Fake It to Make It: Championing the value of design exploration in HRI

Anonymous Author(s)

ABSTRACT

Exploratory prototyping techniques are critical to devising new robot forms, actions and behaviors, and to eliciting human responses to designed interactive features, early in the design process. These low-cost, iterative methods evolve to rapidly respond to the designers' subjective sense of what works and what does not. We present and examine case studies of three design methods: video prototyping, virtual modeling and simulation, and Wizard of Oz prototyping interactions. In each case, we have applied these methods to author interactions for users to have with chair robots that we are in the process of designing.

In this work, we establish the contribution of design explorations such as these to the field of human-robot interaction, arguing design exploration *should instead* be focused on flexibility rather than specificity, possibility rather than replicability, and design insights as incubated subjectively through the designer rather than dispassionately proven by statistical analysis.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design process and methods**.

KEYWORDS

design, prototyping, experimentation, evaluation, Wizard of Oz, Virtual Simulation, Video prototyping

ACM Reference Format:
Anonymous Author(s). 2020. Fake It to Make It: Championing the value of design exploration in HRI. In *Proceedings of HRI '20: ACM Symposium on Human Robot Interaction (HRI '20).* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The need for and use of low-cost prototyping and elicitation techniques in design is well-established in human-computer interaction [8, 15].

In human-*robot* interaction, however, discussion of design techniques such as physical, video, and virtual prototyping, has inappropriately focused on meeting standards of verifiability and replicability [19, 38, 51] more properly applied to experimentation in research.

HRI '20, June 03–05, 2018, Cambridge, UK
2020. ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

Figure 1: What is the right way to interact with a ChairBot? Let us find the ways.

While the use of Wizard of Oz, video prototyping techniques, and simulation in experimentation is valuable, it is important not to apply the standards of research studies to every use of these methods. In this piece, we argue that as a *design* technique, the emphasis of methods such as Wizard of Oz experimentation of human robot interaction *should instead* be focused on flexibility rather than specificity, possibility rather than replicability, and design insights as incubated subjectively through the designer rather than dispassionately proven by statistical results.

In this paper, we argue for the importance of exploratory prototyping techniques in devising new robot forms, actions and behaviors, and in eliciting human responses to designed interactive features early in the design process. These methods are ideally low-cost, iterative, and should evolve to rapidly respond to the designer's subjective sense of what works and what does not. We present and examine case studies of three design methods: video prototyping, virtual modeling, and Wizard of Oz prototyping interactions. In each case, we have used these methods to author interactions for users to have with chair robots ("ChairBots") that we are in the process of designing. We then zoom out to argue for the need for the HRI community to consider the subjective reflections and reports from design explorations of human robot interaction as valid and valuable contributions to the community. Subjecting design explorations such as these to the standards of controlled research experiments prevents academic HRI researchers from sharing methods or insights early, when they are the most valuable, before hardware decisions have been finalized and are expensive to revisit.

2 CASE STUDIES: METHODS TO EXPLORE HUMAN-ROBOT INTERACTION

To provide specific fodder for our discussion, we share case studies of three different methods we have used to explore how we want to design and program the behavior of a set of chair robots. The human-robot interaction community seems to view actual autonomous interaction with actual robots in actual space as the ideal, most informative manner of prototyping interaction. However, actual interaction with actual robots in actual space costs a great deal of actual time and actual money, not to mention investment in design details that may well be off-target. Moreover, putting the actual robots performing actual interaction in front of the actual users that will eventually be using the robot may be highly impractical. It is often necessary and even preferable to use lower-cost, faster, and less actualized prototypes to develop robots and determine their behaviors, actions, and responses.

These case studies are focused on *formative ideas* rather than *conclusive results*—the intent is to provide insight into design directions worth pursuing further.

Physical prototypes with human-in-the-loop control allow for observing responses "in the wild", but are slow, technically challenging, and/or expensive to build. Physical prototyping enables observable field studies whose participants are free to move about, look around, and notice and interact with things as they would in any physical environment. Creating a novel physical system can require substantial engineering work, even if it employs off-the-shelf parts. Answerable questions include "how do people react to robots in the immediate physical environment?", "how do people respond to this motion?", and "how do human wizards respond to common participant responses?" Follow up studies should deploy more closed-loop control to enable repeatable motion—but early on, this adds complexity and development time, and does not contribute to the initial goal of understanding human behavior.

Video prototypes are pre-recorded videos of humans interacting with robots that are shown to participants, who are subsequently asked about their interpretation of the shown interaction. Answerable questions include "is the intent of this motion clear and/or interpretable?", "what does this motion look like to a third party observer?", and "how do subtle contextual changes affect the perception of the robot's motion?"

