaction, which often includes gaze, smiles, pointing gestures, etc. Visual designers of robots then need to decide how far along the humanoid spectrum to place their design (Fink, 2012). This can affect the choice of shape, materials (e.g., metal or silicone), body and facial parts (DiSalvo, Gemperle, Forlizzi, & Kiesler, 2002), as well as the amount of detail sculpted into the robot's form.

In both cases, the robot's expressive quality of **movement** is developed later in the process, if at all. Once the robot is completed, a **3D model** of the robot is usually generated. This model is then used to programmatically develop the way it moves. In some cases, the robot is modeled in a 3D animation program and a pipeline is created to translate the animation from the 3D model to the physical robot (e.g., Gray, Hoffman, Adalgeirsson, Berlin, & Breazeal, 2010).

In contrast, designing interactive robots from their expressive movement up takes the **communicative power of movement** into consideration early on in the design process. This is not to say that **visually aesthetic and pragmatic considerations** are not taken into account. Using an

iterative design methodology, these are factored in throughout the design process, as demonstrated in the case studies below. However, the expressive nature of the robot's movement is not added on after the robot is designed, or—more commonly—completely built. Instead, it is factored in from the onset and converses with both the visual and the pragmatic requirements of the robot.

In addition to its outcomes in terms of expressivity, we have found that this design approach can lead to a different kind of robot than that usually found in research labs: One that displays formal simplicity and abstract geometric shapes, while exhibiting its complexity and sophistication primarily through carefully designed movement qualities.

2. Why Movement Matters

Regardless of the appearance of an object or organism, its movement is a **powerful interaction and expression medium**. Humans, like most animals, are highly sensitive to perceived motion, which is a fact utilized by conspecifics, other animals, and designers of artifacts.

In interaction, humans use **kinesics**—the generation and interpretation of **nonverbal behavior** expressed as movement of the body—to communicate information, reason about mental states, accompany speech, denote real and imaginary objects in the world, express emotions and attitudes, self-present, accomplish rituals, and more (Argyle, 1988; Ekman & Friesen, 1969).

People also use **proxemics**—spatial distance—to define themselves in relation to their environment and to people around them, to set up expectations and communication channels, establish boundaries, claim territory, and signal context (Hall, 1969). Movement and gestures are important to the coordination and performance of joint activities, where they serve to communicate intentions and refer to objects of common ground (Clark, 2005). A significant body of work concerning the analysis of nonverbal acts and resulting perceptions exists, e.g., Knapp and Hall (2002); Moore, Hickson, and Stacks (2010).

2.1 Human sensitivity to abstract motion

Human sensitivity to motion is not limited to the perception of other humans, but applies to the movement of **abstract shapes** as well. Research in point-light displays shows that humans are inclined to, and extremely capable of, extracting information from a minimal set of visual features. In these studies, actors are fitted with point lights at joint positions and recorded as a sequence of white dots on a black background (Figure 1). Participants in these studies have been able to extract complex information, such as activity classification (Johansson, 1973; Thornton, Pinto, & Shiffrar, 1998), recognizing specific individuals (Loula, Prasad, Harber, & Shiffrar, 2005), distinguishing gender (Kozlowski & Cutting, 1977), and more. This demonstrates the expressive power of accurate motion even in the absence of static visual cues.

Beyond classification and recognition, humans also tend to **assign internal states and intentions to abstract movements**. This is part of the human inclination to attribute intention to animate and inanimate objects (Baldwin & Baird, 2001; Dennett, 1987; Malle, Moses, & Baldwin, 2001), a capability usually referred to as *Theory of Mind* (Baron-Cohen, 1991).

For abstract movements to be interpreted as intentional, they do not need to follow human form in detail, as in the case of point-light displays. Humans readily recognize, classify, and attribute intention even to purely abstract moving shapes. In their seminal work, Heider and Simmel (1944) found that when shown an animated film including dots, lines, and triangles moving in an intention-suggestive way, only one participant out of 34 described the film in geometric terms. Instead, most used person-like language, and over half used a narrative, including human actors, emotions, and intentions. This tendency has been extensively replicated and analyzed in detail over the years (for a review see: Barrett, Todd, Miller, & Blythe, 2005; Gao, Newman, & Scholl, 2009; Scholl & Tremoulet, 2000). Blythe, Todd, and Miller (1999) also showed that viewers could recognize intentions animated by other people, even when they were enacted on simple two-dimensionally moving non-human forms. These attributions of causality and animacy are considered to be an **automatic and irresistible part of visual processing**, from a very **early developmental stage**, and

across cultures, despite the fact that they involve impressions typically associated with higher level cognition (Gao et al., 2009; Michotte, 1946).

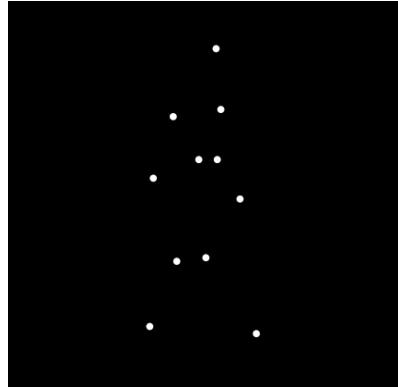


Figure 1. Point light displays are attained by attaching lights to joints of actors, and recording their movement. The resulting video consists of a sequence of moving white dots on a black background, similar to the depicted frame.

2.2 Movement Design in Character Animation and Robotics

Based on this propensity, movement has been used as a communication medium not only by natural organisms, but also by manufactured objects, simulated characters, and robots. Throughout the history of character animation, for example, animators have used motion trajectories of human, animal, and object shapes to convey character, narrative, and emotion expression (Lasseter, 2001; Thomas & Johnston, 1995).

At times, the animated character was purely movement. A famous example is “*The Dot and the Line*” (Jones, 1965), in which complex narrative and character development are implemented using two very simple characters, a dot and a line. All expression is performed in motion, timing and staging alone, earning the film a special place in animation history for distilling the core principles of expressive cartoon acting with few visual aids. In fact, a classic exercise in animation school is to portray two different characters using the same shape (usually a circle) and differentiating the characters only by their movement.

In the field of 3D character animation, John Lasseter produced “*Luxo Jr.*” (Lasseter, 1986, 1987) which features two simple desk lamp characters. For lack of appropriate 3D modeling and animation software, Lasseter used only simple three-joint rigid characters with an articulated neck and expressed both character and narrative through rigid motion in these joints alone. Here, too, the appeal of the protagonists lies in their movement rather than in their appearance.

We cite these examples from character animation as relevant to robot design, even though they do not describe physically embodied characters, as they have been designed for the sole purpose of communicating to an audience. We believe that comparable devotion to the design of physical robot motion can yield similarly successful results.

2.2.1 Robots with Movement in Mind In recent years, several robots have made use of accurately designed motion for expression and interaction. One example is Keepon, a small robot originally designed for interaction with autistic children but also used in expressive musical behavior (Kozima, Michalowski, & Nakagawa, 2009; Michalowski, Sabanovic, & Kozima, 2007).

Movement has been considered in the context of robots and theater in a number of works and position papers (Hoffman, 2005; Hoffman, Kubat, & Breazeal, 2008; Knight, 2011; Murphy et al., 2010). In early robotic installation art, abstract motion has also been used for expressive effect

and interactive communication with audiences (Inhatowicz, 1970). Recently, the motion qualities of a quadcopter have been explored in light of a system of movement used in improvisation theater (Sharma, Hildebrandt, Newman, Young, & Eskicioglu, 2013). Human-robot proxemics has been investigated in a number of research projects (Alami, Clodic, Montreuil, Sisbot, & Chatila, 2005; Mumm & Mutlu, 2011) mostly with respect to distance and path, but without specific attention to movement quality.

2.3 Movement as Dynamic Affordances

However, as humans interact more with robots, this kind of performative and communicative motion becomes important beyond the context of theater. A robot's motion can clue users into what actions and interactions are possible.

In the design of consumer objects and user interface, these cues are called "affordances," defined as qualities designed into objects and interfaces to help cue users on potential actions and interactions that a designed object is capable of (Norman, 1999). There, affordances are usually based on the perception of unchanging features. However, Gibson's original conception of affordances had a far more dynamic aspect: affordances indicate possibilities for action (Gibson, 1977) and hence the perception of affordances is dependent not only on symbolic or formal markers but also by dynamic motions that help actors identify the locus and range of action—see, for example, the recent design of the automatically extending door handles of the Tesla S automobile or that of automatically opening doors (Ju & Takayama, 2009).

In that context, Tomasello (1999) writes that the perception and understanding of these affordances are not inherent or static but rather socially constructed. Robots can thus perhaps move in ways to aid the social construction of possibilities for interacting with them.

3. Techniques for Movement-Centric Design

Based on the above, we propose to design robots along a process anchored in their expressive movement. As presented in our discussion, this approach could lead to mechanically and formally simple robots that display sophistication in the way they move instead of in the way they look.

In support of this design process, we have been using a variety of techniques:

3.1 3D Animation Gesture Studies

In many of our design processes, we use 3D animation studies to accurately design robot movement and to evaluate the expressive capabilities of the robot-in-the-making. These studies are conducted with no concern for mechanical considerations. Instead, motion is key, and often the only factor driving this exploratory stage, which can be thought of as motion sketches for the robot. Through a host of iterative modeling and animation steps, we explore a wide range of DoF placement and movement possibilities vis-a-vis the expression and functionality required by the robot. These studies are also used to evaluate different shapes for the robot, insofar as they support the movement described in the animations.

We conduct animation studies on a spectrum of detail realizations. In some cases, purely abstract shapes are animated to find an effective DoF setup supporting required behaviors and attitudes. In other cases, we use rough robot forms to simulate how the form moves through the animated behaviors. Close-to-final detail models are also used to evaluate whether the expressive movements and behaviors set out in the robot's requirements were actually met by the design. The output of this technique is a collection of clips that can then be evaluated by designers, users, and audiences alike (see also: Section 3.4).

3.2 Skeleton Prototype

Before building the full robot, a skeletal prototype of the robot can be used to closely mimic the robot's movement in physical space. Such prototypes can be rapidly cut out of wood or prototyped

using 3D printing methods and can be assembled in a matter of days with little waste in material. In some cases the skeleton is actuated, and in others it is passive.

A skeleton prototype enables both the designer and the computer scientists working on the robot's behavior to evaluate solutions for the final robot, without committing to the exact details of the robot's exterior shape and finish. Scale models of the robot can be used for movement prototyping in case the completed robot is too large to model at full scale.

3.3 Wizard of Oz

Wizard of Oz (WoZ) is a technique for prototyping and experimenting dynamically with a system's performance that uses a human in the design loop. It was originally developed by HCI researchers in the area of speech and natural language interfaces (Kelley, 1983) as a means to understand how to design systems before the underlying speech recognition or response generation systems were mature. However, it is particularly apt for human-robot interaction (Riek, 2012) and design (Maulsby, Greenberg, & Mander, 1993). In contrast to the prevalent use of WoZ in HRI, we use Wizard of Oz not to fake technologies that have yet to reach maturity, but to explore the wide range of possibilities for how their behavior and movements could be designed. We often involve users in a real-time collaborative design exploration using Wizard of Oz techniques.

The incorporation of the human-in-the-loop allows for improvisation (Akers, 2006) and spontaneous selection of different alternatives. It can inform system engineering, for example by identifying which features and gestures need to be recognized by a computer vision system (Höysniemi, Hämäläinen, & Turkki, 2004), or which sensors best predict contextual needs (Hudson et al., 2003). It can also be used with built-in evaluation systems so as to automatically reach design outcomes. Unlike Wizard of Oz techniques employed in experimental studies, here, the user of the system is often aware of the Wizard and interacts under a shared suspension of disbelief rather than under deception.

3.4 Video Prototyping

Video prototyping builds on all of the previous techniques by using recordings of physical movement performance, actuated skeleton movement, or rendered clips from 3D animation studies. The addition of video to the improvisational aspects of the above methods allows the designer to be more collaborative—to build a corpus of movements and gestures to reflect, classify, categorize, and refine. The video also allows more controlled experimentation, as designers can make sure that multiple people can view the exact same motion, or that the same people can evaluate varying versions of interaction (Woods, Walters, Koay, & Dautenhahn, 2006). Hendriks, Meerbeek, Boess, Pauws, and Sonneveld (2011) interviewed participants who had just watched a video prototype to gain insights into user perception. Crowdsourcing has also been used to scale up the number of people giving assessment and feedback on the interactions depicted in the videos (Ju & Takayama, 2009).

