Research through Evocative Play:
Play-based Methods for Drawing out Contextual Complexities and Understanding Power

Anna Michele Kasunic Das

June 30, 2019

Doctoral Dissertation, CMU-HCII-19-102

Research through Evocative Play

Play-based Methods for Drawing out Contextual
Complexities and Understanding Power

Anna Michele Kasunic Das

Committee

Geoff Kaufman, HCII, Carnegie Mellon University, **Chair**
Jeff Bigham, HCII, Carnegie Mellon University
Laura Dabbish, HCII, Carnegie Mellon University
Jessica Hammer, HCII/ETC, Carnegie Mellon University
Saiph Savage, Human-Computer Interaction Lab, West Virginia University

June 30, 2019

Anna Michele Kasunic Das

Research through Evocative Play: Play-based Methods for Drawing out Contextual Complexities and Understanding Power

Doctoral Dissertation, CMU-HCII-19-102, June 30, 2019

Chair: Geoff Kaufman

Committee: Jeff Bigham, Laura Dabbish, Jessica Hammer, Saiph Savage

Carnegie Mellon University

The eHeart Lab

Human-Computer Interaction Institute

School of Computer Science

Pittsburgh, Pennsylvania, USA, 15213

Abstract

In this dissertation, I present *research through evocative play*, an empirical research method inspired by and drawing upon design research (research through design, critical design and ludic design) and play research and theory, and that uses declarations of play to reveal nuances of a context and investigate power dynamics therein. Viewing play as a research tool rather than a design end, the approach facilitates the revelation of participants' relationship with that context and others in the context, as well as the ambiguities, conflicts, uncertainties, and discontent those relationships might encompass. In *research through evocative play*, the researcher designs and declares play in the context to encourage participants to reflect on and engage with the context in novel ways. In this manner, *research through evocative play* also positions the researcher-designer as an active and integral participant in the study whose perspectives and actions should be critically analyzed and reflected upon as part of the research process.

To demonstrate my path using this approach, I present three related projects. As a precursor to my *research through evocative play* approach, I first present my mixed methods work on the subreddit r/RoastMe, an online forum community in which people post photos of themselves to be harshly ridiculed by others. I show how the play declaration of "comedy not hate" casts online self-presentation behaviors and harsh humor as play, and explore how this play declaration reveals participants' views on and relationships to standards of behaviors for self-presentation and politeness in related contexts, and ambiguates power dynamics and ludic consent within the space of RoastMe.

Next, I discuss Turker Tales, a Google Chrome extension implemented with 171 participants on Amazon Mechanical Turk (MTurk). MTurk is a platform where remote crowd workers produce labor for requesters, often for low pay and with limited platform-supported means of voicing their concerns and communicating with peers. Turker Tales allows participants to anonymously create, view and share short identity-based narratives with workers completing the same or similar labor tasks on MTurk. With Turker Tales, I suggest novel directions for supporting crowd workers, beginning fleshing out *research through evocative play* as a research approach to promote criticism of and reflection on a context and its attendant power dynamics,

and highlight the ethical implications of playing with and within a capitalist structure where power is imbalanced.

Building off my research findings and approaches in both RoastMe and TurkerTales, I lastly present YouMercials, a concept and functioning prototype for a design that declares YouTube advertising as a space for play, encouraging participants to manipulate YouTube advertisements by replacing the original audio tracks or by creating short identity-based imagination exercises for viewers to consider while watching advertisements. With YouMercials, I further explore and directly manipulate elements from RoastMe and Turker Tales, including direct play declarations, the use of roasting humor, the implications of play declaration within a capitalist context, and the anonymous sharing of short narrative-based shared artifacts. Through YouMercials, I analyze participants' ambivalent relationships to YouTube and YouTube advertising, reflect on the role of the researcher in *research through evocative play*, and discuss both the values and limitations of *research through evocative play* as a study approach.

My work contributes methodologically to human-computer interaction (HCI), design research, and play research by proposing the *research through evocative play* approach. In addition, as a side product of my pursuit of this approach, my work also contributes recommendations about the effectiveness of specific forms of play in engaging participants that can be useful to researchers and practitioners in human-computer interaction and play.

Acknowledgements

What follows is only a partial list; there are way too many people to thank.

First and foremost, I need to thank my husband, Sauvik Das. [1] During these last few years of the PhD, he has seen me at some of my worst, has stayed by my side when running away was probably the more logical and lucrative action to take, and has shown me steadfast acceptance, love, and support. So much has changed for us both in the past few years, and I still can't quite fathom that all of those changes have actually taken place. I can't promise that it will be smooth sailing from here on out, but I am so looking forward to figuring out the next phase of our lives together, including, of course, the impending birth of our first child together.

I also want to thank my family, especially my mom, Mary Claire, who has been a constant emotional support for me throughout my life and has continued to be so during my PhD; my dad, Jon, who has been my at-home office mate in Pittsburgh during the past year and a half and has generously put up with me being annoyed by him invading "my" space in his house; my sisters, Raquel and Elise; my brother Nick, who has been the most involved in my research of anyone in my family; my siblings-in-law, especially Louis, who was the person who most encouraged me to apply to the Human-Computer Interaction Institute (HCII) in the first place; my niece and nephews; and other extended family, including my cousin Julia, who helped make semi-living with my parents as I split my time between Pittsburgh and Atlanta during the end of the program less stressful. My family has helped me stay grounded, reminding me that there is meaning and value in life, and there are people who I care about and vice versa, no matter how poorly things might have been going at times in the PhD program. They also served as guinea pigs for testing out nearly every new study protocol I undertook, while still permitting me to otherwise refuse to discuss any aspects of my research with them. Also thanks to my in-laws in Atlanta, especially Sangeeta, Samir, Sunny, Amisha, Ari, and Mini, who put up with me being MIA for many family occasions in the time leading up to the defense.

---

[1] Whose last name, incidentally, I have not yet formally/legally adopted, but nonetheless have chosen to use in this document.

on the difficult questions he tends to pose have strongly influenced my research approach.

I also want to thank the professors of the courses I took while an HCII student, especially Kristen Kovak of the Drawing for Non Majors course, encouraged me to pursue the vending machine project as my final project for the course, and whose guidance and support were very valuable to me at a time when I was just starting to recover from a period of especial struggle; Mark Stehlik of 15-112, and Jeffrey Epstein of Web Development, where I gained more confidence in coding; and Peter Scupelli of Design Ethos & Action, a course that helped me reflect on my ethos, values, and what I want to do in my life, in addition to allowing me to explore a project analyzing my own engagement with online advertisements that later guided me in my study of YouMercials.

Thanks to all the CMU faculty who talked with me and passed on their wisdom, guidance, criticism, and support, especially but not limited to Carolyn P. Rosé, Jason Hong, Nikki Kittur, Laura Dabbish, Jessica Hammer, Amy Ogan, Chinmay Kulkarni, John Zimmerman, Jodi Forlizzi, Robert Kraut, Scott Hudson and Jeff Bigham. A huge thanks goes to all the HCII and IRB staff (past and present) who have been so incredibly helpful and whose work is often behind-the-scenes but invaluable, especially (but not limited to) Lauren Hardwig, Justin Puglisi, Ja'Ron Pitts, Ryan Ries, Indra Danti, Nicole Willis, Ebony Dickey, Nancy Beatty, Angela Huang, Valerie Quickel, Lindsay Olshenske, Gail Kusbit, and Doug McFarland. I also want to thank Jim Skees, Paul Stockhausen, and Jamie Gregory from facilities, who were always super helpful.

I now want to thank my peers (PhD students and other students/colleagues I have met along the way), and I immediately feel overwhelmed, because there are just way too many people to thank and too many things to say; I know I will be leaving people out. Because it might be strange to have an Acknowledgements section that is as long as the rest of my dissertation, I will also cut myself off from too much gushing. I often have trouble starting up, feeling comfortable in, and maintaining friendships (and you know, engaging in social interactions, in general), so I feel so fortunate to have made so many friends here that I truly cherish.

First, a humongous thanks to my cohort, Judeth Oden Choi, Xu Wang, Anhong Guo, Nathan Hahn, Eunki Jung, Qian Yang, and Amy (Shannon) Cook. Judy and Eunki were my first officemates in the HCII, and I am so grateful for that and have very fond memories of that time. Judy has been a true friend through the years. I used to hide my emotions, my mental issues, etc. from friends; Judy was one of the first friends I started opening up to in more ways, and I've continued that trend of opening up with people while at CMU, to the point that now I think I overshare

and am sometimes too open and honest. Xu and Anhong, one or both of whom have been my office mate or unofficial office mate every year since year 2, have been incredible friends to me, there to support, encourage and have fun with me (even if that meant stopping whatever they were doing to: talk with me, console me, help me find my lost shoe by scouring the entire CMU campus, go on "explorations" to discover hidden gems on campus such as secret showers, or inadvertently steal sodas from Robotics, take an impromptu ten mile walk as I with me as I "process my feelings," be a bridesmaid in my wedding, etc.). I also want to thank all the present and past members of the eHeartLab, Alexandra To, Joseph Seering, Fannie Liu, and Joseyln McDonald, who not only were awesome and supportive research colleagues and collaborators, but also among my closest friends in the HCII.

Other graduated and current students and colleagues that have been especially amazing: Samantha Finkelstein (who not only helped me through some of my darkest times at CMU, but also encouraged me to pursue a relationship with my now-husband); David Gerritsen (who was my go-to "senior student" support in my first couple years of the PhD); Diyi Yang (who I look forward to being an ATL neighbor with soon) as well as all the LTI and HCII members of the social computing group; Mary Beth Kery; Siyan Zhao (especially fond memories of TA-ing PUI together); Kenneth Holstein; Beka Gulotta (a model for what senior students should be to junior students); Erik Harpstead; Jenny Olsen; Franceska Xhakaj; Kelly Rivers; Jeff Rzeszotarski; Tatiana Vlahovic; Dan Tasse (who was especially helpful in giving me tips on how to deal and navigate when your spouse take a job in a different city); Michael Rivera; Steven Dang; Nesra Yannier; Nikola Banovic; Annie Maholtra; Sougata Sen; Yasmine Kotturi; Michal Luria; Annabel Sun; Rushil Khurana; Michael Madaio; Judith Odili Uchidiuno; Joseph Chee Chang; Michael Xieyang Liu (who recently saved my butt by directing me to a fabulous tutorial about Google Chrome's CORS changes when I found that my YouMercials extension was suddenly completely non-functional); Lynn Kirabo; Samantha Rieg; Gierad Laput; Yang Zhang; Vikram Kamath; Amber Horvath; Cori Faklaris; Felicia Ng. The list goes on— at this point, I really should just say "ALL THE STUDENTS." Honestly, I could easily go through everyone's names and talk about at least one reason for which I am grateful to them.

I also want to thank the many master's and undergraduate students that I've had the pleasure to work with and get to know as a teaching assistant for PUI and UCRE, and as a student instructor for the intro to programming master's in human-computer interaction (MHCI) course (also shout out to Jason Hong and Nicole Willis for helping to make that possible); the "RoastMe Part 2" independent study team (special shout out to Lily (Hsin Yu) Pai for prototyping an idea related to YouTube advertising that greatly influenced my work with YouMercials, and thanks also to the rest of the team for their creativity and feedback: Xi Jin, Jerry (Jee-Uk) Fu,

Sarah Kim, Karen Kim, and Max (Yutong) Zhu); to the Character Creation project duo, Anurati Sodani and Xiaofei Zhou; to all the students in the VR biosignals team (special shout out to co-instructor Fannie Liu, who was the real driver of the course and a pleasure to work with); and to the students in the exploratory VR trans embodiment independent study (who were especially patient and understanding when the whole project basically fell apart). It seems impractical to list all the students here, but I just want to express how much those experiences meant to me, and how motivating and invigorating it was to work with so many hard-working and inspiring individuals. Teaching/ TA-ing were highlights of my time in the HCII.

In addition, I'm grateful to the GroupX staff that helped me work out my stress and frustrations, including Karen, Lisa, Katie, Dabney, Willy, and Pattye; to students and professors from other universities I've met (virtually or in person) that have provided feedback and inspiration, including Saiph Savage (who I first met at CHI 2017, who motivated me to stay in the PhD program, and who eventually became first a regular collaborator and later a valued committee member), Josh Miller, Neha Kumar, Andrea Stevenson Won, Mark Riedl, Sahana Srivatsan, and Azra Ismail; to my former professors at the Heinz College and boss at the Homewood Children's Village who wrote me the recommendations that helped me get accepted in the first place, Jonathan P. Caulkins, Robert Strauss, and Shannah Tharp-Gilliam; to my former neighbor and University of Pittsburgh post-doc Morteza Ghasri for his encouragement; to the organizers and attendees of Meaningful Play 2018, which was perhaps the only academic conference I attended that I actually enjoyed on multiple levels (especial thanks to Alec Holowka, whose poignant and honest talk on his game Night in the Woods, as well as the game itself, has stuck with me and continues to inspire me, as well as to Katherine Isbister for her talk on tangible gaming for LARPs, which helped remind me of the reasons that I am proud to be a part of the human-computer interaction community).

My very sincere but also very anonymous gratitude goes to all my study participants in the works presented here and in other studies that I have conducted during my time here. I have been so encouraged, inspired and touched by the ways in which people have participated in my studies, and have never been able to truly do justice to my participants' engagement in those studies.