Virtual simulations straddle these worlds. While harder to develop than video prototypes, VR is however more immersive and allows for the observation and capture of participants' live responses to interactions. Pre-programmed motions ensure that participants are responding to precisely designed cues, while human-in-the-loop control allows observation of improvised, or better timed and adapted, wizard responses to participant interactions. But the lack of physicality prevents participants from interacting with the objects themselves, and the simulation itself may omit real-world physical effects, such as object inertia. [22]

3 PHYSICAL PROTOTYPES WITH HUMAN-IN-THE-LOOP CONTROL

A fundamental part of human-robot interaction is understanding and designing the physical interaction that occurs between humans and the physical robot. Through our motor memory, humans have an intuitive way to understand and engage with physical motion in the immediate environment. Physical prototyping allows us to explore how robots' motions are perceived through this process and implicitly understood. These interactions are mediated by the *System 1* mode of thought, defined by Kahneman as an automatic or fast way of thinking, requiring practically no effort [23].

A physical prototype should afford the kinds of physicality and interactions representative of those of the eventual robot. First, the robot should look similar to the type of object for which the motions are designed—a chair-robot should look like a chair, a sofa-robot like a sofa. A representative appearance is important because shape and size and their affordances have a large influence how humans interpret motion. Second, the degrees of freedom of the physical prototype should match those of what the intended ultimate object would similarly be capable of.

Prior Work

Low-cost physical prototyping has a rich history of use in evaluating different robotic objects, including a mechanical ottoman [43, 44], trash barrel robot [53], chairs [26], sofa [42], drawers [34], and an autonomous vehicle, in Ghost Driver [39].

Similarly, the use of live enactments as a generative design method is well-established in HCI, with variations including Burns et al.'s "Informance Design" [6] and Odom et al.'s "User Enactments", in which researchers use short embodied experiences with design concepts to explore "radical alterations to technologies' roles, forms, and behaviors." [35] As Oulasvirta et al. argue, a more accurate understanding of the implications of context can be supported by "bodystorming," that is, "carrying out design sessions in the original context, 'in the wild', instead of in the office." [37]

Prior work has also addressed the specific question of how to engage in rapid iteration with physical prototyping. One memorable example is Bartneck et al.'s exploration of how to address questions of form, purpose, social role, and environment through prototypes ranging from paper and mechanical mock-ups to Wizard of Oz remote control. [4]

Case Study

We used chair robot to prototype communicative expressions at a nearby university cafeteria. The robotic chair motion was controlled by a human operator, Wizard of Oz-style, to communicate its availability to be “sat in” to cafe patrons that had just paid for their food and were looking for a place to sit. If cafe goers noticed or interacted with the chair robot in any way, researchers would approach them after the interaction with a consent form and a survey. The survey asked about first impressions of the robot, and what was and was not “alright” about the interaction.

Our questions were intentionally open-ended so as to not prime for specific types of answers. Additionally, the space was video recorded with multiple cameras to allow us to reconstruct the detailed motion and interaction between the chair and participants. We found several themes in viewers’ responses ($N = 33$), and many, including social engagement and safety, came up in both the “alright” and “not alright” sections of the questionnaire. Some viewers indicated their preference for more pro-social behavior with comments including “it tries to sit in the table with us, it feels like the robot knows us”, while others wanted less social behavior, offering “it came to us a bit too slowly so it was surprising and a bit creepy...we weren’t able to use it (we were already sitting on regular chairs).”

Figure 2: A birds’-eye perspective of the cafe in which physical prototyping experiments took place.

In our experiments, we found that it was incredibly difficult to convince participants to sit in the robot chair. From participant responses and human controller observations, it seemed that the timing of the interactions was critically important to engaging participants.

Discussion

The most important outcome of the case study, for us, was to reveal the importance of timing in creating a successful interaction, and to suggest a critical question for follow-up exploration. Physical prototyping enables the discovery of this type of response, that only physical interaction can elicit, and that requires participants to encounter the robot without advance knowledge of their participation in a research study.

In general, physical prototyping is useful for developing an understanding of how humans will interact with a prototype robot physically, “in the wild,” in a space populated by other real humans. Through physical prototypes, we also get close to a “representative fidelity” that allows our results to transfer easily—but for this, we must pay special attention to designing the appearance of the robot. In the chair robot case, the form should appear as "a chair that is robotic" rather than "a robot that is a chair." This transferable design, in combination with unsuspecting bystanders, can yield a powerful research tool to explore new types of robots inhabiting various spaces.

Human-in-the-loop control enables researchers to experiment with different kinds of motion without engineering automation, however, it also limits fidelity as the human operator’s ability to respond might not reflect the final robot’s response times or agility. These constraints also affect the researcher’s ability to record and repeat identically the motions that contribute to successful interactions.

One challenge is the difficulty of controlling, recording, and then repeating the context in which all interactions take place. As a result, it can also be difficult to understand which factors in the environment—time, movement, lighting, number of bystanders—affect the interaction outcome.

4 VIDEO PROTOTYPES

Video prototypes can be very inexpensive to create, but robots depicted in the scene cannot respond to viewers of the video.

A basic version of this method involves video recording a human moving a robot—little more is needed than a phone with a camera and an object that is plausibly robotic.