Video prototyping can be done on a number of detail levels: on completely abstract shapes, such as flat elements blocks or cardboard cutouts (Vertelney, 1995); on virtual animated models of the robots (Takayama, Dooley, & Ju, 2011); on skeletal models; or by using the final built robot, using either Wizard of Oz puppeteering techniques or autonomous robot behaviors. In addition, video prototypes can be set in various locales, which can be important in situations where there is a contextual element to how the motion would be interpreted (Sirkin & Ju, 2012). When using filmed video, as opposed to simulation or animation, the designer has to be concrete in terms of scale, background, speed, and limits of the robot. As a result the movement of the robot is necessarily evaluated with relation to the context it is in.

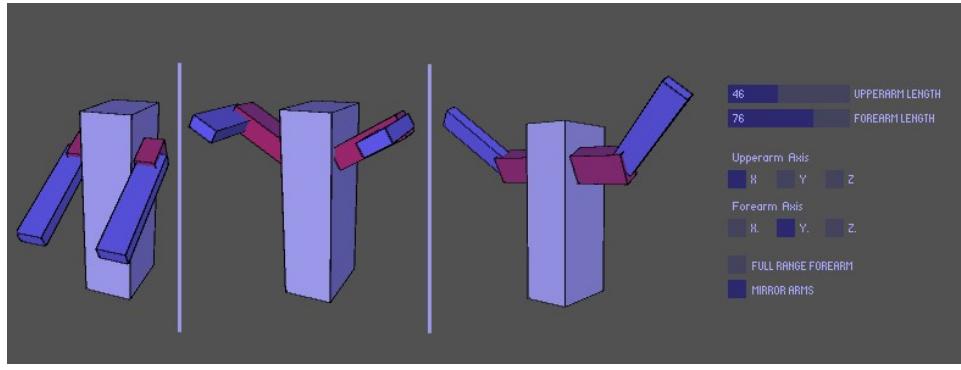


Figure 2. Screenshot from an interactive DoF configuration tool, used to explore DoF positions, link lengths, and relative orientations. The program screen is shown in the right panel of the Figure. For illustration purposes, we have included two additional screenshots of the resulting configuration using different parameters.

3.5 Interactive DoF Exploration

As part of the design of a new robot, we have developed a software tool for the interactive exploration of DoF configuration for expressive robots. This tool complements the use of 3D modeling and animation mentioned above. Instead of an iterative process of modeling the robot, and subsequently animating the structure, this tool continuously plays through a given gesture over and over while enabling the designer to change the DoF configuration in real time and view the resulting expressive qualities.

In a first implementation, seen in Figure 2, we configured the tool to be used to design a two-arm, two-DoF-per-arm robot used for emotional expression. This is a case of high abstraction in terms of the robot's form. Geometric shapes represent each DoF link. The configuration parameters were the relative chain placement of the two DoFs in the arms, the link length, and the joints' relative orientations. The gestures explored here were a number of short cyclical moves with the intention to express various emotions, such as happiness, sadness, and anger.

4. Case Study I: Marimba-Playing Robot Head

The remainder of the paper discusses four case studies putting expressive movement-based design into practice, exemplifying the techniques described above and reporting on insights and outcomes of the proposed design process.

The first case study is a socially expressive and communicative head for Shimon, an interactive robotic marimba player (Bretan, Cicconet, Nikolaidis, & Weinberg, 2012; Hoffman & Weinberg, 2011; Weinberg & Driscoll, 2007).

Shimon is a platform to explore and research robotic musical improvisation in human-robot joint ensembles. It has been used for a number of robotic musicianship-related research projects, in which the robot listens to a human musician and continuously adapts its improvisation and choreography, while playing simultaneously with the human.

4.1 Prior Design

Originally, Shimon was constructed as a music-playing module only, with mostly pragmatic design considerations, specifically to be able to produce high density and quickly changing note sequences. The robot is comprised of four arms, each actuated by a linear actuator at its base, and running along a shared rail. Each arm contains two rotational solenoids controlling two marimba mallets, one for the bottom-row (“white”) keys, and one for the top-row (“black”) keys. Figure 3 shows two views of the robot.



(a)

(b)

Figure 3. Overall (a) and detail view (b) of the robotic marimba player *Shimon*. Four arms share a voice-coil actuated rail. Two rotational solenoids per arm activate mallets of varying firmness.

4.2 Motivation for a Socially Expressive Head

While supporting several research outcomes in human-robot joint musicianship and improvisational intelligence (Bretan et al., 2012), the robot's functional design lacked socially expressive and musically communicative capabilities. After all, musicianship is about not just note production but also communicating gesturally with the audience and with other band members.

This led to the decision to design an additional channel of embodied and gesture-based communication, in the form of a socially expressive head. The robot would use the head to communicate internal states, such as rhythm or emotional content and intensity. Other head gestures could be used to manage turn-taking and attention between the robot and human musicians, supporting synchronization and joint musical interaction. Finally, a functional consideration was to add vision to the robot's perceptual system, through the use of a built-in camera. This would enable *Shimon* to detect social cues related to the musical performance and respond to them.

In specifying the overall design direction of *Shimon*'s head, we decided early on against a humanoid head. The motivation was in line with the argument for abstract moving shapes we make in the Discussion section. In addition, we felt that it would make more sense to match the aesthetics of the arm mechanism by an equally mechanical-looking appearance.

4.3 Design Process

The design process of *Shimon*'s head included five stages: (a) abstract 3D animation exploration; (b) freehand appearance sketches; (c) detailed DOF placement animation exploration; (d) scale exploration models; and (e) final robot solid design and construction. Note the iterative alternation between movement and form design in this process.

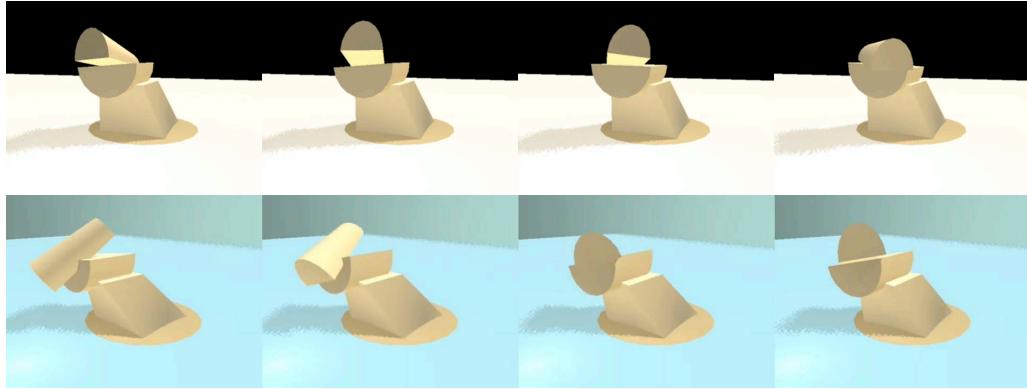


Figure 4. Still frames from abstract animation explorations for *Shimon*'s head design.

4.3.1 Abstract 3D Animation Exploration To evaluate the number of degrees of freedom necessary for the intended musical social expression as well as their spatial relationship, the design process started with a series of animation explorations using abstract 3D volumes, as described in Section 3.1.

We positioned a number of cut-off cylinders and trapezoids in different configurations and attempted a set of expressive gestures using these cylinders. Examples were: responding to the musical beat, showing surprise at an external event or shifting attention from one band member to another. Figure 4 shows several frames from these explorations. We tried out different parameters for each of the following variables: number of DoFs, DoF hierarchy, and DoF relative orientation. For each configuration, we tried to express the desired behaviors and evaluated intuitively how readable and expressive the specific configuration plays out to be.

Our exploration suggested four degrees of freedom: two in the neck (pan/tilt), one tilt below the neck, and one pan at the base of the whole structure. Given the static position of the head with respect to both instrument and band, we felt that much of the required expressivity could be achieved using this minimalist approach.

4.3.2 Freehand Appearance Sketches In parallel to the animation tests, we developed the appearance of the robot through a number of freehand sketches. The aim of the sketches was to explore the appearance possibilities unconstrained by mechanical considerations. Figure 5 shows some of the sketches, exemplifying the gamut of design directions explored at this stage.

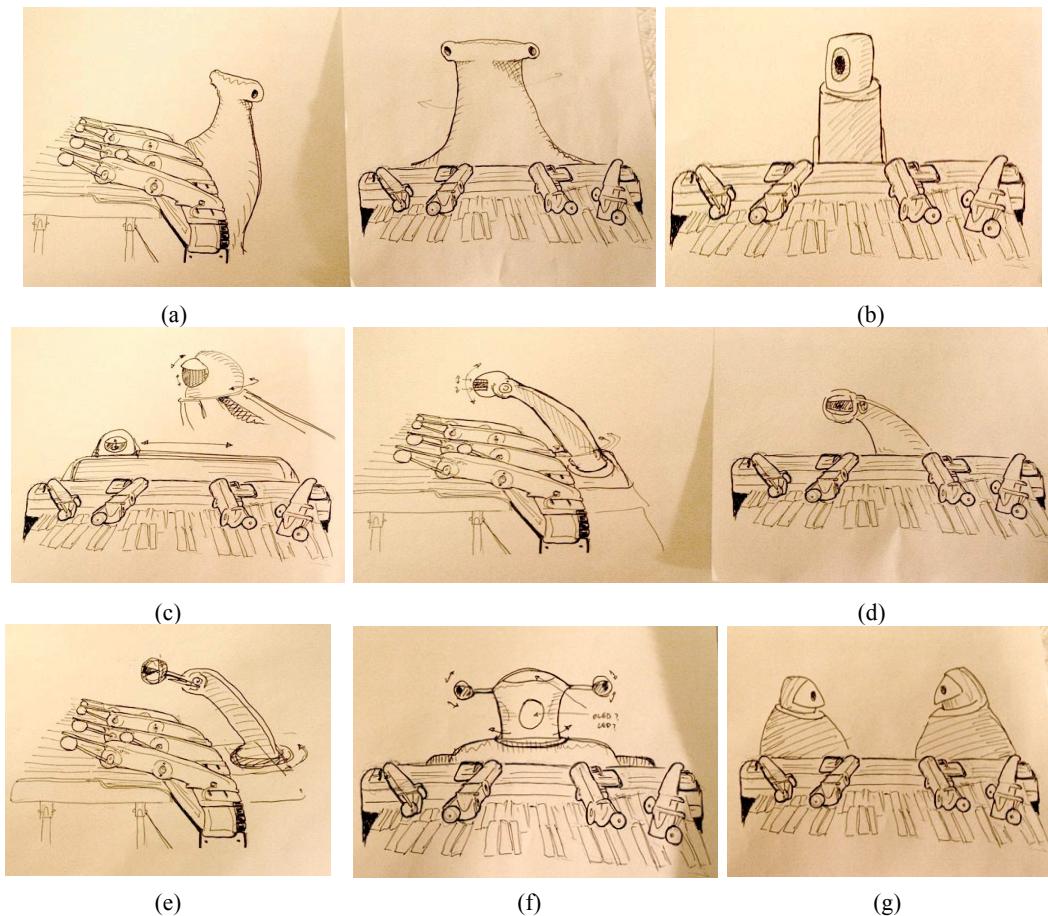


Figure 5. Freehand sketch samples for Shimon head design.

Eventually, the sketching process converged on an interpretive reproduction of the robot's existing mallet arms. Resisting more creature-like appearances, the design centered around an arm-like head using the same materials and rough proportions of the arms as well as mirroring the round mallet heads in an equally proportioned round head (Figure 5d and 5e). In one of the sketches, a headphone-like design emerged when considering the placement of the head tilt motors, which was a lucky coincidence that matched the musical nature of the robot (Figure 8b).

Most of the designs steered clear of explicit facial features. To compensate for this lack, and in order to restore expressive capabilities, we opted for a more abstract emotional display in the head. In part inspired by the expressive capabilities of the iris design of AUR (Hoffman & Breazeal, 2009), an opening mechanism was introduced into the head's design (first indicated in the sketch depicted in Figure 5d, left). The idea was that opening and closing a central "space" in the robot could be attached to a range of emotional meanings. We tied this design to the opening and narrowing of eyes and mouth in human faces associated with a variety of emotional states (Ekman & Friesen, 1969).