Finally, I'm grateful for the financial support of the National Science Foundation's Graduate Research Fellowship Program, as well the CMU's Graduate Student Small Project Help (GuSH) funds, and to Suzie Laurich-McIntyre for organizing, facilitating and guiding me through the logistics of both of these funding sources.

Again, I could go on (and on, and on), but I suppose at some point I'll need to present the research. To summarize: thank you to EVERYONE who has helped me

along the way, whether your name appears here or not. As I end this section of the dissertation and this chapter of my life and move onto the next, I do hope you'll remain in the picture and that I can return even the smallest portion of the support you have given me.

Contents

Preface: A Reflection on my PhD Journey

<div align="right">1</div>

When I first came to the Human-Computer Interaction Institute (HCII) at Carnegie Mellon University, I probably should have had a better idea of what human-computer interaction (HCI) research was. It was perhaps not a promising sign for a budding researcher that I entered into the program not having done sufficient research into the program itself, or into human-computer interaction research, more generally.[1]

1.1 Making Friends with Robots

In my first several months in the Human-Computer Interaction Institute (HCII), I held a series of meetings with obliging HCII faculty where I pitched different ideas of what I thought might constitute human-computer interaction (HCI) research. I distinctly remember one idea that I tried out on multiple faculty as well as friends outside of HCI. The idea centered around setting up researcher-participant dyads that communicate anonymously or pseudonymously online, wherein either the participant is made to believe that the other is either a robot, or another person. The two are meant to continue communicating, perhaps over several months, until the participant feels a kinship with the person/robot (researcher), by some metric. At this point, I proposed that the researcher would reveal to the participant that they had been conversing with either a person or a robot all along, in direct upheaval of what they had originally been told. My plan was then to essentially ask the participant, "So uh, how do you feel about this?" The reactions that I received to this proposal, needless to say, generally ranged from perplexed to skeptical to (politely) dismissive.

---

[1] I recently found a handwritten list I had made during the time I was considering different options, including the PhD. In it, I enumerated pros and cons for every job or position I had held from ages 16-26. A con written for one position was that I "hated research." An underlined pro for starting a PhD was that I would get a free bus pass. I do not always employ very logical prioritization strategies in making major life decisions.

But I must add: I also had other things on the list, including desires to return to the Carnegie Mellon community, where I had previously studied public policy and management and enjoyed very positive social and educational experiences; to learn; to collaborate; to gain new technical skills; to create a future for myself that would allow for the space, time and financial freedom to explore creative pursuits; and to expose myself to new people and ideas.

However, in the past year, this inchoate idea has been returning to my mind again and again, not because I find the initial idea all that compelling (and I certainly don't feel very comfortable actually implementing such a study design for obvious ethical reasons), but because of what it represents to me. It took me essentially the entire five years of my PhD, but I've come to realize consistent patterns in the ways in which I attempted (and often failed) to approach research, and/or stubbornly refused to deem certain pursuits as research.

For example, my newbie, naive notion of HCI research contains elements that I eventually integrated into the *research through evocative play* approach that I present in this dissertation. Namely, the idea I proposed brings to the forefront power ambiguity and the role of the researcher, whom I purposefully positioned in an ethically-ambiguous role, part of which stemmed from a personal discomfort and ambivalence around assuming the role of researcher and the power that entails. In my proposed idea, the researcher is as much an active and subjective participant in the study as the participant. I also recognize in the study plan a curiosity around anonymous and pseudonymous communication, the intention to push the boundaries of and complicating such communications while still operating within their confines, as well as a general interest in simultaneously encouraging and questioning digitally-facilitated social interactions, all of which have stuck with me through the years. In the past year, I've been reflecting on ways in which such a study design (after many alterations and substantial re-framing), *could*, in fact be HCI research, and how this and other "non-research" projects I've considered or pursued during my time in the HCII have implicitly and explicitly influenced my research path.

1.2 The Alien Capture Project

For example, in an early exploratory investigation, which I call the "alien capture project," I role-played a "Friendly Alien" via email (friendly.secret.alien@gmail.com) over the course of ten days with a small number of participants, most of whom I knew personally. I had participants first send me (to my CMU email email account) a photo of themselves in response to the following message.

> **Email subject line:** this is not a pilot study
> **Email content:** Imagine you have been captured by aliens. The aliens throw you in an isolated chamber and snap a photo of you from a hidden camera inside the chamber. Respond to this message with the picture.

All other communications, I conducted as the alien. For example (for those who were assigned to a "brainwash-resistant" group), I first greeted them with the strange, and

Fig. 1.1: Friendly Alien, as represented by an alien mitten (mitten courtesy of Qian Yang).

rudimentary photographic self-presentation of Figure 1.1, and a message explaining how aliens had taken over the earth and, recognizing the connective power of human narratives, had stolen all of humanity's stories through mind control. I told the participants they were one of the select few humans who were immune to those powers, and that they thereby had the power to free other humans through the use of stories (both telling their own personal story, and creating stories for others). (*"I know this is a lot to absorb at once, and you are probably feeling overwhelmed. But we don't have much time. But please, do hurry! I will be able to tell you more soon."*)

Permission for photo-sharing was likewise obtained through the alien.

> **Subject:** psst!
> **Content:** Over here, can you hear me? Put your ear against the wall.
>
> I am sorry you have received this fate, but I hope in the future to figure out a way to rescue you.
>
> To do that, I may have to share the photo from the chamber and your name with another human.
>
> **Do I have your permission?** I cannot promise anything at this point, but know that no matter what happens, there are creatures out there fighting for you.
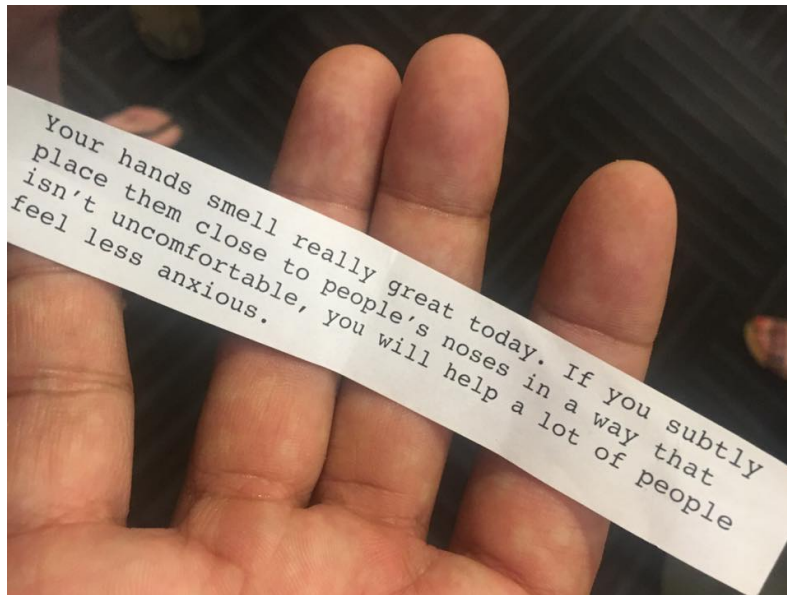
Fig. 1.2: An example of a fortune received from the vending machine

I then assigned participants to a mission to unlock another participant's brain through a storytelling activity. I controlled all communications between the two for the remainder of the pilot. I found all of participants' responses fascinating, from the photos and stories they submitted, to the ways in which they reacted to the stories received from others, to the ways in which they interacted with the alien, including wariness (*"I'm skeptical of you, 'Friendly,' but I have completed your task. Though I dare not hope."*) and earnest appreciation (*"Of course! Please help me, whoever you are. Thank you!"*). However, I settled on photographic self-presentation and semi-fictionalized self-presentation as the most "research-y" contributions of the pilot study, which led me down a path of exploring the effects of semi-fictionalized self-presentation (e.g. [114]). I felt particularly stuck because what had compelled me with the alien capture pilot was not so much a specific interest in online self-presentation and what was revealed therein, but rather, the use of a research approach that centered around playing with and manipulating that context as a researcher, and studying the impact and implications of that play. I just didn't know how to make such an approach "research-acceptable" at that time.

1.3 The Monster Vending Machine

During this period of stagnancy, I bought a $40 mechanical (non-electric) vending machine from Craigslist, updated it with some craft materials to look like a monster, and filled it with snacks with random little fortunes I wrote, many of which I now find trite and cringe-worthy, but all of which involved some level of ceding to an external, anonymous power (see Figure 1.2).

Fig. 1.3: The monster vending machine.

I also placed an extra little box decorated like a mini-monster on top of the machine's head, occasionally filled the mini-monster with free snacks, and periodically checked to see if anyone had taken any (curiously, they never did). I wanted the machine itself to inspire skepticism, to encourage the suspension of disbelief, and to pose as a position of authority or power, but was agnostic about how I wanted people to respond. In my ultimate (unrealized) vision of the vending machine, I wanted the machine to go in some dark hidden basement where people would encounter the machine serendipitously, and where I would never know if anyone had even interacted with the machine at all until I went in months later to check on the supply levels. I didn't actually go so far as to smuggle the machine into an abandoned basement somewhere, but still, I was very resistant to conducting formal follow-up inquiry on people's interactions with the machine (e.g., post-experience qualitative interviews, which I had conducted for the alien capture pilot). Essentially, I resisted any attempts to transform the vending machine project into a research project, because I feared that doing so would obviate the playfulness of the project's design and implementation. As a result, the vending machine now perches at the top of a high bookshelf in the office that I unofficially share with my colleagues Xu Wang, Anhong Guo, and Gierad Laput, empty and unused.

1.4 Beginning to Recognize Myself as Ludic

It also took me five long years and countless firm declarations to potential collaborators that "I do not like games or play," to recognize how central play is to how I approach curiosity and exploration. As I reflect back on these "non-research" projects,

I most vividly remember the pride and satisfaction the processes of facilitating those projects brought, even when I felt (both at the time and in retrospect) that the premises themselves and the specific elements I included were quite puerile or even unethical. I only realize now that I was experiencing the joy of play, and more specifically, the pride I took in a resistant form of play. My stubbornness in refusing to translate these projects into concrete research was in part a resistance against the conventions and assumptions of research (or at least, what I viewed as research at the time), and a critique of the impartial, unassailable role of the researcher in that process.

On a personal level, *research through evocative play* thus represents an attempt to reconcile (a) my resistance to and critiques of research and the attendant powers of the researcher, and (b) a personal inclination towards play as a means to learn, explore and gain insights. I have spent my entire PhD and then some (including before I officially started the program) vacillating between quitting and staying in the program,[2] and in my own self-constructs, have been adamant in defining myself as decidedly *not* a researcher, despite how I might feel obligated to present myself in papers and presentations. Ask me how I came to join the program, and I will (as I have already done in this section's footnotes) trivialize all those reasons, focusing on the frivolous so as to avoid identifying myself as a researcher. My day-to-day experience in the PhD has often been far from playful and my attitude far from lusory.[3] Nonetheless, I have attempted to cement the notion both to myself and to others that my researcher self is not my "real" self or identity; a ludic self, for all intents and purposes, but woefully lacking the playful attitude that ought to accompany such a self.

Despite my obstinate resistance to research and "being a researcher," the research path has taught me a lot. In addition to acquainting me with a supportive and stimulating community of students and professors that have enriched my life in ways that extend far beyond career or academic concerns, pursuing a research path has helped me discover and decipher aspects of myself and how I relate to and strive to understand the world around me. I do not see myself as continuing formal research once I complete this PhD, at least until I've experienced an interlude of contemplation, exploration, and both success and failure in a less formally academic or research-based context. But as I reflect on my time in the program, there are so many aspects and lessons that I want to carry forward as I continue to figure out life. One of these is the spirit of *research through evocative play*, and of the idea of playing with power, context, trust, and anonymity as a way to generate new insights, gain understanding, and discover new questions. Whether I choose to present myself and

---

[2]Causing me, my advisors, and others around me undue distress.

[3]See the Introduction and Background chapters for a more detailed definition and discussion of the lusory attitude.

identify formally as a researcher or not, I plan to continue to play and explore as I enter a new stage of my life.

Introduction

> *In disrupting the normal state of affairs by being playful, we can go beyond fun when we appropriate a context with the intention of playing with and within it. And in that move, we reveal the inner workings of the context that we inhabit.*
>
> — **Miguel Sicart [200]**
> (Play theorist)

2.1 This Is Play, Now: Drawing Circles

Fig. 2.1: A (magic?) circle on the sidewalk

The metaphor of the magic circle, first introduced by John Huizinga [108] and coined and extended as a game design concept by Salen and Zimmerman [189], is described at its most basic as the place where games and play takes place. In entering the magic circle of play, a new reality is created and is subsequently circumscribed by the rules of that specific form of play.

Part of the construction of such a magic circle, then, is the knowledge that a context is defined as a place of play, and that there exists a set of rules that circumscribe that play. However, it does not follow that we can go about drawing magic circles wherever we please simply by declaring a context a play space, and laying out the rules of behavior therein. If we exclaim, "This is play, now!" a new circle does *not* magically appear; saying so does not make it so. Rather, the wizardry of a given circle depends on those who enter the circle *perceiving* and *accepting* it as magical. The magic of the circle depends on players voluntarily entering the space. Suits [207] describes the voluntary nature of play through the concept of "lusory attitude;" when entering into a magic circle, players willingly and of their own volition adopt an attitude specific to that bounded circle that is separate from their ordinary selves.