A higher-fidelity version of this technique involves recording in front of a green screen and removing the human controller from the scene afterward, using video editing software. Removing the otherwise visible human controller imparts an illusion of autonomy to the robot, but takes more effort to create.

Prior Work

Researchers in HCI have long adapted video prototyping to their own needs. The inherent flexibility of the medium allows researchers to explore a range of visual expression and to realize ideas early in the design process [49]. Video has also been used, as we do here, in combination with Wizard of Oz techniques [32], to aid brainstorming, and as a design material itself [7].

Video prototyping as a technique in HRI has also been the subject of several validation studies [45, 52] showing that, although not suitable as a replacement for live field experiments, video-based HRI trials do report similar results to live trials under controlled conditions [52].

Figure 3: *Top*: Video prototype of an "I'm Available" interaction, using a visible human in a painter's coverall to move the chair. *Bottom*: Video prototype of the "Let's Help Out" interaction using a green screen and green suit to key out the human controller.

Case Study

We used an ordinary office chair with casters, manipulated by a human, to prototype communicative expressions for our chair robot design.

We created video prototypes to understand how chairs might communicate requests to engage with humans (such as "follow me"), or offer updates on their status ("I'm available") using their motion alone. The videos depicted a human controller physically guide chair's movement, steering from the top of its backrest, together with an actor in the same scene. In some of the videos, we digitally removed the chair's human controller, leaving only the chair and the actor visible.

To gain an understanding of the interpretability of the chair's motion, we recruited participants through Amazon's Mechanical Turk to view and comment on each scene. Each set of participants was asked to watch a randomly selected set of videos, and then to answer both open-ended questions about their interpretations of the motion, such as "Describe what you saw in the video" and "How would you characterize the chair?", and to rate, Likert-style, statements like "The chair was responsive to the person"

Our examination of viewers' responses yielded a few potential avenues for revision and future inquiry. We learned that certain messages were much easier to understand through motion than others, suggesting an investigation in what facets of a message might yield greater or lesser interpretability. Observed correlations between ratings of chairs as *responsive* and *expressive* suggest that we might aim to design movements that capture both traits.

Discussion

Video prototypes allow researchers to evaluate different motions and scenarios with the same chair in the same physical context. Scenario-building enables quick production of 2D video, which in turn allows quick collection of interpretability information. Videos are easy to make, but do not allow for interactive discovery—they are pre-recorded, and thus the videos' subjects can not react to observers.

The nature of video prototyping additionally provides limited opportunity for observing "in the wild" responses of bystanders because of the obvious human that is only removed in post-processing. Thus, the questions best answered by video prototypes include: "is the intent of this motion clear and/or interpretable?" and "what does this motion look like to a 3rd party observer?"

In this approach, it is possible to be flexible in several ways, including in the movement of the chair and its contexts (chair appearance, number of bystanders), and the method is inexpensive, quick, and does not require advanced technical skills. Researchers can adjust the videos in subtle ways to find differences in perception, while keeping other elements of the video constant.

Limitations. Results from cognitive psychology suggest that two people participating in a conversation are far more accurate, compared with bystanders, in understanding the content of the conversation—even with the same amount of information. This result can be attributed to grounding, a "process by which people establish a mutual understanding through their collaborative efforts to make their meanings intelligible to each other" [24]—suggesting that there are limits to what we can learn from bystanders' understandings of an interaction.

Two other limitations bear mentioning: a short video alone is not a complete representation of an interaction, and thus it is necessary to set the context with a cohesive story; and, participants may not empathize with the people they see in the video, as they themselves are not the ones interacting with the environment.

5 3D ANIMATION & SIMULATION

We use the term "3D simulation" to refer to two similar mechanisms:

(1) An immersive VR world with human-operated or code-controlled chairs.
(2) A screen-based "video game"-style world, shown and interacted using a web browser on a computer.

In each mechanism, participants are in control of some aspect of the environment, while most (or all) of the rest of the environment is simulated.

Prior Work

Simulated environments have proved a valuable tool for the development of autonomous robotic systems. High fidelity simulations have been used to train algorithms for tasks such as vehicle motion [13, 33, 41], complex object manipulation [29], and indoor navigation [27]. Interactive virtual environments have been used to train robot teleoperators [11]. There is an increasing body of work intended on making simulation for human robot interaction more accessible. Integrations between ROS and popular game engines [28] enable responsive robots to exist in the same virtual world as a human participant. Leveraging work from animation has been useful in designing social robots and their behaviors [2, 3, 46].

VR environments, while immersive, also suffer from limitations that appear inherent to the form. For example, proxemics are challenging to study in VR because, although proxemic preferences exist in VR [54], those preferences appear to be inconsistent between live and VR experiences [30].

Case Study

In a recent study, we explored how two control affordances—the top speed of a chair robot, and the perspective of the first-person camera view offered to the controller—affect the ways in which controllers navigate a crowded space to reach an goal location.