After testing several configurations and scales for the opening feature, we decided on a coupled garage-door-like design, in which both top and bottom covers would open simultaneously. This solution was economical but still sufficiently expressive—as we had found in another series of animation tests (shown in Figure 6).

It was intentionally left ambiguous whether this opening is more akin to a single eye with eyelids or a large mouth with opening jaws, in order to not connect a direct anthropomorphic interpretation to the design. Various cyclic movements were designed and tested for the opening to indicate affective state and liveliness—akin to breathing, blinking, or nervous ticks.

4.3.3 Detailed 3D Animation Studies After the general form was determined, the design process shifted to the next, more involved, step of 3D animation studies. A rough model of the robot was built in a 3D animation program, replacing detailed elements with geometric approximations of these parts. The overall layout and components of the robot were based on the freehand sketches.

The goal of this stage was twofold: to verify the expressive capabilities of the design and to decide on the exact placement, relationship, and range of each of the robot's degrees of freedoms. To that effect we iterated through a variety of robot models, each with their own set of DoF relationships and structural parameters.

Figure 7 shows still frames from some of the animation sketches. Each iterative robot model was "directed" through a number of different emotional and musical gestures: moving the head to a variety of beat genres, making and breaking eye-contact, surprise at an unexpected musical event, focus on the robot's own improvisation, approval and disapproval of musical events, and other gestures.

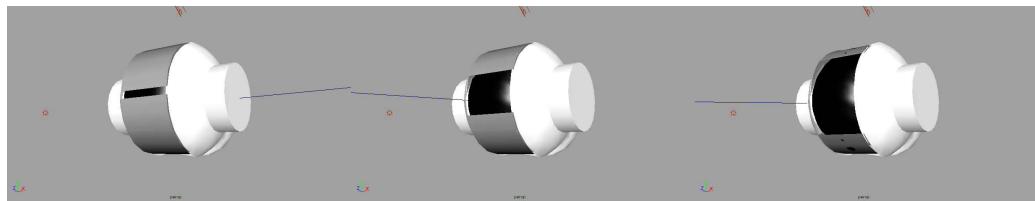


Figure 6. Still frames from animation tests of Shimon head opening mechanism, intended to steer clear from an explicit anthropomorphic face model while still enabling expressive "facial" expressions.

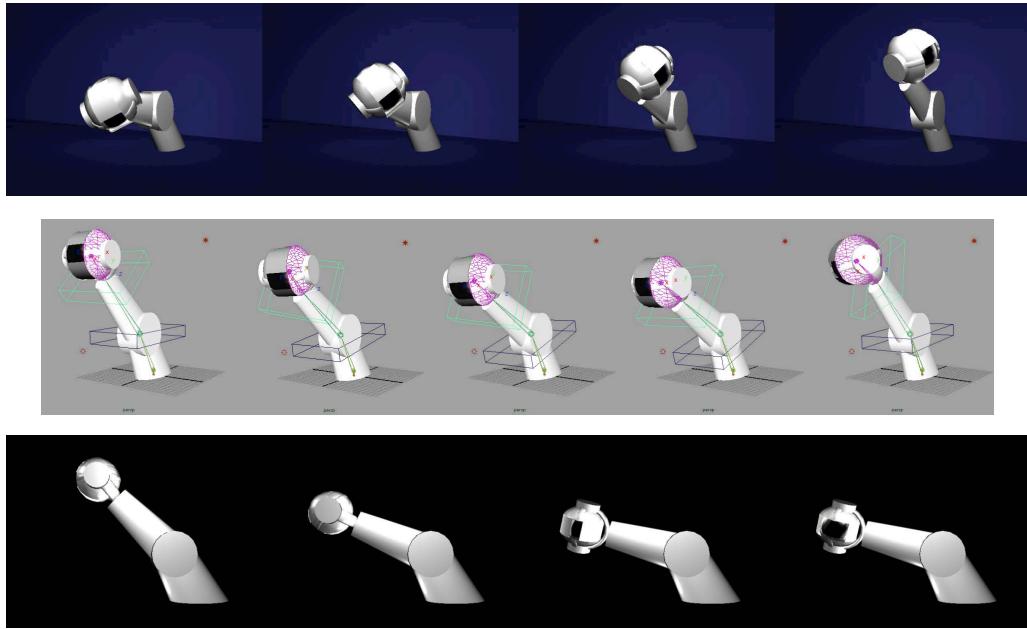


Figure 7. Still frames from detailed animation tests of Shimon head design, exploring DoF placement, hierarchy, and orientation.

The main points of design deliberation at this stage were the position of the base pan DoF (above or below the lowest tilt DoF), the height of the base tilt DoF, determining the breaking point in the neck, and the relationship between the head pan and tilt joints. For example, through these explorations, we decided that the “headphones” housing the pan motors would anchor the head sphere without moving themselves.

A notable outcome of this design stage was to use a non-orthogonal angle between the pan and the tilt motors, in combination with a seemingly right angle relationship between the joints reflected in the shell (Figure 8a–b). As the pan DoF rotates, the straight neck appears to “break” and create an illusion of a fully articulated 3-DoF joint. This is due to the fact that in the off-right-angle placement, the pan creates both a horizontal and a vertical movement, creating a “sideways-and-up” effect (Figure 8c). We noticed, moreover, that the precise choice of angle had a significant effect on the character expressed by the robot’s movement. After several angle studies, the design settled on a 40-degree offset between the two joints for a somewhat mischievous personality, befitting a jazz musician.

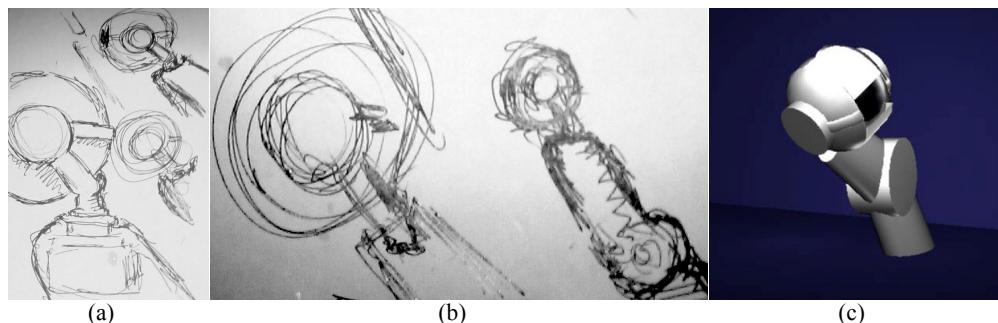


Figure 8. Deceptive placement of pan-tilt mechanism creates a surprising and organically moving joint simulating a fully articulated neck. (c) shows the resulting characteristic “sideways-and-up” gesture suggesting Shimon’s mischief.

4.3.4 Scale Studies Before committing to the final design, we conducted a number of scale studies to set the robot in its context of use and to evaluate the performative and communicative outcomes of the robot's size. This was done using an architectural design program set against models of humans and human environments.

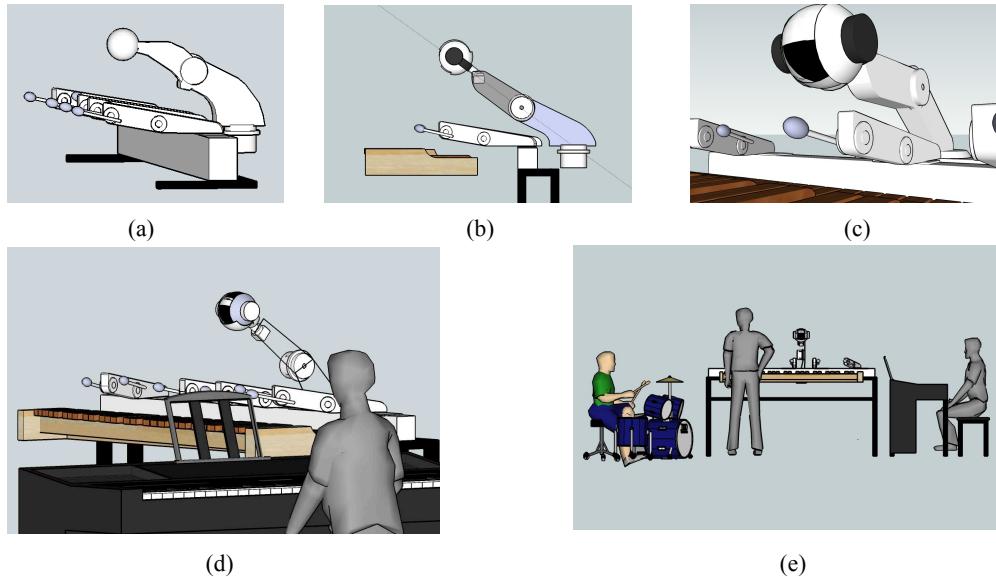


Figure 9. Shimon head design scale studies. Evaluating the robot scale with respect to the instrument and existing arms (a–c), band members (d), and the audience (e).

By rendering through a variety of camera positions, including at eye-level of humans in the robot's proximity, we tried to explore and test Shimon's head scale both as it related to the existing robot arm (Figure 9a–c) and with respect to the human band members (Figure 9d). Relative to the robot's arm, we wanted to strike an aesthetic balance. The head should be noticed, but not overshadowing. For the band members, the scale should support equality and ease of nonverbal communication. Finally, we were also concerned with the performative aspect of the robot's scale, and in particular, how the head would be visible by an audience when the robot is on stage (Figure 9e).

4.3.5 Solid Design and Robot Construction Given all of the above design stages, the final solid design created the structure and shell to support the design decisions described above and resolved issues of physical constraints and dynamic properties of the motors used.

Figure 10 shows the fully assembled robot. The final shell design closely follows the animation tests and strikes a balance between geometric shapes and expressivity. It includes four chained movement DoFs: a base pan moving the whole head left and right, a base tilt roughly 40 percent up from the base allowing for a bowing motion and the 40-degree coupled head pan-tilt mechanism. In addition, a servomotor controls both upper and lower shutter to support the opening and closing of the head. The head also contains a single high-definition digital video camera.

In practice, the robotic head is today used in a number of ways: the head bobs to signal the robot's internal beat, allowing human musicians to cue their playing to the robot's beat. The head makes and breaks approximate eye contact, based on fixed band member positions, to assist turn-taking. For example, when the robot takes the lead in an improvisation session, it will turn towards the instrument, and then it will turn back to the human musician to signal that it expects the

musician to play next. The head also tracks the currently playing arms, by employing a clustering algorithm in conjunction with a temporal decay of active or striking arms. And finally, two animation mechanisms—an occasional blinking of the shutter and a slow “breathing”-like behavior convey a continuous liveliness of the robot.



Figure 10. The socially expressive head of the robot “Shimon,” fully assembled and installed by the marimba-playing arms.

4.4 Summary

The expressive-movement based approach employed in the design of Shimon’s head started from very abstract shape animations and went through several iterations of movement, shape, and scale studies. Each movement design stage inspired the appearance appropriate for its phase: pencil sketches, 3D models, and architectural simulations. Our methodology led us to steer away from a humanoid head which would have included eyes, a mouth, and a neck, and instead end up with a shape roughly made of simple geometric shapes. These shapes, while simple, were meticulously placed in relation to each other in order to achieve the intended expressive movement. The specific interplay between shape and form also enabled us to achieve a complex set of movement from a minimal number of DoFs.

5. Case Study II: Robotic Speaker Dock and Listening Companion

The second case study is Travis, a robotic smartphone speaker dock and music listening companion. Travis is a musical entertainment robot computationally controlled by an Android smartphone, serving both as an amplified speaker dock and as a socially expressive robot. Travis is designed to enhance a human’s music listening experience by providing social presence and audience companionship, as well as by embodying the music played on the device as a performance. We developed Travis as a research platform to examine human-robot interaction as it relates to media consumption, robotic companionship, nonverbal behavior, timing, and physical presence.