Nonetheless, there is little to stop me from drawing circles, and if I so choose, *calling* them magical anyway, regardless of whether those who enter the circles see them as magical or not. To take a very literal and physical example, let's suppose I go outside and draw a circle on the sidewalk. Let's also suppose the circle I draw is located on a high traffic street in New York city, and even more specifically, right outside of Trump Tower. I make a sign pointing to the circle that says "Magic Circle," and write a rule in chalk inside the circle: "Anyone who enters the circle must do a silly dance expressing how they feel about their current environment before proceeding to the next sidewalk square." I sit myself down on a nearby stoop, and wait for the fun to begin.

However, because the circle is on a busy street, most people enter the circle without even realizing the circle's existence. Even when people do see it, many still choose to ignore it. So to make sure people know the magic circle is there, I ask people to step inside, verbally repeating the rule of play.

I enjoy a bit more success. A few people find it a quirky activity and obligingly (maybe even joyfully) do a jig. Others take pity on me and my strange request and shimmy a bit as an act of compassion. Others accidentally make eye contact with me, and then feel obligated, in the way that some humans feel obligated to engage after eye contact has been made, to perform a lackluster two-step, cringing in embarrassment all the while. Some people seems uncertain as to how to act, their eyes jumping from me, to Trump Tower, to the small crowd of spectators that have gathered. Still others sullenly shoot me a dirty look and continue on their way. A very angry person enters the circle and spits in my direction. Another very angry person enters the circle and spits in the direction of Trump Tower, spins three times, and exits the circle. A smattering of people clap in response, followed by an equal number of boos. A new person enters, and hops back and forth while repeatedly bowing to the tower. Serendipitously, two people enter at the same time. The first person is on a conference call, patently unaware of their surroundings.

The second person has seen the mini spectacle I created from a block a way and very purposefully enters the circle, links arms with the (appalled) person on the conference call, and do-si-dos.

For many of the participants who entered the circle I drew, I would argue that the circle is but a circle; not a magical one. Yet for me as the drawer, orchestrator, and observer, the circle is always magical. *My* magic circle is the act of drawing the circle, and watching what happens when passersby in New York city enter into this purportedly magic circle. Within the context of a New York city sidewalk in front of Trump Tower, I get to observe for whom the circle is magical and in what ways; what are the implications when the magic of the circle, or the circle itself, is visible to some and not others (as in the do-si-do example); the ways in which people choose to rail against the circle itself and/or its alleged magic. In observing how sidewalk passersby in New York react to a specific activity declared as play overlaid on part of their daily life, I gain new insights into the dynamics of the context (a sidewalk in New York) and how people in New York relate to that sidewalk, to one another, to power embedded in that context (both as represented by me, the researcher-designer, and the contextual presence of Trump Tower) and to notions of playfulness within that context (which in turn, further reveals their relationship to that context). I'm interested in understanding the specific ways in which they choose to play or not play within the circle I have drawn not because I want to understand my own circle, but because I want to understand the people, this NYC sidewalk context and its attendant power relationships and ambiguities, and the implications of specific forms of play in this context. In this way, I consider my approach (inspired by and in conversation with research through design and other research methods related to design within human-computer interaction, as well as play theory and research) as *research through evocative play*.

2.2 Using Play-Based Methods to Study Power Dynamics (in Digital Contexts)

In *research through evocative play*, the circles of play themselves are *declared* as playful, but research participants may not necessarily see, interpret, or respond to the play declarations with lusory attitudes. I use the term "evocative" to describe the (declared) play of the method I propose because I want to emphasize that this approach can *evoke* (draw out) reflections, criticisms, and novel insights into the power dynamics and nuances of a context. By forcing participants to make choices about whether and how to play in a context, *research through evocative play* evokes reactions to and reflections on the context and power dynamics therein that might be occluded by say, observational or self-report methods (see Chapter 7 for

a more detailed comparison of *research through evocative play* to other research methods). By presenting circles of declared play that are uncertain in their magic, *research through evocative play* asks participants to grapple with the ambiguities (including ambiguities of power) of the context in which they are situated, evoking responses that can help researchers to better understand the context. As a Human-Computer Interaction (HCI) researcher, I focus my attention on digital contexts of power ambiguity and imbalance, with a sub-focus on how varying degrees of anonymity and digital communication affordances and limitations operate within and influence power dynamics. However, as my very literal drawing circles on the sidewalk example implies, *research through evocative play* methods can be applied to non-digital contexts, as well.

In my research, I ask not what is play itself, but what declaring a space as play reveals about a context, the individuals inhabiting the context, and their relationship to the context and one another by means of the behaviors and attitudes that emerge (both playful and non-playful) in response to the declaration of play. My research identifies, analyzes and considers the implications of assuming lusory attitudes in specific contexts, but does not presuppose a lusory attitude among participants, nor does the success of *research through evocative play* depend on engendering these lusory attitudes.

As the researcher, I am the designer of declared play in *research through evocative play*, and I am simultaneously an active participant in the study. In my work using *research through evocative play* methods, and influenced by research and theory in design (especially research through design, critical design and ludic design) and play (see Chapter 3), self-reflection on the part of the researcher is a crucial part of the research process. This includes acknowledgement of one's own biases and assumptions that affect the study design and declared play itself, which in turn impacts participants' engagement and reflection.

2.3 Research Contribution

My work explores how overlaying "circles" (not necessarily magic) of declared play onto behaviors or digital contexts can elicit critical reflection on a particular context or set of behaviors, providing novel insights into and access to participants' opinions about and relationships to the original context or set of behaviors. I have done this both by studying online communities that use design cues to re-frame a set of behaviors as playful within a redefined context, as well as designing and overlaying cues for play in digital contexts where power is especially imbalanced or ambiguous, and means of communication among context inhabitants are limited.

2.3.1 Contribution 1 (Methodological): *Research through Evocative Play*

My research contributes to the field of Human-Computer Interaction by proposing the use of play not as a design end, but as a research tool to elicit reflection on and criticism of power dynamics within a particular digital domain, revealing nuances of the context and inhabitants' relationship with that context. As an HCI researcher, I apply the *research through evocative play* approach to digital (specifically, online) contexts, wherein at least some degree of anonymity is common practice and subsequently influences power dynamics, and where the design of technology platforms also can directly influence power dynamics in a space. Following, my work in using *research through evocative play* models how play-based methods can be used to draw out and better understand power dynamics of varying forms in contexts in which— like many digital contexts— the technology designs may inhibit, alter, or suppress communications and attendant power dynamics among the context's inhabitants. In my work, I show how the *research through evocative play* approach can be particularly relevant to revealing negative and critical sentiments related to power inequities and ambiguities that might otherwise be suppressed. Understanding power dynamics and revealing nuances of power that might otherwise be occluded, then, is central to the *research through evocative play* approach.

In this way, I propose that *research through evocative play* can be an effective empirical research method to study and elucidate different forms of power dynamics and ambiguities that can benefit many researchers in HCI, especially those who are especially interested in dynamics of power (and attendant complexities present in such dynamics). Many conscientious HCI researchers have explored and critiqued diversity, inclusion, and equity (and lack thereof) both within contexts that are relevant to HCI, as well as within the guiding methods, approaches, and assumptions that underpin HCI research practices and directions. To give just a few scant examples, HCI researchers have studied power, oppression, marginalization and diversity [70] from various different perspectives and focal points, including the ethics of conducting different forms of HCI research [65], HCI for development and under-served populations [51], algorithmic fairness and bias [224], and power and oppression as it relates to: gender and sexuality [123]; race [191, 93, 212]; class [190]; and other forms of identity-based diversity such as neuro- and physical-diversity [158, 211, 48, 115]. *Research through evocative play* proffers researchers of power dynamics a study approach that uses play-based methods to further unpack and elucidate aspects of those power dynamics. My own work that I present in this dissertation, for example, concerns power with respect to identity axes such as gender, sexuality, and race (Chapter 4), labor operations and relationships in online

crowd work (Chapter 5), and tensions between capitalist production, individual privacy and autonomy, and leisure (Chapter 6).

My work also extends research in play theory by proposing novel applications for play methods in research contexts, but the use of *research through evocative play* certainly need not be restricted to HCI or play scholars. For example, as my original simplified example (drawing literal circles on the sidewalk) suggests, *research through evocative play* can be used to study non-digital, collocated contexts. Thus, researchers interested in the study of power dynamics from multiple academic disciplines— from play theory and research to psychology to public policy to religious studies to race studies, et cetera— could benefit from applying and adapting *research through evocative play* to their work. (For more guidance on applying *research through evocative play* methods, see Chapter 7).

2.3.2   Contribution 2 (Theory and Practice): Lessons Learned about Specific Cues for Play

In *research through evocative play*, play is not the end, such that engaging in successful play is not the goal or metric of success of a given research project. However, as a side product of my research, I've learned some critical lessons about specific cues for play that can benefit researchers and practitioners of play. In this way, my research provides insights into how specific cues for play can be effective (or ineffective) in engaging individuals in play and/or eliciting critical reflection through play. I chose to narrow the scope of the of my *research through evocative play* designs so as to iteratively gain a firmer grasp on how these cues might operate and affect participant engagement within and across different contexts. Specifically, I focus on the following play cues in my work:

1. Direct declarations of play ("This is play, now").

2. Persona embodiment and encouraging of perspective-taking

3. Encouragement of mocking styles of humor

4. The creation and sharing of artifacts, especially through anonymous or pseudony-mous channels

In Chapter 3, I discuss in more detail how I chose these four specific cues, drawing from prior literature in play theory, psychology, and human-computer interaction. I also present the results of engaging participants using these four play cues in

the chapters to follow. I'll note that where each of the studies I present use direct declarations of play, depending upon the context and the specific of the study design, the following three cues (persona embodiment and encouragement of perspective-taking; encouragement of mocking styles of humor; the creation and sharing of artifacts through anonymous or pseudonymous channels) are present as either (a) designed aspects of the study protocol, (b) precursors to such designed aspects, or (c) emergent forms of behavior and engagement among participants.

2.4 Document Structure

I present research I conducted in three different contexts: the subreddit RoastMe, Amazon Mechanical Turk, and YouTube video advertising. For each of the three chapters below, I summarize the contexts, and briefly describe how each chapter relates to the development of the *research through evocative play* approach.

2.4.1 Chapter 3. Background

In this chapter, I present the research from design, play, HCI, and other fields that informs and inspires my approach.

2.4.2 Chapter 4. RoastMe: The Implications of Casting Harsh Criticism and Online Photographic Self-Presentation as Play

Context: Reddit and RoastMe

A sub-community (subreddit) of Reddit.com, r/RoastMe (or more simply, RoastMe) has participants post photos of themselves to the community to be "roasted," or harshly ridiculed with humorous intentions, by other participants in the subreddit. Despite the presence of language that contains markers of racism, sexism, homophobia, violence, and other ostensibly hateful forms of speech, r/RoastMe defines itself very specifically as a community centered around "comedy, not hate." Reddit is already a community of play, but in RoastMe, a specific type of play is defined and encouraged. Within RoastMe, a circle is drawn, and sets of behaviors that would not be considered playful in some other offline and online contexts— online photographic self-presentation and harsh language— are brought inside the circle, such that players' engagement with the elements in the circle also reflect their relationships to those elements as they exist beyond the circle (for example, online

photographic self-presentation in other social media contexts, such as Facebook and Instagram, and the use of harsh language on the internet, more broadly). The community includes direct declarations of play ("comedy, not hate" = "this is play, not hate"), and mocking styles of humor; the heavy use of perspective taking is also an emergent behavior in the community.

Communications on Reddit are pseudonymous, and dyadic exchanges are discouraged by the platform's user interface, such that communication between subreddit participants has limitations. The existence of "throwaway" accounts compounds the communication limitations of Reddit. Throwaway accounts are created for short-term use only, and deleted or abandoned afterwards, as Reddit puts no limitations on the number of accounts that can be tied to an individual email address or identity. Such throwaway account behaviors are especially common on RoastMe, wherein those posting photos of themselves may want their other communications on other subreddits to remain fully anonymous, untied from their visual-physical identities.

Moreover, power dynamics within RoastMe, especially pertaining to those who post photos of themselves to be roasted, are ambiguous. As is revealed by my study of this existing online community that re-frames a set of behaviors as play, the extent to which participants voluntarily consent to and participate in the play circumscribed by RoastMe is often unclear, as are the boundaries of acceptable humor.

RoastMe and *Research through Evocative Play*

With RoastMe, I present a mixed-methods study of an existing online space where play is declared. In this way, I see my study of RoastMe as a precursor to and inspiration for the *research through evocative play* approach I explore and refine with Turker Tales and YouMercials (Chapters 5 and 6). Play on RoastMe is declared directly via the directive "comedy, not hate," thereby casting both harsh humor and online photographic self-presentation as play, though the two do not always operate as play in other online contexts. In this way, RoastMe also directly encourages a mocking style of humor. Moreover, as I present in my findings, emergent behaviors in RoastMe indicate a heavy use of perspective-taking. In my research on RoastMe, I conduct qualitative interviews with RoastMe participants (N=14) and also analyze a large dataset of RoastMe images, posts, and comments.

Where r/RoastMe takes elements present in other digital contexts and places them inside their declared play circle, for the other two projects I present (Turker Tales and YouMercials), I injected declarations and cues for play into existing contexts. Inspired by but deviating from existing communities like RoastMe that transport

elements and behaviors found in other contexts into a new circle, with the latter two projects I present, I essentially draw circles atop the original contexts.