Using an online simulation environment, we drew on the social instincts from a large number of geographically distributed human robot controllers—(N = 30) online participants from Amazon Mechanical Turk—asking them to pilot a robot through a crowded 3D environment in order to elicit a variety of motion paths. To elicit socially appropriate behavior, participants were advised to "Try not to hit or scare anyone. Avoid getting in the way!" and to control the chair motion using the keyboard arrow keys.

We subsequently used another set of online participants (N = 30) to evaluate the social appropriateness of the elicited robot motions. The pilots were given one of two "chair's eye views," while the evaluators were placed in a third person perspective recreated in the same virtual world. We used a 2x2 within- and between-subjects study design, varying the participant's perspective—either above and behind the chair, or at seat level facing forward—as well as chair top speed—either the same speed as the humanoid agents, or 3 times quicker.

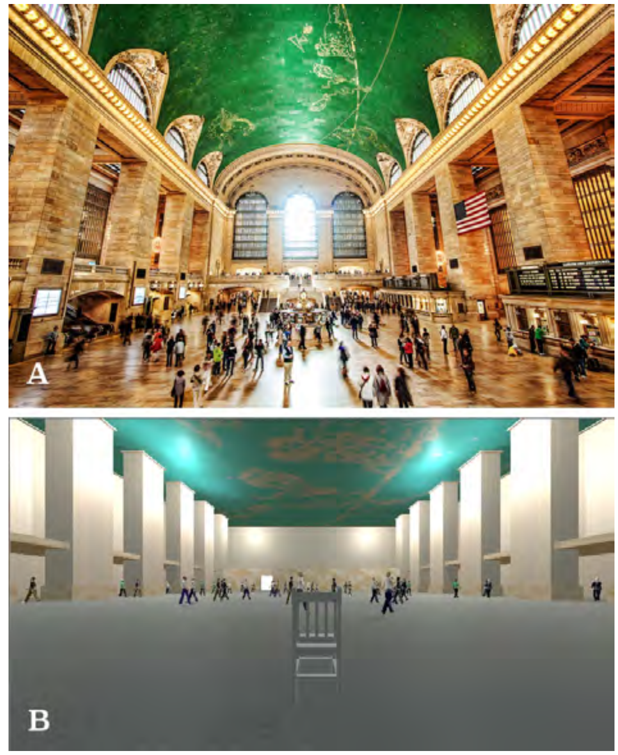

Figure 4: *A*: Grand Central Terminal, NYC. [1] *B*: Grand Central simulation, WebGL; note chair in foreground.

To develop an intuitive sense of this data, we plotted (Fig. 5) all the trajectories taken by each participant, including stops and collisions with passersby. Contrary to our expectations, we found that a high or low camera view perspective did not seem to make a difference to the paths controllers chose, but that controllers offered a higher top chair speed collided with passersby less frequently and for shorter durations.

Evaluators generally supported our internal predictions of which paths were the most socially appropriate, noting that paths that had more impacts with passersby also had lower mean ratings of social appropriateness.

Discussion

The primary lesson we learned from this case study was truly one we did not expect: our robots were too slow, and we plan follow-on studies to explore the effect of speed on the ability of robots to avoid collisions with passersby.

Though we did not take advantage of this property in our study, the capacity to replay an interaction from an infinite variety of perspectives is both valuable and unique to this method. This capacity offers tools for prototyping, reflection and analysis of interactions that are impossible in other methods. A good simulation can provide insight

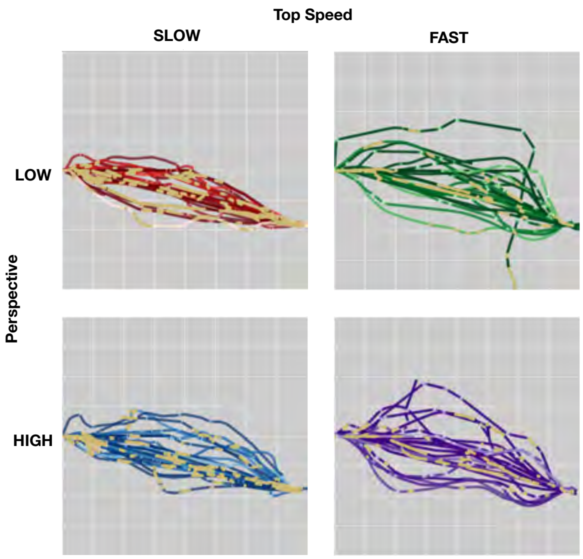

Figure 5: Our visualization of all controllers' trajectories, separated by affordance conditions. Trajectories start at left and end at right.

into the subjective feeling of the interaction and the actual negotiation between parties; the simulated robot can respond to the participant, and respond to the participant's responses too.

The fidelity of any simulation should be matched to the questions being asked. There will be numerous limitations inherent in accessible, high-quality 3D simulations for the foreseeable future. However even relatively low-fidelity 3D simulations can provide an opportune middle ground for design exploration between video prototyping and physical prototyping—and in some ways surpassing both: In a virtual world, the initial conditions of any interaction are perfectly reproducible while remaining fully interactive. This makes it simple to do iterative prototyping for interaction to test many variations with slightly different initial conditions or different affordances. Although we argue against the emphasis on replicability in design prototyping, 3D simulations do offer an avenue for scientific inquiry when a single variable needs to be tested.