There were a number of commercial speakers that embodied music through mechanical movement, mostly in the one- and two-DoF robotic toy realm, such as the Sega iDog. More sophisticated robots that provided sound amplification included the ZMP miuro (Aucouturier, Ogai, & Ikegami, 2008) and the SONY Rolly (Kim, Kwak, & Kim, 2009). However, these robots were designed as non-anthropomorphic mobile robots with only entertainment in mind. Travis extends on these designs in several ways: first, by being a smartphone controlled robot and docking station (Hoffman, 2012), then by being capable of gestures and not just mobility, and finally, by being designed as an open-ended research platform, extendible by mobile applications and cloud computing.

5.1 Motivation

In part, Travis was a result of our experience with Shimon, and in particular, with Shimon's socially expressive head as described in the previous section. In performances and public presentations of the Shimon system, audiences and critics have repeatedly commented on the prominent effect of the robot's social head and its expressive power. Indeed, it often seemed that audiences were more attuned to the social behavior of the robot as it related to the musical performance than to the algorithmic details of the musical improvisation itself. The robot was said to be "enjoying" or "experiencing" the music rather than just responding to it.

This suggested the notion of perceived robotic experience companionship (Hoffman & Vanunu, 2013)—the effects on humans of a robot responding to an external event with them. We wanted to design a robot that distilled only the musically responsive aspects of Shimon, and have it respond not only to music played by it but also to music in general. Thus, Travis was, as a first application, intended to play music from the smartphone, amplify it, and communicate gesturally with the music and with humans around the robot.

Initially, a scaled down version of Shimon's head was considered and partially designed, as can be seen in Figure 11. However, given the different behavior requirements and design parameters along with a general sense of a mismatch between the scale and the shape of a miniaturized Shimon head, this direction was abandoned and a new design process initiated.

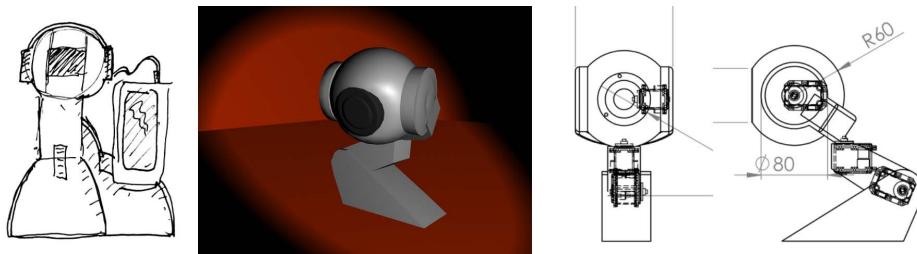


Figure 11. Initial design for Travis, as a scaled-down version of Shimon's expressive head.

5.2 Design Considerations

The robot's appearance was designed with a number of parameters in mind: first, the robot's main application is to deliver music and to move expressively to the music. Its morphology should therefore emphasize audio amplification, and support expressive movement to musical content: the speakers should feature prominently and explicitly in the robot's design; the robot's parts should be placed and shaped for musical gestures.

Second, the robot needs to be capable of basic nonverbal communicative behavior, such as turn-taking, attention, and affect display. The robot's head should be capable of several of these behaviors. Also, when placed on a desk, the robot's "face" should be roughly in line with a person's head when they are seated in front of it.

Finally, the robot's appearance should evoke social presence and empathy with the human user. We therefore wanted it to be sized and shaped to evoke a pet-like relation, with a size comparable to a small animal, and a generally organic but not humanoid form.

5.2.1 Relationship to Mobile Device When designing a smartphone-based robot, an inevitable design decision is the integration of the mobile device within the overall morphology of the robot. Past projects have opted to integrate the device as either the head or the face of the robot. *Mebot* uses the device to display a remote operator's face on a pan-tilt neck (Adalgeirsson & Breazeal, 2010). Other projects (Santos, 2012; Setapen, 2011;) have converted the mobile device's screen into an animated face inside the robot's head, an approach similar to that taken by the designers of the *Tofu* robot (Wistort & Breazeal, 2011).

In contrast, we have decided to not make the mobile device part of the robot's body but instead to create the appearance that the robot "holds" the device and is connected to it through a headphone cable running to its head. This is intended to create a sense of **identification** ("like-me") and empathy with the robot, as Travis relates to the device similarly to the way a human would: holding it and listening to the music through its headphone cable. Moreover, this setup allows for the device to serve as an object of common ground (Clark, 1996) and joint attention (Breazeal et al., 2004) between the human and the robot, setting the stage for **nonverbal dialog**. The robot can **turn the phone's front screen** towards its head and towards the human discussion partner (Figure 12a). In the final application, for example, we use a gaze gesture (Figure 17) as a nonverbal grounding acknowledgment that the device was correctly docked.



Figure 12. *Travis* sketches, showing concepts of (a) common ground and joint attention; (b) "holding" the phone and headphone cable, as well as musical gestures of head and foot. These sketches still show the initial Shimon-like design.

In addition, the decoupling between the robot's brain—its computational and sensory core—and its core structure suggested a theoretical point on any robot's relationship between physical embodiment, actuation, sensing, and computational location. New models of cloud robotics explore similar questions (Goldberg & Kehoe, 2013).

5.3 Design Process

The design process for *Travis* included four stages: (a) freehand appearance sketches; (b) DOF placement animation exploration; (c) abstract skeleton prototype; and (d) final solid design and construction.

5.3.1 Freehand Appearance Sketches Having abandoned the original Shimon-like design for the robot and given the decision that the phone should be placed in the robot's possession rather than being a part of its body, a series of freehand appearance sketches explored a wide range of form factors for the robot as well as for the relationship between the robot and the mobile device. Figure 13 shows a sample of these sketches. The aim of this step was to be free from economic and mechanical limitations, and, following a brainstorming paradigm, allow for surprising elements to emerge.

Some of the designs (for example Figure 13g–h) suggested a passive mode for the robot, in which the design closely resembles a traditional speaker or radio system. This mode would alternate with an active mode, in which the robot would "wake up" to transform into more of a creature-like shape (Figure 13i). Other designs (e.g., 13c, d, f) were further along on the creature-like scale.

Eventually, a notion of speaker prominence emerged, in which the speakers were the central part of the robot's appearance (Figures 13a, b, e–i). A non-anthropomorphic design made up of abstract shapes connected to a roughly creature-like figure was determined (e.g., Figure 13a, e). This was to strike a balance between the appliance and the companion nature of the robot. Moreover, by positioning the speakers in place of the eyes, the design evokes a connection between the input and output aspects of musical performance and enjoyment.

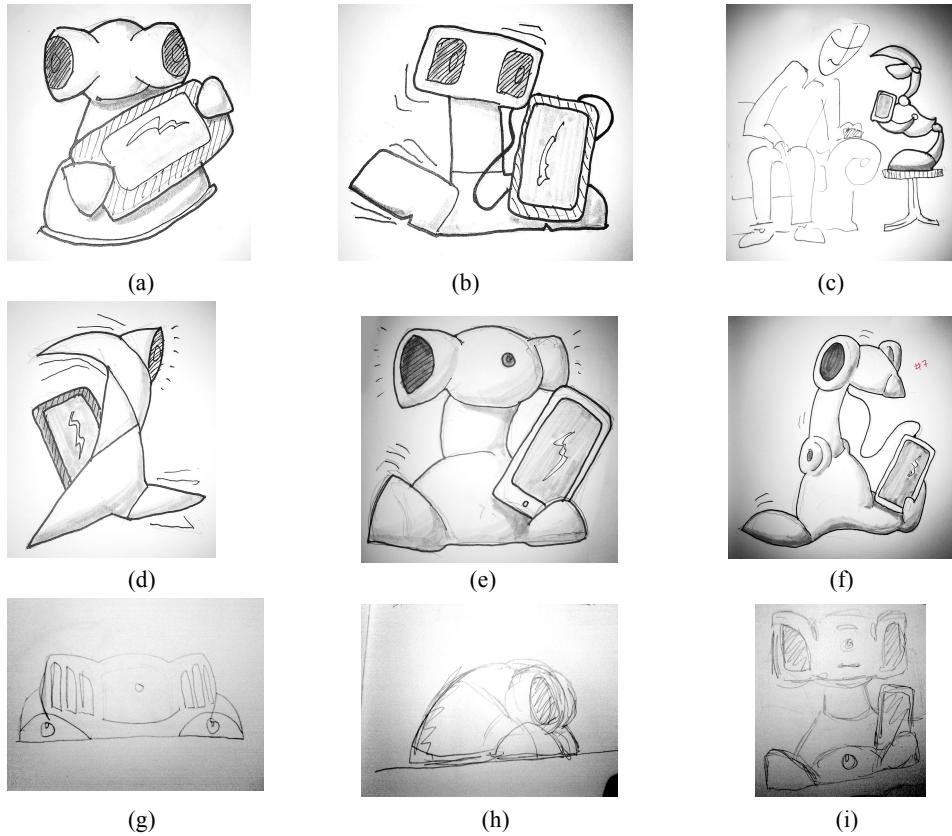


Figure 13. *Travis* sketches diverging from the Shimon-like design. Sketches g–i show the passive and active mode in the appliance-inspired design. Sketches c, d, and f show the more creature-like explorations of the robot’s form. Sketch e is the closest to the final form.

5.3.2 DoF Placement Animation Studies With the general appearance direction set, we continued to explore the scale, relationship, and DoF placement of the robot in a series of 3D animation studies. In terms of precision, these studies were in-between the first and third stages in the Shimon head design. The robot was represented by abstract shapes with no detailed features, but the shapes were intended to resemble the final design of the robot, as set by the freehand sketches.

Figure 14 shows stills from some of the animations used to explore the DoF relationships for *Travis*. Some of the major appearance parameters considered at this stage were the bulk of the base, the width-height proportions of the robot, the relationship between head and body, and the position of the most prominent DoF, the neck tilt. Each of the models was subjected to a number of animation tests, which included: waking up from the passive to the active state; responding to music using a variety of gestures; turning the robot’s attention to and from the human; relating to the robot’s mobile device; and several affect displays. The most effective design eventually chosen can be seen in Figure 14d.

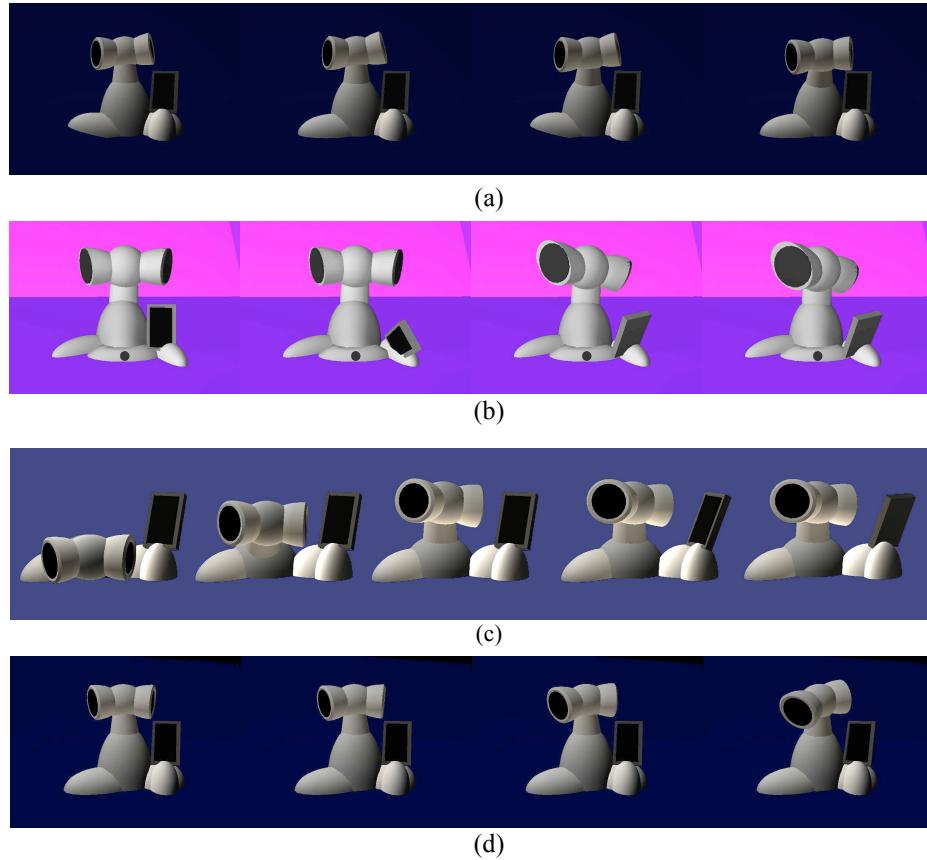


Figure 14. Animation stills from expressive movement design stage of *Travis* robot, demonstrating the “waking up” and “acknowledging phone” gestures, as well as the exploration of shape and scale relationships between the robot’s parts.