2.4.3 Chapter 5. Turker Tales: Revealing Workers and their Relationships to Amazon Mechanical Turk through Play Instructions

Context: Amazon Mechanical Turk

Amazon Mechanical Turk is a crowd platform often criticized for its maltreatment of the crowd workers that are its backbone. In Amazon Mechanical Turk (which I will often shorten to MTurk), there are thus unequal power dynamics, especially related to low wages earned by crowd workers. Moreover, communication is suppressed first between workers, where platform-supported communication is non-existent. In addition, communication between workers and requesters is also limited, and workers have little venues to self-advocate when they feel they are unfairly treated by a requester.

Turker Tales and *Research through Evocative Play*

In this chapter, I present Turker Tales, a Google Chrome extension implemented with 171 participant on Amazon Mechanical Turk that encourages participants to create short identity-based scenarios for other workers to imagine as they complete HITs (perform work) on MTurk. Turker Tales directly declares its activity as play. Further, Turker Tales also features persona embodiment and perspective-taking through its "imagine yourself as" prompts, and supports the creation and sharing of artifacts (tales) while maintaining worker anonymity.

2.4.4 Chapter 6. YouMercials: Exploring Relationships to YouTube Video Advertising through Play Directives

Context: YouTube Video Advertisements

On YouTube, users are subject to the power of a capitalist system in that they are targeted with advertisements, thereby contributing revenue to a large, corporate-owned platform. Power dynamics are complex in that users have the power to generate and influence content on YouTube, but that power is simultaneously leveraged by

YouTube to ultimately benefit the corporation financially. Users may have mixed advertisements towards advertising on YouTube; for example, while some may revile advertisement-targeting as a form of oppression, others may view advertisements ambivalently or even positively.

In the specific context of viewing video advertisements on YouTube, communication is also suppressed in that there is no way for users to directly communicate with other advertisement viewers.

YouMercials and *Research through Evocative Play*

In this chapter, I present my design of YouMercials, a *research through evocative play* concept implemented as both a controlled survey experiment (N = 156) and an "in-the-wild" field deployment (N = 57; subset of N = 156). Drawing on and combing elements from both RoastMe and Turker Tales, YouMercials declares play with YouTube advertisements by allowing participants to create, share and view "YouMercials," which are content overlays for YouTube ads; these artifacts are created and shared anonymously (or semi-anonymously, in the case of audio-dubbed YouMercials). They can be produced in one of two forms: "imagine yourself as" YouMercials, and audio-dubbed YouMercials. Through "imagine yourself as" YouMercials, in which participants create scenarios for other viewers to imagine themselves in while watching the advertisements (very similar to the design of Turker Tales), the design directly encourages perspective-taking.

In YouMercials, I also add another layer of perspective-taking by varying the conditions to which participants are assigned. Either they are primed to think of YouMercials from the perspective of helping advertisers improve content, or from the perspective of taking power back from advertisers. Moreover, because YouMercials are created for and shared with other users, perspective-taking also becomes an inherent part of the creation process for both "imagine yourself as" and audio-dubbed YouMercials. In my study design, I also randomly tell a subset of participants to specifically use a roasting style of humor, extending my observation of the use of harsh humor as a play cue to study power dynamics (building especially on my research of RoastMe, as well as that of Turker Tales).

2.4.5 Chapter 7. A Guide to Applying *Research through Evocative Play* Methods to Your Own Projects

In this chapter, I provide (what is meant to be) a user-friendly researcher-practitioner's guide to applying *research through evocative play* to one's own projects and research. The aim of this chapter is to save future researchers the several years of struggle it took me to develop and apply the *research through evocative play* approach to my own work. In this chapter, I therefore lay out who is "qualified" to use *research through evocative play* (hint: everyone), recommended background reading, identify important steps to take in designing a *research through evocative play* study or project, and make other attempts to compile and condense lessons I've learned about conducting research using *research through evocative play*.

2.4.6 Chapter 8. Conclusion

In the Conclusion, I summarize the findings across my work with RoastMe, Turker Tales, and YouMercials, and propose ways in which to build upon and extend the research I present in this dissertation.

Background

3

My research draws from and builds upon work in design research in human-computer interaction (HCI), play research and theory, and context-specific research related to (a) Reddit and online self-presentation, (b) crowd work, and (c) online advertising.

3.1 Design Research in HCI

I want to first begin with work in design in HCI research— specifically, work in research through design, critical design, and ludic design— which has greatly influenced my perspectives on research in human-computer interaction, and directly inspired the *research through evocative play* approach I take in my work. As in design, I see the role of *research through evocative play* as intending to study and address "wicked problems" [24, 182], or problems to which there is not a clear scientific or engineering solution. Specifically, *research through evocative play* intends to illuminate and provide insights into complex power dynamics.

3.1.1 Research through Design

Research *through* design, especially as forwarded by HCI scholars Zimmerman, Forlizzi and colleagues [232, 233, 231], separates itself from design practice and research *for* design in that artifacts created as part of research through design are not intended to be commercially successful. Rather, research through design as a speculative and exploratory practice [77] creates artifacts intended to be "carefully crafted questions" that "stimulate discourse around a topic by challenging the status quo" [232]. Similarly, my concept of *research through evocative play* uses play declarations as a means to stimulate discussion and critical reflection; the goal is not to create play interventions or games that are commercially viable, or even necessarily enjoyable to those who engage with the form of designed play. In this way, *research through evocative play* can be seen as quite distinct from research through design. Where research through design seeks to create not the commercially successful thing, but instead the "right" thing that moves us to a preferred future state [233, 231], *research through evocative play* might specifically create the "wrong," or at least the "ambivalent" or "ambiguous" thing.

3.1.2 Critical Design

Research through evocative play is inspired equally by research through design as it is by critical design in HCI, which I will view loosely as subsuming other related forms of design research such as oppositional design, adversarial design, and reflective design. Note that in discussing critical design, I refer to the practice as described in HCI research such as [175], not Dunne's more specific school of critical design [57]. Through critical design, researchers design for provocativeness, and cultivate deep relationships between researchers and research participants [9]. Under adversarial and oppositional design, artifacts "work by not quite working" and through the study design, operate as hypothetical products you imagine using as part of the study design rather than actually use [161, 174, 175]. *Research through evocative play* artifacts may similarly be designed to "work by not quite working," asking participants to take part in play that they may view ambivalently or negatively, and not necessarily expecting or hoping that participants actively engage and take on lusory attitudes. For example, the declared play experiences I designed for Turker Tales (for crowd workers on Amazon Mechanical Turk) and YouMercials (for viewers of YouTube advertisements), even when implemented for short time periods "in the wild," were never intended to function as long term, sustainable, or desirable play interventions representing preferred future states.

3.1.3 Ludic Design and "Probology"

In proposing *research through evocative play*, I also draw from ludic design and "probology" as described by Gaver et al. [79, 78]. In ludic design, Gaver et al. advocate designing for ludic (playful) rather than utilitarian activities, and describe "an attitude of engagement in the exploration and production of meaning" as central to notions of ludic activities [79]. Ludic design should not be "for" specific purposes, but instead should allow participants to explore a range of options and meanings. *Research through evocative play* and ludic design share a focus on encouraging playful engagement as a design and research method, but *research through evocative play* study designs may choose to explicitly state a "for" and then observe how that explicit "for" is interpreted, reacted to, complied with and subverted by participants. Moreover, unlike ludic design, *research through evocative play* does not aim to enable play directly, but rather, uses play declarations as a research tool. Also related to ludic design is Gaver et al.'s description of "probology" [78]. Cultural probes are "collections of evocative tasks meant to elicit inspirational responses from people—not comprehensive information about them, but fragmentary clues about their lives and thoughts" [78]. Under the concept of "probology," Gaver et al. advocate using cultural probes to encourage "subjective engagement, empathetic interpretation, and a pervasive sense of uncertainty," thereby positioning probology as a critical approach

that thrives from uncertainty. In a similar manner, *research through evocative play* observes and studies the uncertainties and ambiguities revealed, highlighted, and complicated by the engagement with play declarations to better understand power dynamics within a specific context, participants' relationship with that context and one another, and notions of play within that context. Further, where most research methodologies seek to minimize the role and bias of the researcher, the probes approach "purposely seeks to embrace it" [78]. Likewise, *research through evocative play* specifically views the researcher as an active participant in the study design. In *research through evocative play*, participants' uncertainty regarding the intent of the form of play they are asked to engage in directly draws the researcher and the role of researcher to the fore, as I will show in my work with Turker Tales and YouMercials. Rather than trying to minimize the Hawthorne effect and demand characteristics [165, 151], participants' subjective interpretations of the researchers' intents become integrated into the study design and analysis.

In this way, research through design, critical design, and ludic design in HCI all serve as fundamental inspirations for the *research through evocative play* approach.

3.2 Research in Play

I define *research through evocative play* as using play directives (including cues for play) to study power dynamics within and elicit critical reflection on a context. While the success of a *research through evocative play* study design depend on neither the acceptance of the circle circumscribing the study's artifact design as "magic" (play), nor the adoption of a lusory attitude among participants, assessing such acceptance and attitudes becomes an integral and reflective aspect of the research analysis. I therefore turn to play theory to understand how play reveals context, and to consider its application to declared play. In addition, I draw from research in play and capitalism due to its particular relevance to my work with Turker Tales and YouMercials, and literature on specific forms of play to explain the forms of play most relevant to my work.

3.2.1 The Magic Circle and the Lusory Attitude

I opened the Introduction (Chapter 2 of this dissertation document with a short description of the magic circle concept; I return to a more detailed discussion here. As previously introduced, Huizinga first mentioned the phrase "magic circle" in his seminal play theory work, *Homo Ludens* [108], and the phrase was later adopted, formalized and extended by Salen and Zimmerman [189]. For Salen and Zimmerman, the magic circle is the time and space (physical or meta-physical) in

which play takes place. Entering into the circle happens once one begins engaging in play. Within the circle, "a new reality is created, defined by the rules of the game and inhabited by its players" [189]. Where Huizinga defines play as a space separate from one's ordinary existence, a realm unto itself, Salen and Zimmerman acknowledge that a magic circle may be seen as either closed or open, and they place more emphasis on the open, rather than the closed, aspects of game and play systems. A closed magic circle exists, in Huizinga's sense, as separate from reality, such that observing activity within a closed circle focuses on behaviors intrinsic to the play circumscribed. By contrast, an open system incorporates players' identities, experiences, and social relationships from outside the play world into the magic circle.

Since Salen and Zimmerman first articulated the concept of the magic circle, many other scholars have applied this concept to different game and play domains, and have also criticized and debated the concept. Criticisms often center on Salen and Zimmerman's use of the concept of a closed circle as "separate from the real world" [188]. For example, Taylor [210], Malaby [144], Copier [43, 44], Pargman and Jakobsson [166], Calleja [29] and Lammes [131] have all contested the notion that we can define clear divisions between play and ordinary life.

Supporters of the magic circle concept, however, far outnumber the critics. Many scholars continue to refer to and rely on the magic circle due to the usefulness of the metaphor as a shorthand. Some scholars have also extended and built upon the concept of the magic circle. For example, Montola [156, 155] re-envisions the magic circle as a "metaphor and ritualistic contract," and Juul [110] redefines it as a boundary that players negotiate, emphasizing the notion that play boundaries themselves are permeable. In my own understanding of the magic circle, I strongly favor the synthesis of many of these critiques and extensions as offered by Stenros [205], wherein the magic circle is "a special space with a porous boundary [that] is created though social negotiation" and where the focus is not just gameplay, but play, more broadly.

I bring up these debates both to acknowledge complications surrounding the use of the "magic circle" concept and also to clarify what I refer to when I reference the magic circle of play in discussing *research through evocative play*. Another crucial component of the magic circle concept is the "lusory attitude" as defined by Suits [207]. Suits defines this attitude as a novel, playful disposition that players assume when entering into the (bounded) space of play. Salen and Zimmerman directly reference the lusory attitude in their definition of the magic circle, and maintain that voluntary entry into the magic circle is a precondition of the magic circle's existence [188]. While the lusory attitude can be interpreted to refer specifically to a closed magic circle, in that the play self diverges and is separate from the "ordinary" self,

Stenros [205] encourages us to take a broader perspective, to view the adoption of a lusory attitude as a personal choice in which a participant accepts the rules of play. Under this view, Stenros puts forth that players can still have ulterior motives, but must "adapt an attitude where they take the rules seriously in order for the game [or more broadly, play] to take place" [205].

In my formulation of *research through evocative play*, the lusory attitude is not a prerequisite in that players need not fully voluntarily participate, nor must they take the rules seriously or fully abide by those rules of play. As a result, the circles created through declarations and cues for play under *research through evocative play* are not necessarily magic, as not all those who enter circle defined by *research through evocative play* necessarily accept the circle as magic or adopt a lusory attitude. Moreover, where the concept of a magic circle may be seen as closed (existing separate from "reality") or open (incorporating and reflecting elements of players lived experiences that they bring with them into the boundaries of play), the circles drawn in *research through evocative play* are magnifying glasses placed over the contexts in which they are drawn: the *research through evocative play* circle exists to illuminate power dynamics and other elements of that context.