Like video, 3D simulation has the capacity to scale. A WebGL-based simulation can be shown to thousands of of people on a crowd-work platform in hours; simple Mobile-based VR has been made incredibly cheap [31] while new high end VR headsets continue to enter the market. As VR becomes more popular, the tools will become more accessible and creating an immersive simulated environment will become an easier skill to master and use in prototyping.

This capacity for scale enables questions that could prove difficult with a physical system: simulations can offer identical interactions in different cultural contexts, and researchers can test systems in physical spaces where permission would be required and unavailable, or where robot scale prevents real-world evaluation.

Limitations. Though powerful, 3D simulations are not well-suited for exploring phenomena that are inherently physical or that require real-life contexts. For example, anything haptic—we can not yet simulate what it is like to touch things; physically impossible robots are often easier to program than realistic ones, and researchers can ignore physical limitations. This ability has advantages as well as disadvantages, but one disadvantage in particular is that while bodily harm is impossible, knowledge of that fact can act as an encouragement for participants to do things that are unrealistic. Lastly, the constructed nature of simulations makes this method unsuitable for "field studies" of unsuspecting participants.

6 DISCUSSION

In their classic work, "What do prototypes prototype?," Houde and Hill identify the key reason why multiple forms of prototyping can be necessary:

> Since interactive systems are complex, it may be difficult or impossible to create prototypes of a whole design in the formative stages of a project. Choosing the right kind of more focused prototype to build is an art in itself, and communicating its limited purposes to its various audiences is a critical aspect of its use. [21]

Understanding the purpose of a given prototype, and designing it to answer a specific question, is critical to developing knowledge of the greater set of possibilities within a given design space. In this section, we detail the rationale for prototyping robot interaction.

Figuring out the right design

The British Design Council's double-diamond design process [10] sets out four phases in the design process: discover, design, develop and determine. Design exploration is the primary goal of the third phase, where designers seek to maximize the design alternatives presented to address the problems and opportunities found in the first two phases of the process. As depicted in Fig. 6, the development phase is a divergent process, intended to generate a wide set of alternatives and solutions. If the search space for possible solutions or solution features is not cast widely enough, the outcomes will be stunted or sub-optimal [9, 16]. Hence, the goal of prototyping methods is to help designers to generate more alternatives, more divergence in alternatives, and to enable them to elicit stronger and more detailed feedback from users [48].

While in traditional product design the emphasis is often placed on physical form, the advent of computers and robots

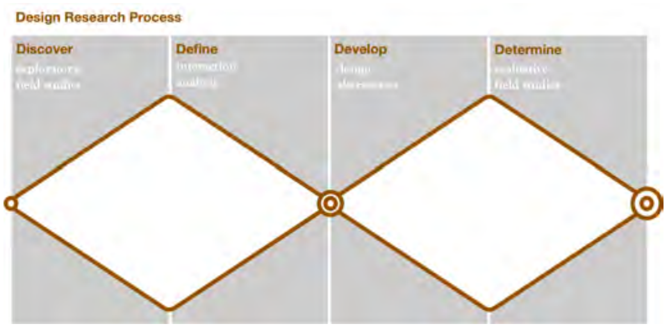

Figure 6: The British Design Council’s double-diamond design process

has broadened the space of exploration to include motion, action, and behavioral response. Hoffman and Ju [20] have described numerous methods to design the movements of robots at the outset, since the movements are necessarily constrained or enabled by the physical form of the machine.

In the following discussion, we describe why we believe the methods described above are powerful, and should be used by human-robot interaction researchers to generate early design insight.

Intangible benefits

The oft-heard critique that design explorations did not “use a real robot” belies the principle, well understood in traditional design and engineering contexts, that increased fidelity early in the design process does not make for better outcomes [18]. The value of low-cost and rapid prototyping is obvious, but low-fidelity prototyping is less well known to be a critical practice used by designers to construct design knowledge, communicate and explore ideas in order to make decisions later in the design process [40, 50].

Prototyping without an actual robot also helps robotics researchers understand where the lower bound lies: what is *truly needed* in a robot? Does this arm need three joints? Does this delivery robot need omniwheels? Does this turntable need high-power stepper motors? Perhaps not, but we are liable to make use of these features if they are available to us, even if a lesser alternative would suffice.

Realizing that a “real robot” is not necessary to make key discoveries and contributions to what kinds of interaction and design are desirable is a liberating insight. It is also a matter of equity; low-fidelity prototyping allows researchers without the resources to buy an expensive robot platform, or access to a programmer, to still engage in HRI research. Greater participation means more voices and more knowledge of alternatives and how they compare with one another. Dow et al. note that within a collaborative team, sharing multiple designs improves exploration, group rapport, and results [14]. In the larger academic and industrial context, sharing design explorations similarly should allow researchers to explore and improve community dynamics, the general knowledge base, and eventual design outcomes [17].