One notable feature to come out of this stage of the process is a slight asymmetry of the speaker horns with respect to the rotational DoF of the head tilt. This, in combination with the perfect spherical shape and central alignment of the head core with its driving DoF, leads to a deceptive effect: as the head moves up and down, the speaker horns seem to both rotate independently with relation to the head and also deform. This flexible, organic motion achieved with the use of a single direct-drive motor would have been difficult to imagine were it not for the animation tests leading to the final determination of the physical design.

5.3.3 Abstract Skeleton Prototype Before constructing the final robot, we used an additional rapid prototyping step to explore the robot’s expressive movement in physical space. Having set the positions and relationship of the DoFs in the previous step, we designed a wooden model of the robot outlining the core parts in the form of an abstract skeleton. The solid model and constructed prototype can be seen in Figure 15.

This step enabled us to experiment with the gestures and software system that were developed for the robot within a few days of the 3D animation studies. In fact, the solid design of the wooden model spanned a mere two days and the laser-cutting and construction took one additional day. As a result, most of the motor control software and expressive gesture system were developed and tested on the wooden skeleton long before the final solid models for the robot were completed.

5.3.4 Final Solid Design and Construction The final solid design stage combined the insights in terms of DoF number, placement, and orientation tested on the skeleton prototype and the shapes explored in the animation stage, resolving issues of physical constraints and dynamic properties of the motors used.

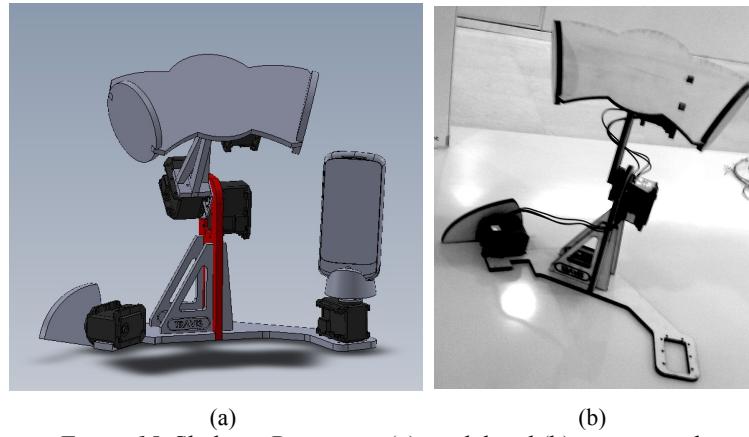


Figure 15. Skeleton Prototype, (a) model and (b) constructed

In addition, detailed proportions and relationships of the shell were explored (see Figure 16). These were subtle changes in parameters such as scale relations, precise drop off angles of shell components, proportions and bulk, and motion range, as well as bevels and accents. Performing this step in a parametric CAD software application lent itself better to making precise changes in the design, and, of course, to taking into consideration physical limits of motors, cables, and other electronics embedded in the robot. The completed robot can be seen in Figure 17.

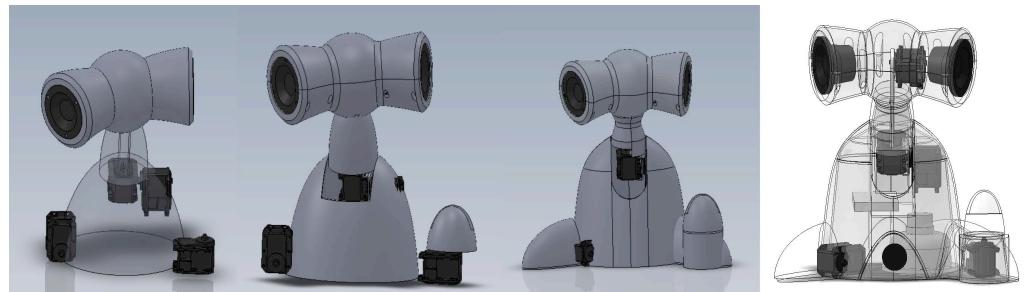


Figure 16. Travis solid designs and finishing details exploration

Using the 3D animation models as guides for the solid design has the beneficial side-effect of a propensity for clean geometric shapes. These shapes suggest abstract relationships, connected by the shapes' movement, and conceal the mechanical structure of the robot. This, in turn, supports expressive movement-centric design by having less distracting features, such as cables and motors, and allowing the attention to be primarily on the robot's movement.

5.4 Summary

In this case study, using expressive movement design techniques led to several design outcomes. First, abstract shape 3D animations encouraged us to focus on the simpler shapes in our freehand sketches, as we saw that much of the expressive power can be achieved with a combination of

spheres, half-spheres, and cut-off cones. Second, the Travis animation studies led us to discover the pseudo-deformation achieved with a single direct-drive motor in combination with the particular shape of the speaker cones. Finally, the use of an abstract skeleton prototype helped us design the animation-inspired software—described in more detail in Hoffman (2012)—before having built the complete robot.



Figure 17. Completed prototype of Travis, a robotic speaker dock and listening companion.

6. Case Study III: Desktop Telepresence Robot

The next case study is the gesture design for our study of a desktop telepresence robot (Sirkin & Ju, 2012). In this case, our goal was not to generate ideas for a new robot that was being designed from scratch, but instead, to explore how robotic gestures could augment or detract from already popular desktop telepresence programs. We sought to understand the role that physical interaction might play in human-robot or human-machine interactions. This case illustrates how contextual aspects of movement and gesture can be studied prior to the design of the full robot.

6.1 Motivation

The telepresence robot study was motivated by the observation that a new class of robots, which we called embodied proxies (Sirkin & Ju, 2012), were coming into vogue. Embodied proxies combine a live video representation of a remote worker with a local physical platform, usually with human-body-like proportions. Following in the footsteps of research systems such as PRoP (Paulos & Cannan, 1998), Porta-Person (Yankelovich, Simpson, Kaplan, & Provino, 2007), RoCo (Breazeal, Wang, & Picard, 2007), and Embodied Social Proxies (Venolia et al., 2010), commercial systems such as Suitable Technologies' Beam, the VGo robot, the Anybots QB and the Double Robotics robot were being developed and deployed. These commercial robots are similar in that they host live video of the remote worker on a flat screen mounted on a remotely steerable base. Smaller remote proxies like Revolve Robotics' Kubi and the Botiful were used on the desktop, similar to the MeBot research platform (Adalgeirsson & Breazeal, 2010). Embodied proxy systems are generally intended for workplace settings, enabling remote workers engage in day-to-day, informal interactions with their centrally located peers.

One issue unique to these embodied proxy systems is the proxy-in-proxy problem where the motion of the remote worker, shown on the video display, can create strange juxtapositions with the articulated motions of the local physical platform. These could portray inconsistent nonverbal facial and gestural cues, which we know from research in face-to-face interactions to cause mistrust (Kraut, 1978) and increased cognitive load (Fiske & Taylor, 1991). We decided to study how inconsistencies between on-screen and in-space movements might affect people's interpretations of the remote worker's nonverbal gestures and their perceptions of the remote worker. In order to

get a broad sampling of how people react to proxy-in-proxy designs, we ran our experiments using online video prototypes and crowd sourced participants from Amazon’s Mechanical Turk service.

This case study features a mix of mechanically puppeteered Wizard of Oz movements, and later, remotely controlled robotic gestures that were captured in video prototypes. Using video prototypes over in-person trials for movement-based design allowed us to more carefully control the coordination of the on-screen and in-person movement, to make sure that different study participants are seeing the actions from the same viewpoint, and to recruit a more geographically diverse audience.

6.2 Design Process

The design process of the desktop telepresence robot included four stages: (a) device prototyping, (b) Wizard of Oz studies, (c) video prototyping, and (d) scenario studies.

6.2.1 Device prototyping In the first stage, we developed a physical simulation of a desktop telepresence robot. Whereas in the previous case studies the forms and movements were designed from scratch using basic shapes, in this case study the robot was prototyped by appropriating an existing device. We originally sketched and enacted how the design of a desktop screen on a “neck” should move and saw that an existing product, the iMac G4, had many of the DoFs we desired.

With its hemispherical base and its three DoF “neck” that supported at 15 inch screen, we were able to quickly prototype a desktop telepresence robot that had a movable screen and a live telepresence feed. The neck allowed the screen to pan in the horizontal plane at the hemispherical base, tilt in the vertical plane and the base, and tip the screen in the vertical plate at the top of the neck-screen connection. To move the screen around, we used two four-bar linkages of dowels, covering the dowels in black tape so that they would blend in against a black background.

While the head was specifically sourced because it had the degrees of freedom we needed to emulate the on-screen motions of telepresence users, the rest of the device was prototyped using found objects. We built the torso out of an IKEA Eiworth stool and used an entry-level OWI robotic arm to enable in-space pointing and gesturing.

6.2.2 Wizard of Oz studies To better understand the possibilities of the robotic telepresence setup, we used our device in research lab meetings to “host” remote research team members who regularly teleconference into discussions every week. During a few meetings, we placed our prototype device on a chair and displayed the live web-camera view of the remote team member full-screen on the screen of the telepresence robot. A member of our design team stood behind the telepresence robot prototype and puppeteered the head and neck of the robot. A separate monitor with the on-screen feed of the remote team member allowed the puppeteer to try to mimic or even exaggerate the on-screen gestures and motion of the remote team member.

By interviewing both the local research group meeting participants and the remote team member, we were able to get informal assessments and observations about how the telepresence robot worked. On the whole, the interactants were positive and said they felt “strangely” like the remote member was more present. The remote team member also reported feeling like more people were watching him and that he was obliged to focus and not take on the other parallel tasks that usually felt natural when he was participating via laptop. Viewers of the robot often giggled at the mirroring of the on-screen and in-space action, not because it was funny but because it viscerally felt right. On the other hand, there was some enjoyment of discrepant moments in the meeting, both when the puppeteer had some fun gesturing the robot in exaggerated ways that the remote participant was not able to see and when major mismatches in action occurred because the person puppeteering the robot incorrectly anticipated what the remote person would do.

6.2.3 Video prototyping After some period of informal experimentation, we developed some intuition for how and when the telepresence robot should move. One aspect, which we thought

would be interesting to test in a controlled experiment, was whether the on-screen and in-space actions were reinforcing; in playing around, it was clear that viewers felt a really visceral feeling of “rightness” when the on-screen and in-space motions mirrored each other in a well-coordinated way.

To evaluate these questions, we recorded a series of short video snippets using gestures and motions that were the most common in our Wizard of Oz experiments and varied them so that they showed on-screen only, in-space only, or both on-screen and in-space.

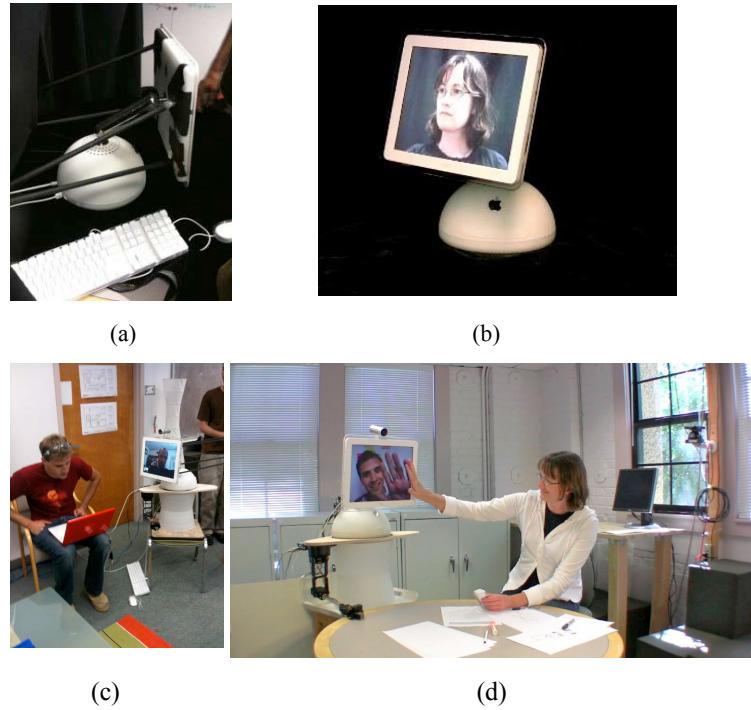


Figure 18. The physically simulated telepresence robot and Wizard of Oz set-up (a) allowed us both to test variations of specific gestures (b) and to experiment with real meetings how people would respond to a telepresence robot (c–d).