At the same time, in *research through evocative play* we analyze whether, when, and how participants adopt lusory attitudes to gain insights into how participants view notions of play within a context, and subsequently, how they relate to that context and to one another. Understanding ways in which participants choose to engage in play within the declared play, and to what extent their play is subversive or resistant to elements of the play declarations or the context itself, are fundamental to the process of *research through evocative play*. In this manner, determining whether lusory attitudes are present in participants is integral to *research through evocative play*. Here, I draw upon Lazzaro's four keys to fun and emotion in games and play [135], which lays out four categories to capture what players like most about games and play. These include (1) the internal experience key, in which players enjoy their internal state during and after play; (2) the "hard fun" or challenge and strategy key, in which players derive enjoyment from problem-solving and overcoming adversity; (3) the "easy fun" or immersion key, in which players experience curiosity about and are drawn into elements of play; and (4) the social experience key, in which players engage with others either directly through the form of play (e.g. player competition and cooperation) or indirectly by viewing the play as performance or spectacle. In my work with RoastMe, Turker Tales, and YouMercials, Lazzaro's four keys of fun guide how I analyze the ways in which participants reflect on themselves; are challenged and rewarded by engagement in the declared play; are made curious about elements of the interactions or the structure; and interact with (directly or indirectly) other participants.

3.2.2 Play as Revealing of Context

Miguel Sicart speaks of play as disruptive and revealing of context. To return again to his quote,"In disrupting the normal state of affairs by being playful, we can go beyond fun when we appropriate a context with the intention of playing with and within it. And in that move, we reveal the inner working of the context that we inhabit," as well as "the seams of behaviors, technologies or situations that we take for granted" [200]. He speaks of playfulness as carnivalesque in that it is "an opening toward critique and satire, toward freedom in the context of mundane activities" [200]. Sicart's notions of play, then, are central to my approach to *research through evocative play*, and point to play as a way to open participants up to critique and satire of a context. For example, critique and satire are explicitly encouraged in r/RoastMe and I extend this in my work with YouMercials through an experimental condition encouraging a subset of participants to using "roasting-style" humor. I also seek and assess the spontaneous emergence of critique and satire in both Turker Tales and YouMercials as I determine how declaring spaces as play and injecting cues for play may differentially affect participants in a context.

In addition, I draw from Sicart's notions of the porousness of play. Sicart defines play spaces as not limited to games, but also inclusive of contexts such as lunch breaks and other "openings in time and space where play becomes possible" [200]. This especially fits into my work with Turker Tales, where the declared play is conducted alongside participants' daily crowd work activities. I also incorporate this notion into YouMercials, declaring video advertising interludes during YouTube watching, a time and space that is typically viewed as interruptive and annoying, as an opening for play.

3.2.3 Dark Patterns

Sicart and others have also written about dark patterns and abusive game design, which disrupt the usual assumption that game designers are advocates for players and aligned with players [229, 221, 199]. Games (and play) are understood as "systems of power in which subjects become voluntarily subordinate to a network of processes, actions, rewards, and values that define what actions are valid, valuable and socially recognized" [221]. In this way, the relationship between the game designers and the players becomes central; in playing (or resisting play), players enter into a personal dialogue with the designers. Here, we can recognize parallels with critical design and "probology" [9, 78], wherein the relationship between designers and participants is personal, central, and recognized as inherently subjective. In a similar manner, the researcher-designer in *research through evocative play* plays uses their power to play with participants and manipulate the context. In *research through evocative*

play, participants' assumptions about the affiliations and intents of the researcher in designing and encouraging the declared play thereby become a central area of inquiry.

3.2.4 Play and Capitalism

With both Turker Tales and YouMercials, I apply play declarations to contexts that involve capitalist systems in which participants either work directly for a large corporation (Amazon Mechanical Turk), or participants otherwise produce value for a large corporation (YouTube, wherein viewers contribute to advertising revenue and, through their liking patterns and comments, feed the algorithms that operate the platform). In my research, I question the implications of declaring play within capitalist systems, drawing on DeWinter et al. [53], who criticize the use of gamification and play to serve capitalist systems. I also pull from the notion that ostensible "play" (voluntary leisure) activities online often serve capitalist interests, rendering the internet both a "playground" and a "factory" of exploitative digital labor [192]. Thus, especially with both Turker Tales and YouMercials, I study how participants interpret and respond to the play declarations, and how they perceive my role as researcher-designer in relation to the larger capitalist system. In so doing, I critically consider the ethics and implications of incorporating any kind of play design within a capitalist system, *including* play intended to understand, empower, uplift, or resist against injurious aspects of that system.

3.2.5 Specific Forms of Play

The three projects I present in this dissertation make play declarations, and also directly instruct participants to engage in at least two of the following: perspective taking; mocking styles of humor; and the anonymous or semi-anonymous creation and sharing of artifacts with other participants. Here, I draw from related literature (primarily in play theory and/or social psychology) to explain the choice of these specific forms of play cues.

Play and Perspective Taking

According to Huizinga and Suits [108, 207], in any form of play, there is a shift in players' identities as they take on the lusory attitude and become their ludic selves. Sutton-Smith too, in defining the "rhetorics of identity," speaks of play as being about the "ontology of being a player and the dream that that sustains" [208]. Through perspective- or experience-taking, participants can achieve high levels of

flow and enjoyment and change their beliefs about others [120, 16], and narrative-based play can serve as a means of increasing social connection and feelings of belongingness [124, 75]. In this way, encouraging perspective taking through short narrative-based activities both draw attention to and manipulate (for the purposes of further reflection) participants' relationships with others in the context. Moreover, in contexts in which communication with other inhabitants of that context (other pseudonymous Redditors, other Turkers, other YouTube viewers) are limited by anonymity / pseudonomity as well as the constraints of the platform, encouraging perspective taking play could allow for indirect engagement with, reflection on, and (for the researcher) observation of participants' relationships to and perspectives about their peers in the context.

Where in RoastMe, perspective taking naturally emerged as part of the roasting process (see Chapter 4), in Turker Tales and YouMercials, I specifically encouraged participants to engage in a form of perspective taking that involves the creation of short narratives. In these short narratives, participants create a character or persona in a scenario within which other participants in the system and platform should imagine themselves.

Play through Harsh or Ridiculing Humor

The brand of humor evidenced in RoastMe could be said to fall into the camp of the "lulz," internet-speak for *schadenfreude* that represents "amusement derived from others' misfortunes" [172]. Especially for someone unfamiliar with the RoastMe terrain, it may be difficult to distinguish between subversive humor and hateful or overtly harm-seeking comments. This falls in line with Milner and Phillips' discussion of Poe's Law [173], which is the idea that it is virtually impossible to extricate earnestness from irony on the internet; sincere extremism is indistinguishable from satirical extremism. Likewise, the effects of both humorous comments and hateful comments on those viewing or receiving the messages may thereby be one and the same; they may both be *received* as hateful comments, regardless of the differences in intentions. In my study of RoastMe, part of my inquiry lies in understanding how a declared circle of play ("comedy, not hate") is perceived and navigated by participants in that community. For whom does the magic circle exist, and what are the implications of playing within a circle that may be magical for some but not others?

In Turker Tales, I also saw the emergence of satirical and critical humor, and in YouMercials, I directly instruct a subset of participants to engage in roasting-style humor, without specifying who their target should be. In identifying the presence of

humor I draw from benign violation theory. This theory proposes that humor occurs only when the following two conditions are simultaneously met: (1) the situation violates a norm, and (2) the situation is considered to be safe or acceptable by those involved [216].

Benign violation theory's supporters argue that situations can be considered benign if a) alternative norms are in place that render the violation acceptable, if b) there is a weak commitment to the violated norm, and if c) there is psychological distance from the situation [154]. The benign violation theory helps to explain why some people view certain content as very humorous, while others are appalled by the same material. As an example, "RIP trolls"— individuals who post purposefully offensive content on Facebook memorial pages for the deceased— may view their violations as benign and thereby humorous because they feel psychologically distant from their prey. Indeed, the literature on trolling has repeatedly found that trolls and cyberbullies tend to emotionally dissociate from their targets [172], which we can view as a form of psychological distancing under the benign violation theory. In my work with RoastMe, Turker Tales, and YouMercials, I seek to understand the specific ways in which humor, especially harsh or critical humor, are both enacted and interpreted by participants, and how the use of and reactions to humor reveal participants' attitudes towards power dynamics in the context in which they are situated. I do so both by studying emergent humor absent of play directives that encourage humor (as in Turker Tales, and for certain YouMercials users, depending on their assigned condition), and humor in response to specific play directives for harsh humor (as in RoastMe, and for those YouMercials users assigned to the "roast" condition).

Social Elements of Play: Shared Artifacts with Other Anonymous Participants

In the three contexts in which I situate my research, social interactions allowed by the platforms are either pseudonymous or anonymous. My research therefore explores the power dynamics embedded and engendered by pseudonymity/anonymity, and using *research through evocative play*, studies reactions to play cues that allow for novel social communications on the platforms.

In RoastMe, those who post photos of themselves step over the line of anonymity by revealing their physical appearances, but as mentioned previously (see Introduction, Chapter 2), may do so with a "throwaway" account, effectively separating their RoastMe participation from their other engagements on Reddit, and making it nearly impossible for other RoastMe members to contact them after their roasting period

has come to a close. Moreover, most participants on RoastMe engage as "roasters," or those posting comments featuring harsh humor and ridicule, and thereby remain obscured by their Reddit handles (pseudonyms) without revealing other aspects of their identity (e.g., their physical appearance).

On Amazon Mechanical Turk, crowd workers may choose to communicate with one another through external platforms (e.g. online forums and communities like TurkerHub or TurkOpticon), but the platform itself has no built-in way for Turkers to communicate with one another. Even communications between requesters and crowd workers are limited and stilted. For example, Turkers have few options for self-advocacy if they feel a requester has unfairly rejected their work.

On YouTube, participants may comment on specific videos (though it should be noted that the large number of comments and Google's algorithmic display may limit sustained or meaningful communications). However, there is no platform-supported way for YouTube viewers to participate in discussions about the video *advertisements* that are displayed to them periodically as they watch videos on YouTube. In this way, the YouTube platform prevents communications among viewers of YouTube advertisements in a similar manner to Amazon Mechanical Turk, essentially isolating peer users on the platform from one another.

Digital platforms that limit opportunities for social interaction defy existing standards for successful online communities as defined by Kraut, Resnick, and Kiesler [128], who draw from Reiss' 16 basic needs, including social interaction [181]. Thus, directing participants to play with the original context in a way that increases social interactions while still retaining elements of the original context (specifically, anonymity/pseudonymity) could be one way to elicit critical reflection on and better understand power dynamics among participants in a context. Moreover, drawing again from Lazzaro [135], the fourth key to fun is the "social experience key," and is one of the primary ways in which players experience, enjoy and benefit from play. Using a *research through evocative play* framework, I thereby analyze the extent to which participants choose to interact with and reflect on other participants in the system in response to socially-oriented play directives.

3.3 Context-Specific Related Work

For each of the three contexts in which I situate my dissertation research (Reddit and online self-presentation, crowd work on Amazon Mechanical Turk, and online advertising via YouTube), I also am inspired by, draw from, and build upon prior work

that has studied these contexts, especially work from human-computer interaction, social psychology, and anthropology.

3.3.1 RoastMe: Reddit Culture and Online Self-Presentation

RoastMe is a comedy-focused subreddit that encourages harsh humor and involves ambiguous power relationships in that consent and the boundaries of acceptable humor are often unclear. Voluntary consent to the magic circle and the adoption of the lusory attitude among participants is ambiguous, placing those who submit their photos to be harshly ridiculed in an especially precarious position. In addition, Reddit (and by extension, the subreddit RoastMe), limits certain forms of communication in that dyadic exchanges not encouraged by the platform, and interactions are veiled by pseudonymity, and compounded by use of short-term "throwaway" accounts. Whereas in Turker Tales and YouMercials, I as the researcher define a new circle of declared play over an existing context, with RoastMe, I study play directives built into a playful platform (Reddit) that essentially takes elements that are not considered playful in other contexts (namely, harsh humor and online photographic self-presentation), brings them into the circle of Reddit and RoastMe, and in so doing, declare, "This is play, now." Thus, understanding both the culture of Reddit and its use of harsh humor as play, as well as expectations for online self-presentation as they exist in contexts outside Reddit guide my research approach to RoastMe.

Reddit Culture

Adrienne Massanari defines play as central to Reddit culture, explaining, "For Redditors, play becomes the primary mode by which Reddit culture is enacted, membership is codified, and community is solidified" [147]. Reddit as a whole shares a loose set of governing rules of behavior (known as "reddiquette") that differ from those that are prominent on other popular, more conventional social media platforms like Facebook, and there are many subreddits— including RoastMe (see Chapter 4)— that embrace aspects of *schadenfreude* [147]. For Massanari, play on Reddit takes place within a magic circle, but she acknowledges the blurriness of the distinctions between play and reality on Reddit. She explains, "...trying to define where play begins and ends in a space like Reddit becomes challenging, as the rules which help define the space are constantly shifting and being negotiated at multiple levels by both players (redditors) and designers (administrators and moderators)."

Using a lens of play, we can consider the norms of Reddit (and of particular subreddits) to be akin to the rules of play for that space. According to literature on norms in online community, one the one hand, anonymity can increase non-normative behaviors in online communities [128]. When viewed through a lens of play, we can see such non-normativeness as "rule-breaking," or rejecting the circle of declared play as magical. However, as explained by the social identity model of deindividuation, group identity can supersede individual identity [180]. This can result in behaviors that are consistent with community norms, even if the community itself identifies as subversive in nature and departs from more traditional norms. Multiple academic studies lend credence to this model. For example, a study of the collection of forums called Something Awful identified specific norms guiding community practices that deviate from the standard norms present in other online communities, such as the use of rough and abusive humor and a focus on high-quality content, regardless of how offensive that content might be [170]. Scholars have also studied other spaces that would be considered especially deviant, such as online pedophilia networks [99], and even in such domains, persistently find a set of norms and standards— albeit aberrant— that guide community behaviors. Viewing communities of play like Reddit as operating under a specific set of play rules, and reinterpreting social deindividuation theory through a lens of play, we can then puth forth that aligning oneself with the community's group identity is akin to accepting and entering into the magic circle of play in that space.