Eliciting the other half

Beyond the knowledge that prototyping is central to any design process, it is specifically important for human-robot interaction design, because the success of such interactions hinges on understanding the human side of the interaction. Wizard-of-Oz and enactment techniques can be used to draw out the many ways a human might respond to various robot behaviors. For example, the elicitation of naturalistic human typewritten and speech responses to faux intelligent natural language systems in the early 1980’s was critical to the development of actual natural language understanding in the following decades [12, 25].

Understanding when and how design generalizes

Ironically, in the quest to obtain “validity” by requiring that the specifications for a Wizarded robot behavior be tightly specified and controlled [38], HRI researchers are likely to sacrifice generalizability. In general, design insights generalize when we can find elements of commonality across design instances—feedback on one prototype’s affordances can offer insights into another prototype’s similar affordances. Overgeneralization, however, can result from or result in invalid assumptions and underexplored design spaces. Understanding the nature of the commonality is made possible by rapidly testing out assumptions across a broader design space.

Beyond a single project, designs can generalize across domains when there is commonality in specific areas; a chair robot prototype, for example, can offer insights into other mobile robots with similar affordances or similar motion, or other furniture, depending on the nature of the feedback and the context of use. Design processes themselves can also generalize: approaches for exploring one design space may apply to another.

Comparing and combining exploration methods

The methods described here each have their own advantages and disadvantages individually, but often complement each other in ways that provide a more complete understanding of the phenomena under study. Three specific factors bear discussion.

Controlling robot motion. Video prototypes and virtual simulations provide the easiest interfaces for producing naturalistic robot motions in a variety of contexts, but at the cost of realism—due to visible human operators or physical constraints on green screen deployment—and responsiveness to participant behavior. The ease and low cost with which a variety of robot motions and interactions can be explored enables very rapid iteration at the earliest stages of design—when the mere act of viewing a video playback of your own best guess at motion alone provides a great deal of insight. These prototypes can be plentiful and generative, and used in the way that sketches are used elsewhere in design.

In our own work, we have often led with storyboards and video prototypes, moving to 3D simulations or physical prototyping after developing an intuition for what types of robot capabilities and affordances are worth examining in real life with real participants. The chair robot explorations described here have a longer history than the individual case studies presented, ultimately beginning with video recordings from a workshop of human-moved chairs dancing in a large atrium space.

Participant immersion; naturalistic environments. In a video prototype, compared with a simulated environment, it is easy to include real people in real physical contexts, which can help viewers imagine the real experience, and avoid the distraction and prejudice that comes with a video game-style rendering environment. Combining video prototyping and Amazon Mechanical Turk, it becomes easy to get an understanding of participants' perspectives as bystanders to nearly any interaction.

A simulated environment, in contrast, loses some of this flexibility but gains the ability to observe how participants respond to pre-programmed or human-operator-controlled robot motion. In addition, a simulated environment makes recording all actions and interactions extremely easy—where a physical environment might require a complex tracking setup to collect data on participants' motions, a simulation already contains, and can record, all data about participants' use of provided affordances. In simulation, researchers additionally have the flexibility and control to violate physical laws, put people in semi-unsafe situations—as is especially useful in driving studies—and program events to happen reliably and repeatably.

A physical prototype affords none of these latter advantages, but does allow a degree of immersion impossible to otherwise achieve. A robot rapidly approaching you in the physical world is a very different from observing the same on a monitor or even in VR, and this can make the physical prototype the more ecologically valid choice. As Rodney Brooks notes, "The world is its own best model" [5].

Physical prototypes complement video prototypes by offering strong evidence of behavior "in the wild" with unprimed participants, limiting priming effects [47] and demand effects [36]. These *in-situ* studies can both validate the insights generated by earlier methods, as well as provide the "next step" of insight—an understanding of how the dynamic human-robot system evolves after the first or second interactions that can be plausibly prototyped in video.

Technical effort. Significant engineering effort may be required to create a physical robot prototype capable of performing the same motions, with a similar level of fidelity, that a human can perform by manually manipulating an armature or by moving a large, heavy "mobile" object such as a chair. Thus, expending this effort has the most value once the potential for design insights from quicker, lower-cost methods has been exhausted.

In general, video prototypes and simulated environments are less work to develop than physical prototypes, though there are many exceptions—in cases where an existing robotics platform may be suitable, researchers may be able to trade off engineering effort for monetary expense. The cost of re-recording a video, or even modifying a simulation, is small compared with the effort required to re-engineer a physical prototype; in the earliest stages of design, iteration speed is critical.

7 CONCLUSION

We argue for the importance of using a variety of prototyping techniques for studying the design space of robots and their use. We believe that a "real robot" is not required for meaningful and critical contributions to the field of human-robot interaction, that this insight is critical to enabling participation from a broader and more diverse set of researchers, and that there is great value in sharing design insights that suggest interesting avenues of exploration, but might not yet have experimentally proven findings to back them up.