Each clip presented one of the following behaviors, which we had observed during typical remote video conference meetings:

- Agree (Nod “Yes”)
- Disagree (Shake “No”)
- Surprise
- Laughter
- Look to One Side
- Lean In to Look Closely
- Look Down at the Table
- Confusion
- Think Carefully

We hypothesized that consistency between on-screen and in-space action would improve observers’ comprehension of the message that the remote participant is expressing when compared to on-screen or in-space action alone.

We deployed these clips in controlled between-subject online experiments and indeed found that consistency between the remote actor’s facial expressions and gestures and the proxy’s physical motions resulted in improved understanding of the behavior portrayed, as well as in higher

confidence levels and stronger responses. There was, however, an interesting variance based on the kind of gesture portrayed, as we reported in detail in Sirkin and Ju (2012).

6.2.4 Scenario studies One important aspect of video prototyping is that it is possible to set the context that the movement is used in. For the next stage of our telepresence robot work, we performed studies looking at how having on-screen and in-space motion changed the relational dynamics, such as the relative amount of power and perceived involvement of the remote participant. The video scenario shows a third-person view of a remote teammate (Eric) asking an on-site collaborator (Becky) for assistance revising the design of a hand-held remote control (see Figure 18d for the physical setup). A brief discussion ensues about how to make the remote work for a wider range of hand sizes, and another local participant is called over for further design support. After a brief period, the three check back in with each other, review the designs they had developed, and choose one that resolves the original problem. The clip is 90 seconds long and includes an audio track of the actors' conversation.

In order to show the robot working realistically in this scenario without a puppeteer in-frame, we actually needed to build robotic drive mechanisms for the neck assembly and experimented with recreating our gesture designs using electronic control. The iMac G4's screen was actuated by three DC motors and a cable drive system to move the neck and screen to positions controlled using a remote interface. Screen motions were controlled gesturally through the orientation of a hand-held Wii remote, so that larger movements of the remote produced more rapid movements of the screen. Pilot trials also revealed the need for arm-based gestures, so we added a Lynxmotion AL5D five DoF robotic arm to provide deictic as well as other symbolic gestures critical to interactive team activities.

As in the previous step, we made several videos of similar scenarios, varying only whether the remote participant had on-screen only or on-screen plus in-space motion and also whether the local or remote participant had the more dominant role in the meeting. We found that the addition of physical proxy motion favorably influenced the perceived dominance and involvement of the interaction. Proxy motion also had a surprising influence on the perceptions of the on-site teammate. When the remote participant displayed proxy motion, the on-site teammate was viewed as being more equal in stature.

6.3 Summary

Our video prototype studies showed that when the remote actor's movements were reinforced by the movements of the local physical platform, viewers had an improved understanding of the expressive movements as well as greater confidence and strength of response. For iconic gestures such as nodding "yes" or shaking the head "no", the gestures were well-recognized regardless of whether the on-screen and in-space movements matched up. However, for semi-voluntary reactions showing laughter or surprise, the addition of on-screen reinforcement of in-space action was critical to help viewers interpret the actions; in-space motion alone was confusing. For movements that indicate orientation of attention, such as looking to one side or leaning in, the on-screen motion alone could be confusing; in-space motion and in-space with on-screen actions were better understood. The scenario-based studies showed that these differences in interpretability have impact in a realistic context. When the remote participant was shown physically gesturing in the scenarios, he was perceived as having greater involvement. On-site teammates were also perceived as being more equal in stature.

An important concern regarding the use of video prototypes in HRI is centered on whether the use of video prototypes is an adequate substitute for live first person interactions, both for conducting human-robot interaction research and for evaluating design concepts. However, Woods et al. (2006) were able to demonstrate in a series of studies that the results of studies with video prototypes are comparable to those with real interactions. These comport with research by Ju and Takayama (2009) that indicate that reactions to video prototyped interactions (in their case of gesturing interactive

doors) have results that echo the reactions to physical interactions but with a smaller effect size.

By keeping movement in mind during the design of these telepresence robots, we are able to positively influence the non-verbal communication of the participants and also to improve the perceptions observers have of the people using the robots.

7. Case Study IV: Animated Personal Robot

The final case study looked at how performative movements associated with anticipation and reaction affected people's ability to read the robot's intentions and actions. Although the details of the study and its findings are well-documented in Takayama et al. (2011), we aim here to discuss the design details that went into the creation of the gestures and movements used in that study.

In this case, the robot portrayed in the animations was the Willow Garage PR2, which is an existing and working robot platform; however, many of the gestures we depicted were faster than the PR2 could actually execute or involved more degrees of freedom. More importantly, the animations in the study all featured the robot's movements in the context of scenarios where the robot was attempting to perform some task—plugging itself in, serving drinks to restaurant patrons, opening a door, and ushering people. This case illustrates how contextual aspects of movement and gesture can be studied prior to the design of the full robot.

7.1 Motivation

The animation studies were motivated by issues that researchers at Willow Garage have faced sharing research space with the PR2 robot they were developing (Bohren et al., 2011). Unlike Shimon and Travis, the PR2 robot was designed with personal robotics applications in mind. The PR2 has the capabilities to autonomously plan motions, reason, navigate, open doors, and grasp objects, enabling it to, for example, fetch a drink from the refrigerator.

Because the focus of the robot development centered on autonomy and safety, the expressive aspects of the robot's motions were not optimal. At times, the lack of expressivity would impact the robot's performance. For example, people sharing the space with the robot would not understand that it was scanning the refrigerator door to identify the door handle as part of the grasping task; it seemed as if the robot was just accidentally parked inconveniently in front of the fridge. Even people who sought to be helpful sometimes inadvertently created unexpected disturbances, causing the nominal plan to be interrupted and reformulated. Other times, researchers assumed the robot was calculating a plan of action, when in fact it was stalled. The lack of performed thought also created safety issues; even roboticists occasionally mistook cogitating robots for idle robots and have narrowly missed being hit by darting arms. While some of these issues could be improved with speedier performance, it became clear that the PR2 could actually help to prevent a lot of these issues by being clearer about what it was doing when it was sensing and planning.

Because the PR2 was already built and lacked a lot of the modalities used by people to express thought—e.g., tilting the head or scratching the temple—we focused on using animated video prototypes to increase the expressive range of PR2 beyond what the actual robot could do. In this way, we would be able to gain insight on what movements were useful and what features would need to be incorporated in future generations of personal robots.

7.2 Design Process

The design process of the personal robot animations included four stages: (a) scenario analysis, (b) scenario selection, (c) animations, and (d) studies.

This work was done in collaboration with accomplished animator Doug Dooley from Pixar Animation Studios. As in the previous case study, we were using the animations as short video clips that we could then use in online studies, allowing us to understand how the robot's expressions and gestures were understood by a wider audience. However, we were using our understanding based in interaction shortcomings with the PR2 as expert inspiration rather than

physically prototyping a life-size robot. Meerbeek, Saerbeck, and Bartneck (2009) have described an approach for designing robot personality characteristics, which is similar to our design process. In their work, professional improvisational theater actors generate interaction scenarios that are then video-recorded, carefully analyzed, and refined into 3D animations.

7.2.1 Scenario analysis As an animator, Dooley had a lot of intuition and expertise on how to make robots seem "real"; however, the robots we were depicting were actually real, but did not move in ways that felt right. Thus, many of our early sessions involved analyzing the activities that the real PR2 robot did, while thinking through what would feel more "right" and why. Whereas Dooley had many ideas of specific movements and gestures that were appropriate for each situation, it took a lot of animated discussion and enactment to generalize some of these specific moves—the forward lean of the body towards interactants, the tilting and scratching of the head, the lifting and slumping of the torso—to the broader themes of forethought and reaction, and to distinguish them from gestures and motions that indicate mood or emotion.

One of the most interesting aspects of the analysis discussions was the switch between high-level concepts, like animation principles of anticipation and reaction, and detailed enactments of specific movements or behaviors that people or animals employ in different contexts. We asked many questions like: do we use the similar or different movements when we think about how to use a tangible thing, like a door handle or power plug, as when we are wondering about how to spell something? These questions usually prompted attempts at acting out different scenarios, and suggestions from the other members of the discussions how to tweak those motions so that they were clearer.

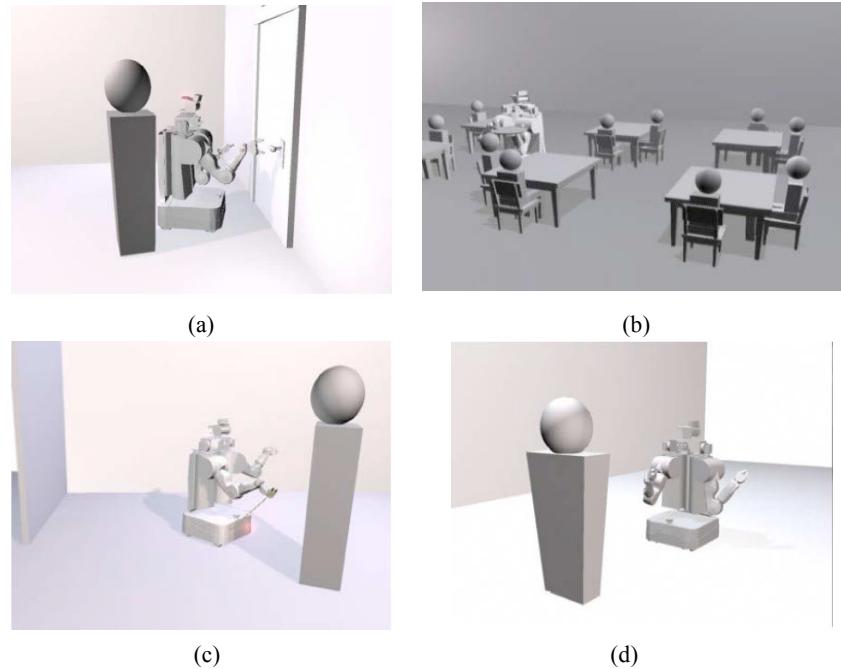


Figure 19. The use of forethought and reaction was studied in scenarios with the personal robot attempting to (a) open a door, (b) bring restaurant patrons drinks, (c) plug itself in, and (d) usher people.

7.2.2 Scenario selection Although the scenarios selected stemmed from specific tasks the PR2 actually did, our study's goal was quantifying the effects of adding movements inspired by the animation principles of anticipation and reaction for future robots. Thus, though we made

sure the robot's actions were technically possible, we did not constrain the robot in the animation to only performing movements that the PR2 could do. We selected four different scenarios in order to have a mix between functional tasks (such as opening doors or offering drinks) and communicative tasks (waving for help, or ushering). By selecting a mix of tasks representative of those being attempted by others in the robotics community, we hoped that people would be able to focus on the principles and motions rather than the specifics of any particular task.

7.2.3 Animation One of the interesting discoveries in this study was how important it was to animate the scenarios we were depicting as simulations, and not narratives. With too much detail study participants tend to shift into the mindset that they were watching a cartoon rather than evaluating a prototype. Hence, it was actually necessary to keep the animations pretty spare and "sketchy," with simple shapes and minimal use of color. Even though the animations were made to look primitive, the motions depicted in each scenario were actually quite complete. In the scenario where the robot is looking for a plug, for instance, the robot seems to notice that his power-low LED is blinking, looks around and finds his plug, pulls it out, and stares at it while deciding what to do. He then looks around for a plug, does not seem to spot one, and then looks over his shoulder, turns his body, contemplates the plug again, looks around and sees a person, lifts his body in anticipation, and moves towards the person, waving the plug in his hand. All of this happens within 15 seconds.

We found in pilot testing that people were strongly influenced by the outcome depicted in the scenario, so we separated the animations into pre-action and post-action parts. In the post-action animation, participants saw the task outcome (success or failure) and reaction or lack of reaction to the task outcome.