However, as Massanari points out, the lines of this magic circle, and the extent to which the circle operates as magical for those who enter the space, is up for debate. A relevant case in point: as part of a recent study of a 2015 ban of two self-identifying hate-speech oriented subreddits (r/CoonTown and r/fatpeoplehate) [33], researchers tracked the subreddits to which participants in the banned subreddits migrated. RoastMe emerged as one the top ten migration destinations for both of the banned subreddits. Thus, despite RoastMe's adamant play directive of "comedy, not hate," it appears that the site's brand of humor may be similar enough to hate speech that it appeals to users who aim to directly harm and provoke their targets. In this way, the play directives, or the rules of RoastMe, may be followed by some participants, but not others; the circle is seen as magic and entered into voluntarily by some, whereas others enter the space (circle) without adopting a lusory attitude. Although I did not design or declare the specific form of play called for on RoastMe, my analysis of how participants respond to and reflect on the play directives of RoastMe strongly informed how I understand and approach *research through evocative play*.

Online Self-Presentation

RoastMe takes cues from and incorporates the rules of Reddit as a whole, but also defines a circle of play that is highly specific to the subreddit. RoastMe's play directives position both posting photos of oneself, and harshly ridiculing those photos, as play. Ways of self-presenting (and attitudes towards others' self-presentations) on RoastMe thereby interact with and reveal participants' relationships not just with the play directives on RoastMe that name such self-presentation "play," but also with the concept of self-presentation as practiced in contexts that do not declare a space of play. I thereby turn prior literature to better understand strategies, concepts and findings related to self-presentation.

Goffman [83] spoke of social self-presentation in offline social interactions as a form of performance. Using theater metaphors, he characterizes humans as social actors who try to actively manage the impressions they leave on others, and that modify their performances based on the situation. He speaks of a front stage, where the performance of the actor/individual takes place, and a back stage, where an individual's private thoughts, needs, and desires reside. Using Goffman's theory, we can think of online social contexts as one particular situation or domain in which individuals perform their selves. The set of norms and expectations that have emerged around social media platforms create additional constraints that users interact with in the curative process of online self-presentation.

For social networks such as Facebook, where the assumed audience is comprised of friends and colleagues, studies have found that people are generally motivated to self-present authentically in online photographs [32, 8, 196]. This remains true for online social settings that involve avatar representation, such as instant messaging; people generally self-present to show their authentic personality and/or appearance [160]. Research has shown that we can reliably predict people's personality from their Facebook profile photos [32], that friend groups present their actual, non-idealized selves in social networks [8], and that individuals self-disclose differently on social media depending on their personality [196]. In fact, people may express their true (not ideal) selves better online than face-to-face [10], although computer-mediated communication may also increase self-awareness and perceiving others as self-centered [162].

However, self-presentation dynamics change depending on the online context. For example, in online dating contexts, users may be more likely to present an idealized version of the self in the hopes of attracting interest [59]. Studies have explored authenticity versus self-idealization. In one study, users that engaged with avatars representing their ideal future physical selves engaged in more healthy behaviors

[125], but other studies have shown that idealized self-fictionalization can impede expression of the "true self," where there is pressure to present oneself positively [219]. In my work with RoastMe, part of the inquiry involves observing how people choose to self-present when that self-presentation is declared as play, and how that self-presentation relates to their perceptions of the circle of play declared by RoastMe, as well as their relationship to self-presentation expectations, more broadly.

In addition, Roland Barthes' Camera Lucida serves as an important backdrop and inspiration in how I approach photographic presentation in RoastMe. Barthes spoke of self-awareness and control in photography, reflecting, "Now, once I feel myself observed by the lens, everything changes: I constitute myself in the process of 'posing.' I instantaneously make another body for myself, I transform myself in advance into an image. This transformation is an active one: I feel that the Photograph creates my body or mortifies it, according to its caprice..." [12]. Although Barthes does not explicitly name photography as "play," we can see parallels to the idea of taking on a novel identity or self through photography in the same way that players take on distinct ludic identities when engaging in play. There are also some interesting power dynamics operating in Barthes' formulation of photography; the photograph itself, even if captured by the individual (e.g., through a selfie), has its own power over the individual in the way it generates the image of the individual. Analyzing self-presentation on RoastMe through a lens of play thus tells us how participants relate to the play directives defined by RoastMe, but again, also speak to their relationship with the more expansive context of (online) photographic self-presentation.

3.3.2 Turker Tales: Marginalization in Crowd Work

Crowd work has been criticized for its low wages, which often fall well below minimum wages in the U.S., especially when taking into account the fact that unpaid elements, such as search time, are part and parcel of crowd work [96, 15, 184]. Although crowd work may offer autonomy and flexibility, it also places individuals at risk of exploitation [52, 28], with crowd work labeled as "invisible work" [145, 204, 104, 76, 105] in that workers often have inadequate support to advocate for their rights, and requesters may dehumanize crowd workers by viewing them as "cogs in the wheel" rather than individual, visible employees [126]. Applying Marxist theories of work to crowd work suggests that alienation from the outcomes of the work and other workers [95] could lead to feelings of meaningless or worthlessness [87]. Although there is little research on mental health in crowd work, research on telework and preliminary findings on crowd work suggest that crowd workers could be at risk for stress and depression [209, 116].

In order to support crowd workers, conscientious scholars have designed and implemented successful ways to assist participation and/or retention [63, 22, 230, 37], increase pay and productivity [195, 111, 30, 187, 186], improve the quality of work produced [56, 163, 136] and improve learning and skill-building [55, 82, 35, 139]. Crowd workers have self-organized through forums to improve work practices, increase income, find new HITs, and provide moral support and encouragement to one another [215, 132, 89]. Building off research from community empowerment [68, 203, 214], studies have also helped bring light to and support crowd workers in actions to self-organize and advocate for their rights as workers [105, 106, 104, 14, 40].

Relatively little work, however, has considered support for crowd workers in ways that are untied from work efficiency, quality, and earnings. Some interesting interventions have embarked on this territory, using methods like play and curiosity to engage crowd workers [133, 47, 61], or focusing on more holistic social support [58], but we note that these research projects remain focused on work quality and/or efficiency as the end goals. In my work with Turker Tales, I inject play directives into a crowd work setting that are designed to be *tangential*, that is, not seeking to improve workers' performance or efficiency. Following, I study the extent to which participants choose to engage in tangential play, and how their different forms of engagement shed light on worker-to-worker and worker-requester power dynamics, and elicit critical reflection on the crowd work context.

3.3.3 YouMercials: Video Advertising on YouTube

Reddit is a space of existing play— a (magic) circle unto itself. By contrast, Amazon Mechanical Turk more neatly falls into a space of "non-play" in that it is a labor marketplace run by a large corporation. Meanwhile, the specific space of video advertising on YouTube (a platform where users can upload, share, and view videos from other users), falls somewhere in the middle.

On the one hand, YouTube itself is a space of play and leisure, allowing for the viewing of a multitude of entertaining and edifying videos. Like Reddit, YouTube can be viewed through a lens of play and participatory online culture [26, 228, 41, 107], with users creating and uploading their own content to entertain, educate, or otherwise share with other users. User-viewers contribute and participate by watching and voting (much like Reddit), which then determines what others view, as well as commenting on videos.[1] YouTube has also been used to bring light to human

---

[1] Although comments on YouTube have a bad reputation, they are still the primary means through which YouTube viewers can connect and communicate, and are used by roughly 12% of YouTube users [193].

rights violations that would be censored elsewhere (though it should be noted that the same lack of censorship has led to the propagation of forms of freedom of speech on YouTube that also impinge on human rights, such as ISIS propaganda [7]).

However, even while "playing" or relaxing on YouTube, the power dynamics of that play are complex in nature. As multiple scholars have argued, [178, 5, 157, 72, 71, 73, 74, 192], "play" on YouTube can simultaneously be viewed as labor for a capitalist corporation.[2] YouTube relies on both its user-consumers (viewers) and user-producers (content creators) in order to operate and produce value and profit. For example, by upvoting and downvoting videos, viewers produce data that feeds YouTube algorithms and allows the platform to operate and produce revenue; upvoting and downvoting could thereby be seen as a form of free labor. Gerlitz and Helmond have termed this dynamic the "like economy" [80], whereby users are induced to "play a game" that ultimately benefits a large corporation. Users' views and likes guide YouTube's "black box" algorithms, which subsequently influence the content to which users are exposed to and encouraged to watch [169]. In this way, by viewing videos on YouTube, users "produce the cultural content of the commodity" in what Lazzarato terms "immaterial labor" [134]. Such immaterial labor is defined as activities that are not normally recognized as "work," e.g. activities "involved in defining and fixing cultural and artistic standards, fashions, tastes, consumer norms and, more strategically, public opinion" [134]. This process of simultaneous consumption and production by users of a platform has also been described as "produsage" [23] and "prosumption" [183].

The more specific context of video advertising on YouTube is especially illustrative of imbalanced power dynamics.[3] Not only do video advertisements interrupt YouTube play, bringing to the fore the role of the corporation in controlling user experiences, but users themselves contribute financially to the platform by viewing such advertisements. YouTube viewers may not be fully in control when using the platform in general, as their views are algorithmically driven, and monitored and harvested to benefit a large corporation. When viewing advertisements, users specifically do *not* make the choice of viewing, and in viewing, they do productive work for the company, as Google and YouTube can essentially sell the user's data and views to

---

[2]Since 2005, and as of the time of this writing, YouTube is owned by Google.

[3]Although my work with YouMercials focuses on users' perspectives and users' perceptions and navigations of power on YouTube, I also want to acknowledge that from the advertising side, too, the situation is complex. For advertisers, issues of copyright and content collide. Because user-uploaded content on YouTube often falls into the realm of copyright infringement, brands and corporations may be wary of advertising on the platform [5, 152]. In addition, because advertisements are targeted and displayed algorithmically through Google, advertisers are unable to control exactly how and when their advertisements are displayed. As a result, many large corporations have recently pulled their advertisements from YouTube after their ads appeared next to controversial and unsavory content, including videos promoting extremist views and hate speech, and videos that either directly appeal to pedophiles, or inadvertently attract a large number of pedophilist comments [202, 227].

advertisers. In my research, I study how overlaying play directives onto the specific sub-context within YouTube of video advertising can reveal how participants perceive and relate to targeted advertising on YouTube.

3.4 Summary

From all of the prior research I have discussed here— design research in human-computer interaction (including research through design, critical design, and ludic design), research on play and play theory, and context-specific research on Reddit, online self-presentation, crowd work, and YouTube advertising— I draw knowledge, inspiration and guidance that greatly influences my *research through evocative play* approach. In the following chapters, the research I present both builds upon and converses with this prior literature. In the following three chapters, I present my research in three different domains— Reddit's r/RoastMe subreddit, Amazon Mechanical Turk, and YouTube (specifically, YouTube video advertising)— to show the path I took to iteratively develop and implement the *research through evocative play* approach. In Chapter 4, I will first begin with my study of RoastMe, which I view as a precursor to the *research through evocative play* approach I adopt and present in the latter two studies (Chapters 5 and 6).

RoastMe: The Implications of Casting Harsh Criticism and Online Photographic Self-Presentation as Play

<div style="text-align: right; font-size: 2em;">4</div>

I present my work with RoastMe, an online community that declares harsh humor and online photographic self-representation as play, as a precursor and fundamental inspiration for my approach to *research through evocative play*.

4.1 Introduction

[1]Few people would likely characterize verbals attacks, decidedly unconstructive criticisms, and otherwise offensive comments based on one's physical appearance as the foundation of a positive online experience. The over one million subscribers to r/RoastMe,[2] a humor-focused sub-community (subreddit) of the large online forum reddit.com (Reddit), might beg to differ. On RoastMe, people willingly post photos of themselves to have other users ridicule, offering themselves up as the target of jokes and comments that could be construed as offensive, impolite, or politically incorrect. For example, responses to photos often include markers of racism, sexism, violence, body-shaming, and homophobia. In this way, RoastMe declares sets of behaviors— online photographic self-presentation and harsh humor— as play ("comedy, not hate" = "this is play, now"). Declaring such behaviors as play, and doing so within the

---

[1]This chapter is based off research published at ICWSM 2018 [113].

[2]1.4 million subscribers as of June 2019. During the qualitative interviews conducted for this work, subscribers numbered roughly 300K; by the following summer, numbers were already up to 600-700K.
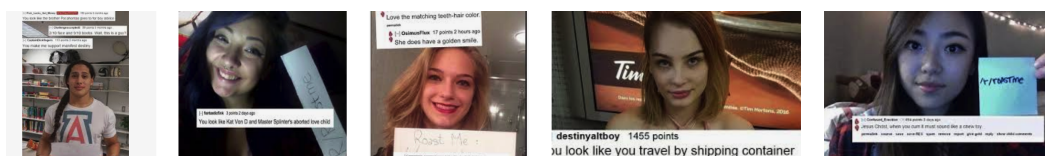

Fig. 4.1: A sample of photos and associated comments (roasts) posted to RoastMe. (Obtained via Google image search, October 2018).

larger context of Reddit, a site that generally encourages play and irreverence [148], as I will show, still does not create a crisply circumscribed magic circle of play on RoastMe. Instead, the declared play of RoastMe impels participants to poke fun at and contemplate contrasting notions of online self-presentation and humor. In participating in RoastMe, members also consider the implications of engaging in the play the subreddit declares, wherein the boundaries of consent to the rules of play are not always clear.