REFERENCES

[1] 2012. *Wikimedia Commons.* https://commons.wikimedia.org/wiki/File:Grand_Central_Terminal_Lobby.jpg

[2] Etienne Balit, Dominique Vaufreydaz, and Patrick Reignier. 2016. Integrating animation artists into the animation design of social robots: An open-source robot animation software. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction.* IEEE Press, 417–418.

[3] Etienne Balit, Dominique Vaufreydaz, and Patrick Reignier. 2018. PEAR: Prototyping Expressive Animated Robots-A framework for social robot prototyping. In *13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018).*

[4] Christoph Bartneck and Jun Hu. 2004. Rapid prototyping for interactive robots. In *The 8th conference on intelligent autonomous systems (IAS-8).* 136–145.

[5] Rodney A Brooks. 1990. Elephants don't play chess. *Robotics and autonomous systems* 6, 1-2 (1990), 3–15.

[6] Colin Burns, Eric Dishman, William Verplank, and Bud Lassiter. 1994. Actors, hairdos & videotape—informance design. In *Conference companion on Human factors in computing systems.* ACM, 119–120.

[7] Jacob Buur, Thomas Binder, and Eva Brandt. 2000. Taking video beyond ‘hard data’ in user centred design. In *Participatory design conference*. Citeseer, 21–29.

[8] Bill Buxton. 2007. *Sketching User Experiences: Getting the Design Right and the Right Design (Interactive Technologies)*. Morgan Kaufmann.

[9] Bill Buxton. 2014. Multi-Touch Systems that I Have Known and Loved. *BillBuxton.com* (2014). http://www.billbuxton.com/multitouchOverview.html. Accessed: 2014-12-24.

[10] Design Council. 2005. The ‘double diamond’ design process model. *Design Council* (2005).

[11] Jeff Craighead, Jennifer Burke, and Robin Murphy. 2008. Using the unity game engine to develop sarge: a case study. In *Proceedings of the 2008 Simulation Workshop at the International Conference on Intelligent Robots and Systems (IROS 2008)*.

[12] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of Oz studies—why and how. *Knowledge-based systems* 6, 4 (1993), 258–266.

[13] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. 2017. CARLA: An open urban driving simulator. *arXiv preprint arXiv:1711.03938* (2017).

[14] Steven Dow, Julie Fortuna, Dan Schwartz, Beth Altringer, Daniel Schwartz, and Scott Klemmer. 2011. Prototyping dynamics: sharing multiple designs improves exploration, group rapport, and results. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Acm, 2807–2816.

[15] Steven P Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L Schwartz, and Scott R Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction (TOCHI)* 17, 4 (2010), 18.

[16] Steven P Dow, Kate Heddleston, and Scott R Klemmer. 2009. The efficacy of prototyping under time constraints. In *Proceedings of the seventh ACM conference on Creativity and cognition*. ACM, 165–174.

[17] Christiane Floyd. 1984. A systematic look at prototyping. In *Approaches to prototyping*. Springer, 1–18.

[18] Elizabeth Gerber and Maureen Carroll. 2012. The psychological experience of prototyping. *Design studies* 33, 1 (2012), 64–84.

[19] Anders Green, Helge Huttenrauch, and K Severinson Eklundh. 2004. Applying the Wizard-of-Oz framework to cooperative service discovery and configuration. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*. IEEE, 575–580.

[20] Guy Hoffman and Wendy Ju. 2014. Designing robots with movement in mind. *Journal of Human-Robot Interaction* 3, 1 (2014), 89–122.

[21] Stephanie Houde and Charles Hill. 1997. What do prototypes prototype? In *Handbook of human-computer interaction*. Elsevier, 367–381.

[22] Yuki Kado, Takanori Kamoda, Yuta Yoshiike, P Ravindra De Silva, and Michio Okada. 2010. Sociable dining table: The effectiveness of a “KonKon” interface for reciprocal adaptation. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 105–106.

[23] Daniel Kahneman. 2011. *Thinking, fast and slow*. Macmillan.

[24] Yoshihisa Kashima, Olivier Klein, and Anna E Clark. 2007. Grounding: Sharing information in social interaction. *Social communication* (2007), 27–77.

[25] John F Kelley. 1983. An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 193–196.

[26] Heather Knight, Timothy Lee, Brittany Hallawell, and Wendy Ju. 2017. I Get It Already! The Influence of ChairBot Motion Gestures on Bystander Response. In *Robot and Human Interactive Communication (RO-MAN), 2017 25th IEEE International Symposium on*. IEEE, 443–448.

[27] Eric Kolve, Roozbeh Mottaghi, Daniel Gordon, Yuke Zhu, Abhinav Gupta, and Ali Farhadi. 2017. Ai2-thor: An interactive 3d environment for visual ai. *arXiv preprint arXiv:1712.05474* (2017).

[28] D Krupke, S Starke, L Einig, F Steinicke, and J Zhang. 2017. Prototyping of immersive HRI scenarios. In *International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines*. World Scientific, 537–544.