7.2.4 Studies Our studies were deployed online using Amazon's Mechanical Turk; participants were shown multiple scenarios, but depending on their experiment condition, saw different versions of the animations. Unlike in the earlier case studies where the intended outcomes were robots with well-designed movement, in this case study the goal was to establish principles and practices that would lead to well-designed movement.

In our studies, we found that animating anticipation into a robot's action can improve how sure people felt about their readings of the robot's intentions, as well as improve people's perceptions of the robot's appeal and approachability. Adding a reaction gesture to a robot's action positively affected the perceived "smartness" (a combination of "intelligence" and "competence") and confidence of the robot.

The animations in the study are not a template for exemplary robot movements; the studies do not obviate the need for experts like Dooley who have excellent intuitions for the movements and mannerisms that indicate what a robot is thinking of doing. However, studies help to show why these intuitions, movements, and mannerisms are important and what impact good design can have.

7.3 Summary

The video prototype studies illustrated that performing forethought greatly improved people's confidence in diagnosing what it was that a robot was trying to do and also increased the appeal and approachability of the robot. Performing reactions to the success or failure of the robot task make the robot seem smarter and more confident even in the scenarios when the robot failed on the task. By designing this theoretical PR2 robot, we were able to design with expressive movement in mind and thereby take into account how that movement would be interpreted in-context and what effect expressive movement would have on the robot's ability to perform the task with other people around.

In an ideal scenario, we would follow up the animation studies described here by deploying these same forethought and reaction routines on the physical PR2, testing these capabilities in physical and real scenarios with other people. However, as we mentioned, the modalities we

needed to express thought were not built into the original PR2 and retrofitting these robots with these features was prohibitively expensive. Originally we planned to incorporate some more of these expressive capabilities in future generations of the Willow Garage robot; however, plans for the follow-on robot to the PR2 were put on hold and subsequently cancelled.

8. Discussion

The techniques and case studies presented in this paper lay the groundwork for an expressive movement centered approach towards robot design. By prioritizing the quality and communicative properties of movement over pragmatic function or aesthetic form, we can give proper attention to the importance of movement in human communication and perception.

The focus on expressive movement encourages designers interested in the robot's ability to communicate its internal states to de-emphasize static aspects of the robot's look and to use movement to dynamically create effects of personality, state, affect, intention, and mood.

In this discussion, we relate our framework to the design of expressive non-anthropomorphic robots and suggest how this could benefit personal robotics and HRI research by availing real-world human-robot interaction. We then proceed with categorizing the challenges of designing robots with movement in mind.

8.1 Designing Non-Anthropomorphic Robots

One of the commonalities across the case studies presented is that most of the robots did not fall into classic humanoid form. We believe this is not an accident, for two reasons: (a) well-designed movement, as key to creating communicative ground, is all the more critical when there are no anthropomorphic features to reference and (b) the focus on quality of movement buys the designer valuable expressive power of which other aspects of the robot's design can be relieved.

Movement-centric robot design may thus open the door to more simple non-anthropomorphic robots, such as abstract volume robots or robotic furniture, that engage with humans primarily through their movement. This can have interaction benefits, such as lowering expectations stemming from anthropomorphic appearance (Duffy, 2003; Nomura et al., 2008) or avoiding the uncanny valley (Mori, 1970).

From a design point of view, we identify additional benefits of building non-anthropomorphic robots for HRI, including freedom of exploration, economic feasibility, and the potential for higher acceptance:

8.1.1 Freedom of exploration By definition, naturalistic and humanoid robots imitate an existing ideal, be it human or animal, and therefore constrain the design exploration towards the natural example. Most of the design process thus lies in deciding which features to copy and how to imitate the ones that are hard to replicate. The resulting design is then evaluated with respect to the original and inherently falls short, always being a lacking simulacrum of the ideal. In contrast, starting the design process without a set ideal guiding the robot design, the object can be imagined through a variety of evolutions and paths. A common object, such as a door, a speaker dock, or a lamp can supersede the original in any number of ways (see, e.g., the mechatronic home appliances of Chambers, 2011). This open-endedness allows both for more creative freedom in the robot design process and for this process to eventually be valued on its own terms.

8.1.2 Economic feasibility and rapid prototyping Second, considering simpler forms with fewer degrees of freedom and less detailed features holds the promise to study human-robot interaction that is relevant to more imminent consumer products and affordable prototypes. This can hasten the road towards HRI outside of the research laboratory, making it more relevant to real-world applications. Simple economic forms also allow for easier replication and thus avail large-scale deployment for field studies and other multi-participant research paths. Finally, economic

form allows for rapid evolutionary prototyping, frequent reimplementation, and hands-on exploratory design. All of the above are out of reach for complex, one-of-a-kind, anthropomorphic robots.

8.1.3 Acceptance Finally, taking another note from the world of character animation, non-realistic characters tend to be accepted even when they lack believable detail (see, e.g., Bates, 1994). Photo-realistic characters are judged on nature's terms, whereas—without a real reference point—reality-detached characters are accepted for what they are, be it a disappointed desk lamp or a talking carpet. Perhaps, the less detailed a character design is, the more space is left for the viewer's imagination, enabling them to project a more forgiving narrative onto the object.

8.2 Design Challenges

Based on our shared experiences with designing robots that emphasize expressive movement, we have identified four key challenges for movement-focused robot design: discovery, implementation, appearance-matching, and validation.

8.2.1 Discovering the right movement The core design challenge of the proposed approach is to explore and discover the right movements for the robot. The designer needs to envision the use of the robot, its personality, the internal states, and other information it needs to communicate as well as how it fits into the human environment. These questions should guide the movement the robot needs to achieve. In our case studies, we show how using simple prototypes (e.g., the desktop telepresence robot and Travis), simulations (e.g., the PR2), enactments (e.g., the telepresence robot), and engagements with experts such as animators (e.g., Shimon, Travis, and the PR2) can aid this process.

The process of finding the right movement can be more of an art than a science and hence it is useful to engage collaborators who can offer specific insights or intuitions on how the robots should move. These can be experts and professionals in movement-related fields, such as actors, dancers, choreographers, or animators, and they can also be lay people or end users. We often find that design processes allowing greater participatory design, such as Wizard of Oz or video prototyping, are useful in designing expressive robot motion.

8.2.2 Implementing the movement Once the movement qualities of the robot are defined, a mechanical system needs to be designed to support the movement requirements. The designer needs to consider actuator performance and set the resulting type and size of actuators. Mechanical and economical considerations further constrain the placement of actuators and resulting degrees of freedom (see, e.g., Sections 4.3.5 and 5.3.4).

Since this stage is driven by the expressive, and not the functional, movement of the robot, unusual solutions can emerge, which diverge from classical mechanisms of industrial robotics. In our work, we found that strategically placed DoFs can support a certain attitude. For example, Travis has a doubly-linked tilt DoF that, when used in synchrony (a common animation technique), affords a smooth organic movement in his beat tracking gesture. Similarly, non-orthogonal placement can provide a surprising effect, such as Shimon's mischievous “up-and-back” rotation. In addition, a combination of active and passive joints can be used to simulate secondary action.

8.2.3 Matching form to movement Once a set of movements is defined and a mechanical plan is outlined, the robot's overall form and detailed appearance should support these movements. At this stage, the designer needs to relate the degrees of freedom to the intended shape of the robot and experiment with the resulting movement capabilities and expressive results. Just as the mechanical design was made to support the intended movement qualities, so should the appearance design support the same goal. For example, Shimon's headphone-like motor housings serve both as mechanical pivots and as a simulated accessory for a robotic musician. The spherical head shape of both Shimon and Travis enable apparent movement across the surface of the robot. Travis's head

shape was designed to work together with the concentric DoF placement to cause the rigidly attached speakers to appear undergoing a deformation.

Since the challenge of designing and implementing movement is intertwined with the challenge of creating the robot's form, the experimentation is inherently iterative. Tools that enable the rapid creation of form and testing of motion, such as the DoF configuration tool presented in Section 3.5, skeletal mock-ups, rapid prototyping techniques, and mechanical "stand-ins" can help designers quickly find form and movements that work together well.

8.2.4 Validating the design Finally, the designer needs to validate that the design has the intended effect. To this end, the means of validation should be tied to the robot's purpose. If the robot is meant to be performative, then it is desirable to have the motions and forms shown to proto-audiences, so that the context of the motion is factored in throughout. Alternatively, if the robot is meant to be interactive, it may be necessary to experiment with having different people view and interact with the robot to see if the motions feel right and read correctly from both a first- and third-hand perspective. Evaluation of the designed movement can and should happen at a number of points in the process. If 3D animations are used in the design process, they can be evaluated for their readability and emotional effect before moving forward to more detailed implementation, as in the case of the PR2 and Shimon's head. Simple skeleton prototypes can be built, either actuated or passive, and used as mockups to evaluate the usefulness of the structure in terms of the movement design, as shown with Travis and the desktop telepresence robot. Then, of course, the built robot can be used in studies to see how it measures up to the intended design goals.

9. Conclusion

In this paper, we argued for a design process for interactive robots that puts expressive movement at its center. Movement is a highly salient, yet widely under-recognized, aspect of human-robot interaction. Movement can give interactive robots a larger dynamic range of expression and enables spatiotemporal affordances. Because humans are extremely sensitive to subtleties of motion—a fact reflected in the importance of kinesic facets of nonverbal communication, in the appeal of abstract character animation, and in our sensitivity to understand motion from minimal cues—well-designed robot motion can go a long way towards people's comprehension and acceptance of robots.

In presenting design techniques and design challenges, we aim to give designers a feel for the range of tools and issues that come into play when we are designing robots with the quality of their movement in mind. We use 3D character animation sketching, Wizard of Oz and video prototyping, interactive DoF placement tools, and fully actuated skeleton models for motion control exploration. Our case studies represent a wide, though not comprehensive, range of expressive-movement centric robot designs and goals. In the design of the social head for Shimon, a robotic marimba player, and of Travis, a robotic speaker dock listening companion, we designed the robot's movement first and followed through every other stage of the robot's design to support the movements as indicated earlier. In our telepresence robot experiments, we designed a variety of gestures and experimented with how the interplay between the on-screen and in-space gestures affected how people interpreted various movements. Finally, in our animated robot studies, we looked at how different movements in different contexts, informed by higher-level goals of exhibiting forethought or reaction to task success, could affect how easily people could understand what the robot was intending to do.

A focus on movement design for interactive robots has the potential to usher in a new era in human-robot relationships. Robot designers can make do with feasible, simple, low-DoF machines, that will be able to communicate intentions, display internal states, and evoke emotions through the way they move. Robotic furniture, abstract companions, and actuated everyday objects can supplement the more commonplace vision of robotic butlers and humanoid assistants. These simple, recognizable objects, moving in an engaging way, might be more easily accepted into people's day-to-day lives when compared to complex anthropomorphic machines, whose crude and unaffectionate motion mainly cues their human counterparts to stay their distance.

Acknowledgements

Both Shimon and Travis were designed in collaboration with Roberto Aimi of Alium Labs as well as in collaboration with the Georgia Tech Center for Music Technology (GTCMT). Shimon was designed and constructed at the GTCMT under the direction of Prof. Gil Weinberg; Travis was designed as part of a collaboration with the GTCMT and Prof. Gil Weinberg. We would also like to thank Ian Campbell for his help designing the skeleton model of Travis.

The telepresence robot study was designed in collaboration with David Sirkin of Stanford's Center for Design Research with a generous research grant from the Hasso Plattner Design Thinking Research Fund. The animated PR2 research was performed in collaboration with Leila Takayama from Willow Garage and Doug Dooley from Pixar Animation Studios.