Of course, RoastMe is not the first context to cast harsh humor and self-presentation as play. Roasting as a form of comedy is nothing new; dark humor and satire have a long history in comedy (take the works of ancient Greek playwrights and Shakespeare, for example). Roasting, or subjecting a singular "guest of honor" to insulting but generally good-natured jokes for an audience's amusement, is a concept that predates both Reddit and the Internet. In the 20th century, it became common practice among (wealthy) clubs and organizations to hold testimonial dinners praising a guest of honor. The Friar's Club, founded in 1904, began putting a spin on these testimonial dinners, introducing jabs and sarcastic humor. By the 1940s, insult humor became the defining features of the Friar's Club roasts [2]. The practice has extended to modern times, with the television network Comedy Central hosting roasts of celebrities like singer Justin Bieber, actor James Franco, and even the current president (then television host) Donald Trump [17].

On RoastMe, within a particular roasting thread, one participant (the roastee) posts a photo of themselves, thereby offering themselves up to be the target of the roast. As in traditional roasts, on RoastMe the "roastees" ostensibly participate of their own volition, and the directive for all members of the community is "comedy, not hate." However, in traditional roasts, roasters are typically friends and colleagues of the roastee, and roasters, the roastee, and the audience (also friends, acquaintances, and general supporters of the roastee) participate together in the event in person. In contrast, on RoastMe, the roasters are identified only by their anonymous Reddit handles (pseudonyms), and are strangers to the roastee. The community may not be everyone's cup of tea, but its situation at the intersection of anonymity and exposition, its declared rules that frame photographic self-presentation and harsh humor as play, and the inter- and intra-personal dynamics it engenders make it a fascinating and instructive milieu for human-computer interaction research. RoastMe provides an opportunity to study how declared rules of play drive and impact participation in a subversive humor community and in turn, to consider how the power dynamics on such a site could inform the design of other online domains as well as inspire *research through evocative play* study designs in other contexts.

Seeking to understand the rules of play declared by r/RoastMe, I embarked on a two-part mixed methods study of r/RoastMe. In the first part of the study, I

conducted semi-structured interviews with 16 members of the RoastMe community. In line with previous work on unconventional or subversive online communities, the RoastMe community highly values abidance by community-specific norms. Through a lens of play, we can see these community-specific norms as rules of play; RoastMe participants highly value and abide by the rules of play defined within the community. At the same time, I discovered unique aspects to the power dynamics operating on RoastMe, especially with respect to participants views towards and understanding of the subreddit's circle of declared play. On RoastMe, roasters make heavy use of perspective-taking, roastees self-present to show or even amplify perceived flaws and find value in harsh judgment, and participants (especially roaster-participants) are highly concerned about harm generation in the community. In this way, it appears that many participants in RoastMe view the circle of declared play as ambiguous and problematic rather than fully magical.

In the second part of the study, I extend my qualitative findings by quantitatively exploring a data set of RoastMe activity from June-August 2017 collected through the Reddit API. The data set consists of roughly 9,000 posts of photos to be roasted, and 230,000 comments on (i.e., roasts of) these posts. In this part of the study, I explore whether my interviewees' descriptions of the community cohere with the activity markers and participant attributes present in the data set, and I cluster topics present in the data set using Latent Dirichlet Allocation (LDA). In addition, I build a linear regression model to predict the popularity of posts based on roastee attributes, with my findings suggesting that those who appear to be female, and have Reddit activity that indicates possible mental health struggles are more likely to to be targeted with more comments (roasts). Together, my qualitative and quantitative explorations point towards design changes that could improve user experiences in RoastMe, especially concerning ways to reconcile differential levels of consent to the declared rules of RoastMe play. This work suggests opportunities for research and design in human-computer interaction to meet "deviant" desires that could extend to domains beyond the context of RoastMe, including ways in which RoastMe's design can inform *research through evocative play* study approaches.

4.2 Methods

4.2.1 Semi-Structured Interviews

For this study, I received IRB approval and recruited 16 participants from r/RoastMe. To do so, I contacted the Reddit account shared by all the group moderators, and a moderator assisted me by pinning a description of the project to the front page of the subreddit. Interested parties were instructed to contact me via Reddit and/or email.

Each interview lasted about 1 hour and took place over Skype or Google Hangouts. Specific questions I asked as part of my interview protocol included: "What, if anything, do you find challenging about participating in RoastMe? Rewarding? How would you describe the RoastMe community? How did you come to be a roaster/roastee (probing on specific circumstances)? What motivates you to participate as a roaster/roastee?" I compensated all participants with a $15 Amazon gift card.

As part of my study, I included a brief, write-in demographic questionnaire; participants in my study skewed young, white, male, and North American. The average age was 22.5, with a median of 19, 13/16 were male, and 9/16 identified as white or Caucasian. Of the remaining seven, three identified as Latino/a or Hispanic, one identified as brown, one identified as African, and two identified as Asian and/or Chinese. Twelve participants hailed from the US, and one each came from the UK, Canada, Russia and South Africa, respectively. For 15/16, English was a primary language; three of these native English speakers also held a second native tongue (one participant each also spoke Chinese, Spanish, and Afrikaans). There was also one Russian speaker who did not identify English as a primary language. For occupation, five identified as students, three as employed in sales, two as self-employed, and two as artists/musicians, with the remaining six employed in consulting, teaching, IT, and penetration testing/ethical hacking. Two of the participants were RoastMe moderators who had also participated as members, while the rest were general members.

There are two primary ways for members in the community to participate in r/RoastMe: either 1) by posting as roastees, in which they upload a picture of themselves with an adjoining title asking for others to roast them, or by 2) commenting as roasters, in which they provide insulting/humorous comments in forum threads about the photos and titles that other members have posted as roastees. Two of my study participants had exclusively posted as roastees, six had exclusively commented as roasters, and the remaining eight participants had engaged both as roasters and roastees.

I audio-recorded and transcribed all interviews, and then conducted open coding on the transcripts, starting with high-level emergent themes and then iteratively refining the codes. I shared a preliminary codebook and a sample of 20% of the responses with my advisor for independent coding. Based on comparisons and discussion of disagreements in the resulting codes, I then revised and refined the codebook and shared a new sample of quotes reflecting about 20% of the data set. Based on this final test set, the Cohen's Kappa measure of inter-rater reliability from this subset was 0.98. Patterns and findings emerged around depictions of the community and interactions and behaviors therein, revealed in low-level codes such

as: (1) "outlet_haven" to describe RoastMe as a safe place to let out non-politically correct rhetoric, frustration, and other types of potentially offensive commentary; (2) "perspective-taking," to reflect ways in which roasters adoped the point-of-view of others while roasting (e.g., putting themselves in the poster's shoes, identifying with other participants, or thinking about the poster's feelings or desires); and (3) reasons why roasts were seen as negative or undesirable, such as "direct_support," reflecting disapproval of comments offering support or flattery to the poster instead of roasting the poster, which violates RoastMe rules.

4.2.2 Analysis of Reddit API Data

The qualitative findings resulted in several further questions that I subsequently explored through quantitative analyses. The Reddit API allows users to access historical comments and posts. Although I initially used the Reddit API in conjunction with the Python wrapper PRAW to collect data, latency and limitations on the number of calls led me to a publicly accessible version of Reddit's historical data on Google BigQuery. Using Google BigQuery's SQL interface, I collected all the image posts and comments (roasts) from RoastMe between June 1, 2017 through August 31, 2017, totalling over 290,000 comments (roasts) and over 12,000 posts (photos to be roasted). The data set included features such as (for roastee posts) time of post, number of comments received, author of the post, and text title accompanying the post, and (for roaster comments) body of the comment, and parent ID of the post with which it is associated. Using the implementation of the Latent Dirichlet Allocation (LDA) in NLTK (Natural Language Toolkit) for Python, I then performed topic modeling on the roastee titles.

I conducted additional queries to mark each roaster and roastee with the number of posts (roastees) and roasts/comments (roasters) they had made since the subreddit's birth in 2015 to more comprehensively determine the nature of roasters' and roastees' participation in the subreddit. Lastly, because several of the interview participants had expressed uncertainty about the mental stability of some of the roastees, and the morality of roasting in such cases, I curated a list of 34 popular subreddits devoted to mental health struggles drawn from various subreddit guides and other online recommendations. This list included subreddits such as "suicidewatch," "suicidenotes," "depression," "eatingdisorder," "stopselfharm," and "survivorsofabuse." I marked any individuals in the data set who had submitted posts or comments to the list of mental-health related subreddits as possibly vulnerable to mental health issues.

Following these steps, I removed untidy instances, corrupt instances, and instances with key data missing. For example, I excluded from my analysis posts and associated

comments that were missing the column for the number of comments received, and excluded comments that were not associated with any posts from the time period (e.g., comments on posts that had been submitted prior to June 1, 2017). This resulted in 234,475 comment instances representing 53,885 unique roasters, and 9,032 posts representing 6,013 unique roastees.

Next, I employed Microsoft's Cognitive Face and Emotion APIs to estimate the age, gender, and emotions of roastees' photos in the data set. Because all Reddit users are anonymous/pseudonymous, the Reddit API does not make available demographic information such as age or gender. However, by applying Microsoft's Cognitive Services to the data, I can estimate these demographic features and incorporate them into the descriptions and analyses. I acknowledge that this is a rough and highly imperfect measure. For example, Microsoft's Face API will predict gender based on one's physiognomy, not one's self-presentation or internal identity. Thus, an individual who self-presents and/or self-identifies as female or non-binary but has physical markers of a masculine facial structure will be classified as male, not female, by the Face API. Nonetheless, human perception shares some of the same flaws as algorithmic approaches, and many viewers on RoastMe may assume the gender of roastees based on their physical features, and roast them accordingly. For the purposes of my analyses, I thus deem my gender estimate to be a flawed but still reasonable proxy for gender. Microsoft does not offer clear metrics about its services' reliability, so I conducted an informal check on the data by taking a random sample of 100 image post instances and manually coding them as "male," "female," or "unknown"; when I compared this to the Microsoft Face API results for these instances, I found a 99% match. I incorporated these newly derived features (estimated age, gender, and emotions) into my data set.

Lastly, I a built a linear regression machine learning model to predict the estimated popularity of a post using WEKA software as well as the ADAMS workflow tools to assist with converting data sets to ARFF format. Here, I used a subset of my data that included only posts for which I was able to obtain age and gender information; posts with low quality, blurry, or deleted photos were thus excluded from the analysis, leaving me with 4,710 instances.

Reddit has features to upvote and downvote comments and posts on any subreddit, but it does not display the actual number of upvotes and downvotes on the site (numbers are "fuzzed" to avoid spambotting) and returns only null values for upvotes and downvotes in its API. The API does make scores available, which are defined as the number of downvotes subtracted from the number of upvotes, but this measure provides limited information; for example, a post with 1,000 upvotes and 1,000 downvotes will share the same score as a post with one upvote and one downvote. Thus, I turn to the number of comments a post receives as a flawed but reasonable

choice for my model's dependent variable, given that it approximates the amount of traffic that a post receives. I first divided my post data into train and test data sets using an 80:20 split, and iteratively refined the model. After tuning my model, I evaluated the final model on my complete data set using 10-fold cross-validation.

4.3 Qualitative Results

First, I present findings that, in keeping with previous literature, show that the RoastMe community supports a specific set of norms. Using the lens of *research through evocative play*, I speak of these norms as a specific set of *rules of declared play* that the RoastMe community supports. Next, I define how RoastMe participants relate to the rules of declared play in unique ways, highlighting a) roasters' heavy use of perspective-taking, b) roastees' values and self-presentation behaviors in relation to harsh judgment, and c) the high level of concern about harm generation in the community. In the Discussion section, I will explain how these findings about RoastMe participants' relationships to the rules of declared play are instructive for considering the design of online communities, as well as *research through evocative play* study designs.

4.3.1 Adherence to RoastMe's Rules of Play

Here, I present an overview of the ways in which adherence and commitment to RoastMe's declared rules of play are manifested in the community, including the high value placed on humor and originality, the expectations around consent, and community- and administration-driven moderation practices. Gaining a general understanding of RoastMe guidelines and behaviors will help the reader to see how my study extends previous findings about adherence to norms in subversive communities, as well as provide context for comprehending the unique aspects of RoastMe's rules of play that can inform novel design directions for online communities and online self-presentation that I will discuss in subsequent sections of this chapter.

As stated directly on the site, RoastMe is about "comedy, not hate." We can consider this statement a declaration of play and the overarching rule of declared play in RoastMe: "You must engage playfully, not hatefully" or in so many words, "This is play, now." My participants echoed this overarching rule of play in their discussion of the site. For example, a participant who was also a moderator (p12) called the phrase their "rule number one." They also expected other participants in the site to adhere to the guiding principle. For example, p16 explained, "RoastMe is about comedy; it's about making people laugh. It's not about starting a hatred thread. . . It's just that: I wish people would understand the rules about RoastMe."