[29] Youngwoon Lee, Edward S Hu, Zhengyu Yang, Alex Yin, and Joseph J Lim. 2019. IKEA Furniture Assembly Environment for Long-Horizon Complex Manipulation Tasks. *arXiv preprint arXiv:1911.07246* (2019).

[30] Rui Li, Marc van Almkerk, Sanne van Waveren, Elizabeth Carter, and Iolanda Leite. 2019. Comparing human-robot proxemics between virtual reality and the real world. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 431–439.

[31] Dan MacIsaac. 2015. Google Cardboard: A virtual reality headset for 10? *The Physics Teacher* 53, 2 (2015), 125–125.

[32] W. E. Mackay, D. S. Pagani, L. Faber, B. Inwood, P. Launiainen, L. Brenta, and V. Pouzol. 1995. Ariel: Augmenting paper engineering drawings. *Conference on Human Factors in Computing Systems* (1995). http://portal.acm.org/citation.cfm?id=223763

[33] Wei Meng, Yuchao Hu, Jiaxin Lin, Feng Lin, and Rodney Teo. 2015. ROS+ unity: An efficient high-fidelity 3D multi-UAV navigation and control simulator in GPS-denied environments. In *IECON 2015-41st Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 002562–002567.

[34] Brian Mok, Stephen Yang, David Sirkin, and Wendy Ju. 2014. Empathy: interactions with emotive robotic drawers. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 250–251.

[35] William Odom, John Zimmerman, Scott Davidoff, Jodi Forlizzi, Anind K Dey, and Min Kyung Lee. 2012. A fieldwork of the future with user enactments. In *Proceedings of the Designing Interactive Systems Conference*. ACM, 338–347.

[36] Martin T Orne. 1962. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American psychologist* 17, 11 (1962), 776.

[37] Antti Oulasvirta, Esko Kurvinen, and Tomi Kankainen. 2003. Understanding contexts by being there: case studies in bodystorming. *Personal and ubiquitous computing* 7, 2 (2003), 125–134.

[38] Laurel D Riek. 2012. Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction* 1, 1 (2012), 119–136.

[39] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. 2016. Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*. IEEE, 795–802.

[40] Michael Schrage. 1999. *Serious play: How the world’s best companies simulate to innovate*. Harvard Business Press.

[41] Shital Shah, Debadeepta Dey, Chris Lovett, and Ashish Kapoor. 2018. Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and service robotics*. Springer, 621–635.

[42] David Sirkin and Wendy Ju. 2014. Using embodied design improvisation as a design research tool. In *Proceedings of the international conference on Human Behavior in Design (HBiD 2014), Ascona, Switzerland*.

[43] David Sirkin, Brian Mok, Stephen Yang, and Wendy Ju. 2015. Mechanical ottoman: how robotic furniture offers and withdraws support. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 11–18.

[44] Marco Spadafora, Victor Chahuneau, Nikolas Martelaro, David Sirkin, and Wendy Ju. 2016. Designing the behavior of interactive objects. In *Proceedings of the TEI’16: Tenth International Conference on Tangible,*

Embedded, and Embodied Interaction. ACM, 70–77.

[45] Dag Sverre Syrdal, Nuno Otero, and Kerstin Dautenhahn. 2008. Video prototyping in human-robot interaction: Results from a qualitative study. (2008).

[46] Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing thought: improving robot readability with animation principles. In *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on*. IEEE, 69–76.

[47] Steven P Tipper. 1985. The negative priming effect: Inhibitory priming by ignored objects. *The quarterly journal of experimental psychology* 37, 4 (1985), 571–590.

[48] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. 2006. Getting the right design and the design right. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 1243–1252.

[49] Laurie Vertelney. 1989. Using video to prototype user interfaces. *ACM SIGCHI Bulletin* 21, 2 (1989), 57–61.

[50] Matthew B Wall, Karl T Ulrich, and Woodie C Flowers. 1992. Evaluating prototyping technologies for product design. *Research in Engineering Design* 3, 3 (1992), 163–177.

[51] Astrid Weiss, Regina Bernhaupt, Daniel Schwaiger, Martin Altmaninger, Roland Buchner, and Manfred Tscheligi. 2009. User experience evaluation with a Wizard of Oz approach: Technical and methodological considerations. In *2009 9th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 303–308.

[52] Sarah Woods, Michael Walters, Kheng Lee Koay, and Kerstin Dautenhahn. 2006. Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *9th IEEE International Workshop on Advanced Motion Control, 2006.* IEEE, 750–755.

[53] Stephen Yang, Brian Ka-Jun Mok, David Sirkin, Hillary Page Ive, Rohan Maheshwari, Kerstin Fischer, and Wendy Ju. 2015. Experiences developing socially acceptable interactions for a robotic trash barrel. In *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on*. IEEE, 277–284.

[54] Nick Yee, Jeremy N Bailenson, Mark Urbanek, Francis Chang, and Dan Merget. 2007. The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior* 10, 1 (2007), 115–121.