References

- Adalgeirsson, S. O., & Breazeal, C. (2010). MeBot: A robotic platform for socially embodied telepresence. In *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 15–22), Osaka, Japan.
- Akers, D. (2006). Wizard of Oz for participatory design: Inventing a gestural interface for 3D selection of neural pathway estimates. In *Proceedings of CHI '06: Extended Abstracts on Human Factors in Computing Systems* (pp. 454–459), Montreal, Canada.
- Alami, R., Clodic, A., Montreuil, V., Sisbot, E. A., & Chatila, R. (2005). Task planning for human-robot interaction. In *Proceedings of the 2005 joint conference on smart objects and ambient intelligence (SOC-EUSAI)* (pp. 81–85). New York, NY: ACM Press.
- Argyle, M. (1988). *Bodily Communication* (2nd ed.). Methuen & Co, UK.
- Aucouturier, J., Ogai, Y., & Ikegami, T. (2008). Making a robot dance to music using chaotic itinerancy in a network of Fitzhugh-Nagumo neurons. *Neural Information Processing*, 4985, pp. 647–656.
- Baldwin, D. A., & Baird, J. A. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*, 5(4), 171–178.
- Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind* (pp. 233–250). Oxford, UK: Blackwell Press.
- Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior*, 26(4), 313–331.
- Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, 37(7), 122–125.
- Blythe, P. W., Todd, P. M., & Miller, G. F. (1999). How motion reveals intention: Categorizing social interactions. In Gigerenzer, G., and Todd, P. M. (eds), *Simple Heuristics that Make Us Smart*. Oxford University Press, New York.
- Bohren, J., Rusu, R. B., Jones, E. G., Marder-Eppstein, E., Pantofaru, C., Wise, M., ..., Holzer, S. (2011). Towards autonomous robotic butlers: Lessons learned with the PR2. In *Proceedings from the IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5568–5575).
- Breazeal, C., Brooks, A., Chilongo, D., Gray, J., Hoffman, G., Kidd, C., ... Lockerd, A. (2004). Working collaboratively with Humanoid Robots. In *Proceedings of the IEEE RAS/RSJ International Conference on Humanoid Robots (Humanoids)*. Santa Monica, CA.
- Breazeal, C., Wang, A., & Picard, R. (2007). Experiments with a robotic computer: Body, affect and cognition interactions. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction* (pp. 153–160). New York, NY: ACM.
- Bretan, M., Cicconet, M., Nikolaidis, R., & Weinberg, G. (2012). Developing and Composing for a Robotic Musician Using Different Modes of Interaction. In *Proceedings of the International Computer Music Conference (ICMC)* (pp. 498–503), Ljubljana, Slovenia.
- Chambers, J. (2011). *Artificial Defense Mechanisms* P. Antonelli (Ed.). New York, NY: The Museum of Modern Art.
- Clark, H. H. (1996). *Using Language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, 7(4-5), 507–525.
- Dennett, D. C. (1987). Three kinds of intentional psychology. In *The intentional Stance* (chap. 3). Cambridge, MA: MIT Press.
- DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal: The Design and Perception of Humanoid Robot Heads. In *Proceedings of the 4th Conference on Designing Interactive Systems (DIS)* (pp. 321–326). New York, NY: ACM Press.

- Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3-4), 177–190.
- Ekman, P., & Friesen, W. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1), 49–98.
- Fink, J. (2012). Anthropomorphism and human likeness in the design of robots and human-robot interaction. *Social Robotics*, 199–208.
- Fiske, S., & Taylor, S. (1991). *Social cognition*. New York, NY: McGraw-Hill.
- Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing : A case study in the perception of animacy. *Cognitive Psychology*, 59(2), 154–179.
- Gibson, J. J. (1977). The concept of affordances. *Perceiving, acting, and knowing*, 67–82, Lawrence Erlbaum Associates, New Jersey.
- Goldberg, K., & Kehoe, B. (2013, January). *Cloud Robotics and Automation: A Survey of Related Work* (Tech. Rep. No. UCB/EECS-2013-5). EECS Department, University of California, Berkeley.
- Gray, J., Hoffman, G., Adalgeirsson, S. O., Berlin, M., & Breazeal, C. (2010). Expressive, interactive robots: Tools, techniques, and insights based on collaborations. In *HRI 2010 Workshop: What Do Collaborations With the Arts Have to Say About HRI?*
- Hall, E. T. (1969). *The hidden dimension*. New York, NY: Anchor Books.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2), 243–259.
- Hendriks, B., Meerbeek, B., Boess, S., Pauws, S., & Sonneveld, M. (2011). Robot vacuum cleaner personality and behavior. *International Journal of Social Robotics*, 3(2), 187–195.
- Hoffman, G. (2005). *HRI: Four Lessons from Acting Method* (Tech. Rep.). Cambridge, MA, USA: MIT Media Laboratory.
- Hoffman, G. (2012). Dumb robots, smart phones: A case study of music listening companionship. In *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Atlanta, Georgia.
- Hoffman, G., & Breazeal, C. (2009). Effects of anticipatory perceptual simulation on practiced human-robot tasks. *Autonomous Robots*, 28(4), 403–423.
- Hoffman, G., Kubat, R. R., & Breazeal, C. (2008). A hybrid control system for puppeteering a live robotic stage actor. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Munich, Germany.
- Hoffman, G., & Vanunu, K. (2013). Effects of robotic companionship on music enjoyment and agent perception. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.
- Hoffman, G., & Weinberg, G. (2011). Interactive improvisation with a robotic marimba player. *Autonomous Robots*, 31(2-3), 133–153.
- Höysniemi, J., Hämäläinen, P., & Turkki, L. (2004). Wizard of Oz prototyping of computer vision based action games for children. In *Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community* (pp. 27–34), College Park, Maryland.
- Hudson, S., Fogarty, J., Atkeson, C., Avrahami, D., Forlizzi, J., Kiesler, S., ... Yang, J. (2003). Predicting human interruptibility with sensors: a Wizard of Oz feasibility study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 257–264), Ft. Lauderdale, Florida.
- Inhatowicz, E. (1970). *Cybernetic Art*. Available from <http://www.senster.com/ihnatowicz/senster/>.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201–211.
- Jones, C. (1965). *The Dot and the Line*. Available at <http://vimeo.com/71888010>
- Ju, W., & Takayama, L. (2009). Approachability: How people interpret automatic door movement as gesture. *International Journal of Design*, 3(2), 1–10.
- Kelley, J. F. (1983). An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 193–196), Boston, Massachusetts.
- Kim, J., Kwak, S., & Kim, M. (2009). Entertainment robot personality design based on basic factors on motions: A case study with Rolly. In *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication, (RO-MAN)* (pp. 803–808).
- Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4), 513–549.
- Knapp, M. L., & Hall, J. A. (2002). *Nonverbal communication in human interaction* (5th ed.). Fort Worth:

- Harcourt Brace College Publishers.
- Knight, H. (2011). Eight lessons learned about non-verbal interactions through robot theater. In *Social robotics* (pp. 42–51). Springer.
- Kozima, H., Michalowski, M. P., & Nakagawa, C. (2009). Keepon: A playful robot for research, therapy, and entertainment. *International Journal of Social Robotics*, 1(1), 3–18.
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21(6), 575–580.
- Kraut, R. E. (1978). Verbal and nonverbal cues in the perception of lying. *Journal of Personality and Social Psychology*, 36(4), 380.
- Lasseter, J. (1986). Luxo Jr. Pixar Animation Studios. Pixar. Available from http://www.pixar.com/short_films/Theatrical-Shorts/Luxo-Jr.
- Lasseter, J. (1987). Principles of traditional animation applied to 3D computer animation. *Computer Graphics*, 21(4), 35–44.
- Lasseter, J. (2001). Tricks to animating characters with a computer. *ACM SIGGRAPH Computer Graphics*, 35(2), 45–47.
- Loula, F., Prasad, S., Harber, K., & Shiffra, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 210–20.
- Malle, B., Moses, L., & Baldwin, D. (Eds.). (2001). *Intentions and Intentionality*. MIT Press, Cambridge, Massachusetts.
- Maulsby, D., Greenberg, S., & Mander, R. (1993). Prototyping an intelligent agent through Wizard of Oz. In *Proceedings of the Conference on Human Factors in Computing Systems (INTERACT '93 & CHI '93)* (pp. 277–284), Amsterdam, The Netherlands.
- Meerbeek, B., Saerbeck, M., & Bartneck, C. (2009). Towards a design method for expressive robots. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, (pp. 277–278), San Diego, California.
- Michalowski, M., Sabanovic, S., & Kozima, H. (2007). A Dancing Robot for Rhythmic Social Interaction. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction* (pp. 89–96). Arlington, VA.
- Michotte, A. (1946). La perception de la causalité (Etudes Psychology, Vol. VI.).
- Moore, N.-J., Hickson, M., & Stacks, D. W. (2010). *Nonverbal communication: Studies and applications*. London: Oxford University Press.
- Mori, M. (1970). The uncanny valley. *Energy*, 7(4), 33–35.
- Mumm, J., & Mutlu, B. (2011). Human-robot proxemics: Physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI)*(p. 331). New York, NY: ACM Press.
- Murphy, R., Shell, D., Guerin, A., Duncan, B., Fine, B., Pratt, K., & Sourntos, T. (2010). A Midsummer Nights Dream (With Flying Robots). *Autonomous Robots*, 30(2), 143–156.
- Nomura, T., Suzuki, T., Kanda, T., Han, J., Shin, N., Burke, J., & Kato, K. (2008). What people assume about humanoid and animal-type robots: Cross-cultural analysis between Japan, Korea, and the United States. *International Journal of Humanoid Robotics*, 5(1), 25–46.
- Norman, D. A. (1999). Affordance, conventions, and design. *Interactions*, 6(3), 38–43.
- Paulos, E., & Canny, J. (1998). PRoP: Personal roving presence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 296–303), Los Angeles, California.
- Riek, L. D. (2012). Wizard of Oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1).
- Santos, K. B. (2012). *The Huggable: A socially assistive robot for pediatric care*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.
- Scholl, B., & Tremoulet, P. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8), 299–309.
- Setapen, A. (2011). *Shared Attention for Human-Robot Interaction*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.
- Sharma, M., Hildebrandt, D., Newman, G., Young, J. E., & Eskicioglu, R. (2013). Communicating affect via flight path Exploring use of the Laban Effort System for designing affective locomotion paths. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 293–300), doi: 10.1109/HRI.2013.6483602.
- Sirkin, D., & Ju, W. (2012). Consistency in physical and on-screen action improves perceptions of telepresence robots. *Proceedings of the Seventh Annual ACM/IEEE International Conference on*

- Human-Robot Interaction (HRI)*. doi: 10.1145/2157689.2157699
- Takayama, L., Dooley, D., & Ju, W. (2011). Expressing thought: Improving robot readability with animation principles. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI)* (pp. 69–76). ACM Press, doi: 10.1145/1957656.1957674
- Thomas, F., & Johnston, O. (1995). *The Illusion of Life: Disney Animation* (revised ed.). New York: Hyperion.
- Thornton, I. M., Pinto, J., & Shiffra, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology*, 15, 535–552.
- Tomasello, M. (1999). The cultural ecology of young children's interactions with objects and artifacts. *Ecological Approaches to Cognition: Essays in Honor of Ulric Neisser*, 153–170.
- Venolia, G., Tang, J., Cervantes, R., Bly, S., Robertson, G., Lee, B., & Inkpen, K. (2010). Embodied social proxy: Mediating interpersonal connection in hub-and-satellite teams. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1049–1058), doi: 10.1145/1753326.1753482
- Vertesney, L. (1995). Using video to prototype user interfaces. In *Human-Computer Interaction* (pp. 142–146).
- Weinberg, G., & Driscoll, S. (2007). The design of a perceptual and improvisational robotic marimba player. In *Proceedings from the 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 769–774). Jeju, Korea: IEEE, doi: 10.1145/70609.70615
- Wistort, R., & Breazeal, C. (2011). TofuDraw: A mixed-reality choreography tool for authoring robot character performance. In *Proceedings from the IDC Conference Idc* (pp. 213–216), doi: 10.1145/1999030.1999064
- Woods, S. N., Walters, M. L., Koay, K. L., & Dautenhahn, K. (2006). Methodological issues in HRI: A comparison of live and video-based methods in robot to human approach direction trials. In *Robot and human interactive communication*, . In *Proceedings from the 15th IEEE International Symposium on Robot and Human Communication (RO-MAN)* (pp. 51–58).
- Yankelovich, N., Simpson, N., Kaplan, J., & Provino, J. (2007). Porta-person: Telepresence for the connected conference room. In *CHI'07: Extended Abstracts on Human Factors in Computing Systems* (pp. 2789–2794), doi: 10.1145/1240866.1241080.

Authors' names and contact information: Guy Hoffman, Media Innovation Lab, School of Communication, IDC Herzliya, Israel. Email: hoffman@idc.ac.il. Wendy Ju, Center for Design Research, Stanford University, Stanford, USA. Email: wendyju@stanford.edu.