The community also values creative, unique, and genuinely funny content; to post unoriginal content, then, is for many one of the highest RoastMe offenses, as investigations of other subversive humor forums have similarly shown [170]. Taboo and offensive comments are perfectly acceptable within RoastMe; being unoriginal, however, violates RoastMe rules of play. One participant explains that roasts that encourage suicide are not off limits, but do violate the rules if they do so in a generic, uncreative way. He explains that such roasts are unacceptable, saying, "If it's in the generic rules, like 'Kill yourself,' because this is against the rules—roasts like that I don't think are allowed anymore. Because it's just too flat-out and generic; if it was more creative then it would be allowed" (p1).

Another participant (p11) directed me towards a previous post he had made outlining the kinds of roasts that were not acceptable; here, he advocated against certain stereotypes because they were generic, overused, and therefore unfunny, violating the value in RoastMe of creative, humorous, and original roasts. These included racist, homophobic, and violent roasts. The key here is that the offensive nature of these roasts was not the racism, homophobia, or violence per se, but rather, the lack of originality in their deployment, in violation of the stipulations of RoastMe's declared play. Six participants actively helped enforce RoastMe rules, stating that they had either downvoted or directly reported comments to the moderators because they considered them to be unoriginal and of low quality, thereby violating the rules of RoastMe play. As one participant explained, this can be a problem with newcomers who haven't yet learned the RoastMe rules: "Most people, like, when new people join the community, it's basically just racist comments and stuff like that until they realize these just get downvoted a lot, so then they come up with original content" (p9).

Of course, under the "comedy, not hate" rule, any comments that are specifically intended to produce harm and not humor are not permitted, either. To this end, six participants also shared that they had reported or downvoted comments that they felt were unnecessarily cruel. As one participant explained, "You know, if you're saying like, 'You look like your grandma just died,' that's funny. But if you say 'I hope your grandma dies. Go burn in hell,' that's a little weird. Those comments usually get downvoted into oblivion, just gone" (p6). Another participant (p9) said they were not okay with people roasting people in the background of the photo, as this goes against RoastMe rules—people in the background did not directly give their consent to be roasted.

Participants also spoke negatively of actions that might be viewed positively in other contexts, but that violated the RoastMe rules of play and were therefore deemed to be inappropriate. For example, one participant (p11) said he disapproved when people offered direct, encouraging support to a roastee such as "I hope you're okay,"

explaining, "What bothers me is when people mess with the process of RoastMe." A moderator (p13) explained that flattery is not permitted: "One of the rules of RoastMe is that you can't hit on any of the Roastees, so like, nobody can be like, 'How can I roast you? You're actually kind of hot' or whatever."

In keeping with the literature on community norms and benign violation theory, seven of the participants specifically viewed RoastMe as an outlet, a safe space where they can release built up frustrations, "satisfy urges" as one participant (p11) put it, and give voice to taboo thoughts that they can't express elsewhere. P11 compared RoastMe to trolling that happens on other internet sites and forums, saying that because so many social spaces encourage us to repress taboo and potentially harmful thoughts, we might have "explosions" that occur on trolling threads, RoastMe, or "any anonymous outlet that people can help unleash that." Others echoed this urge to unleash, saying, for example, "sometimes it's fun to say mean things that you can't say in person" (p3). In this way, these participants saw RoastMe as a magic circle in which it was acceptable to engage in behaviors as their ludic selves that they would not be able to do as their "ordinary selves" in other contexts outside that magic circle.

Fourteen of the sixteen participants cited the expectations of the community to explain why comments they or others might deem as offensive in other contexts were copacetic on RoastMe. Acknowledging that people who post photos of themselves are willingly doing so, and know to expect harsh and insulting feedback, produced the sentiment that everything is "fair game" (p2). As p7 explained, "...it's in a controlled environment where it's expected that you're going to be mean and you just be as funny about it as you can be." Another participant (p15) contrasted it with social media cyberbullying, explaining that even if the comments look similar, the intents and expectations of RoastMe create a new, safer context in which such comments would not constitute bullying. Again, from this standpoint, RoastMe operates as a magic circle in which such behaviors are condoned by the rules of play.

This feeling of safety within RoastMe was further supported by the forum's moderation, as revealed by moderators themselves, as well as opinions other RoastMe participants expressed about moderators. Moderation helped participants feel that this was a safe space to engage in taboo or trolling-like behaviors without causing too much harm. Moderation of RoastMe includes efforts to ensure consent, such as stipulating that all roastees must hold up a hand-written sign reading "RoastMe," and further monitoring the posts to see if PhotoShop may have been used to forge consent. Participants who have posted a photo of themselves and then later regret it or are upset by the comments can contact moderators directly to request deletion, and moderators reported that they will immediately delete the thread; one moderator

stated that several of such situations had happened. Moderators also try to protect youth, requiring that roastees be at least sixteen years of age.[3] While much of this process is self-reported and verifications of age largely subjective, moderators also use bots, such as a bot that helps protect people with poor mental health by crawling through each roastee's posting history to learn if they've posted in any pro-self-harm subreddits. Moderators thereby act as enforcers of the specifics of RoastMe's rules of play, attempting to ensure that RoastMe's circle of play remains magical.

Understanding that RoastMe operates under a specific, defined set of (declared) play rules is important context for delving into the unique nature of how participants interpret, reflect on, and respond to RoastMe's rules that I will now discuss.

4.3.2 Perspective-Taking

Where other target-perpetrator dynamics of subversive humor emphasize emotional dissociation [172], in the interviews, I learned that perspective-taking is a critical component of engagement in RoastMe. This finding indicates that RoastMe's rules of play encourage perspective-taking in unique ways. I saw perspective-taking emerge particularly often during the roasting process. Ten of the 14 participants who had commented as roasters discussed directly engaging in perspective-taking as part of the roast process. Participants often imagined themselves in the place of the person who had posted a photo, and thought about what they would want to hear in their place. For example, "I want them ['roastee'] to feel as content as I would want to feel" (p3). Others (p12) asked themselves questions like, "What would I want to hear about that [physical] feature? What would I think that's funny about myself based on that?" This might include taking on the perspective of the general RoastMe audience such as, "What would I think if I wasn't in the picture and reading the comments?"

Roaster participants sometimes sought out roastees with whom they personally identified; for example, one of the three female-identifying participants I interviewed explained that she specifically tries to comment on other women because she shares their perspective. Identifying with roastees also can have benefits for the roasters. As p3 explains, "Being able to see someone that has the same flaws as me makes me realize that I'm not alone, and then being able to criticize someone else about it makes it a little bit better..." P15 spoke of how he hoped that by making roastees laugh, he could make them feel valuable and give them hope. Another participant (p11) explained, "Yeah, I really believe that laughter is the best thing you can do for someone... if it's terrible and you can get someone to laugh, that's a couple

---

[3]At the time when the interviews were conducted, the rule was still 16 years of age. As of this writing, the subreddit's rules have changed; roastees must now be 18 years of age or older.

seconds at least where you can distract them from something bad and they can enjoy themselves." Others spoke of tempering their roast so as to "zing" without being too harmful, indicating a mindfulness of other participants' feelings.

The use of perspective-taking within RoastMe is further reflected in the high value placed on participating as both a roaster and a roastee within the community. Participants characterized RoastMe as a "two way experience." As p15 explained, "It's not just people posting photos and getting roasted, it's also giving the chance for someone else to anonymously critique or to make a comment about someone else…" This participant (p15) had not yet posted his own photo to be roasted, but still felt there was value in the two-way process. He explained, "I might do it [post] in the near future, because I think that it's unfair for someone in the community to give out these comments that seem kind of hurtful— even if they're funny on some level— I think it's quid pro quo. If you made comments, have gotten upvotes, then you should probably subject yourself to the same thing." Another participant (p3) explained that participating as a roastee helped him become a better roaster: "… being a roastee made it so I could see the other person's point of view when I'm writing comments about them, so I know what they're kind of going through."

Although less prominently represented in the participants' responses, perspective-taking may even take the form of encouraging others to perspective-take (what I'll call "perspective-sharing"). This was true of one participant (p11) who often targets roastees who appear in some way privileged (e.g., those who are judged to be more attractive or generally happier or more well-off than average). As he explained, "And I'll admit that I think that I've had more than my fair share of suffering, and for whatever reason, this comes into play in my roasts; that I want other people to feel that as well."

Thus, we can see that unlike other, more dissociative contexts in which individuals are provoking or ridiculing a target, perspective-taking is integral to roasting on RoastMe. The RoastMe community adopts the declared rules of play on RoastMe in a way that values perspective-taking over dissociation, and sees perspective-taking as an especially lucrative and rewarding play strategy.

4.3.3 Concern about Harm: An Uncertain Magic

Until now, I have focused on the ways in which RoastMe participation is viewed as play, in accordance with the declaration of play on the subreddit. However, the boundaries of this supposedly magic circle were also unclear for many participants, particularly with respect to *for whom* and *to what extent* the circle remains magical. I found that many of the RoastMe participants were quite concerned about the well-

being of other participants. From the perspective of benign violation theory, there were indications that the level of benignity of the humor violations was dubious for many participants; they expressed uncertainty as to whether it was truly safe and okay to engage in certain forms of humor on the site. For example, four of the participants expressed guilt, secrecy, or regret about their participation in RoastMe. The lines between cruelty and humor were not always clear, and participants worried how their actions might affect others negatively. One participant (p15) discussed how he sometimes deletes comments after writing them, such as when he removed a self-authored comment that he felt reinforced an unfair, negative stereotype about gay women.

Participants saw people who posted photos as thick-skinned, that is, as willing and prepared to take on whatever the RoastMe community will throw at them. Indeed, the interviewees who had posted photos of themselves all described themselves as equipped to "handle the heat" in one way or another. Yet some roasters questioned whether they really could trust this assumption. They questioned whether all those posting had truly entered into the magic circle voluntarily, adopting the lusory attitude with full consent. This was especially true if the participant self-identified as insecure or "thin-skinned." Four of the participants who had chosen only to engage as a roaster, not a roastee, echoed this concern. For example, one participant conjectured that perhaps some people post photos of themselves because they have low, not high self-esteem (p7). Another participant stated that some comments on RoastMe can be exceedingly harsh, and that, "I kind of worry about destroying people's self-esteem" (p6). Another worried that there could be site passersby who could be negatively affected by roasts "by proxy," explaining, "It may actually just hurt the people there that are viewing the subreddit based on what they may think is unattractive and they may feel self-conscious about it" (p1). The participant, who stated his race as African in the demographic questionnaire, said he found himself feeling self-conscious about how others view him after reading the plethora of racist roasts of others on the site, wondering, "Is that really what people think the first time they see me?" Others worried that even though there are age cutoffs for the site, there might be young people participating that could get seriously injured. "Honestly, I don't like it when kids come onto RoastMe because I don't think they can handle it" (p5). If those posting their photos on RoastMe did not fully consent to the declared play of RoastMe, then participants worried that their own engagement as roasters was likewise no longer confined to a magic circle.

Participants also expressed the desire to remain within ethical lines. However, descriptions of where these lines actually fell varied greatly by participant. Many expressed that they took no issue with any offensive comments, but others felt certain kinds of comments, such as those that were directly racist, insulting of one's religion or one's sexual preferences, or treating rape or child abuse as a joke, crossed the

line, and subsequently would either downvote or report comments falling into these categories. For example, participants had mixed reactions to posts that referenced suicide or self-harm. As participant p7 explained, "There are a couple where the title is something depressing like, maybe they even reference being depressed, or self-harm or even suicide, and to be honest, those ones make me a little uncomfortable so I don't post on those ones. Because you never know: someone might be joking, but a lot of times even when someone's joking, there might be some degree of truth to it." In this way, the perceived boundaries of the magic circle of play declared in RoastMe was interpreted variously by different participants.

Although first impressions of RoastMe might lead one to assume that roasters are not particularly concerned about the feelings or well-being of their roastee targets, the interviews suggested otherwise. On RoastMe, desires to engage in humor that borders on the offensive and cruel coexist with desires to protect and support other community members. RoastMe participants were eager to engage in RoastMe as play, but were also simultaneously concerned about the boundaries of the magic circle the community circumscribes.

4.3.4 Benefits for Roastees

Heretofore, much of my focus has been on the values of roasters. Now, I zone in on how being a target of ridicule creates value for roastee participants. Participants cited "skin thickening" as a benefit to posting as a roastee; by subjecting themselves to criticism in the RoastMe space, they felt they would be able to handle "real life" insults and abuse. As p13 explains, "It just makes life so much better because even when somebody now insults you in the real world, you can laugh at it. Even though everyone's trying to tear each other down, there's still a feeling you came out as... almost a newer version of yourself. You have this newfound respect for yourself, and thicker skin..."

In fact, seven participants cited finding insights into themselves from anonymous strangers, and learning ways in which they might improve themselves, as key reasons for posting. Some even changed their behavior after posting, such as one participant (p4) who modified how he dressed and wore his hair, and another participant (p6) who said she was glad someone had ridiculed her dark under-eye circles because she hadn't previously noticed them, and had started wearing more under-eye makeup as a result. Another participant (p10) posted soon before going on a date. After receiving jokes about his hair, he decided to get a haircut, noting that, "I actually needed to get a haircut, so it was good for the date." This same participant, who was very nervous prior to the date, found that the process helped give him perspective, stating that it helped him "not take things too seriously."

Participants also enjoyed having people they didn't know being able to accurately pinpoint aspects of their true selves; nine participants found this type of comment particularly humorous. Whereas, as discussed in related work, photographic self-presentation on certain stranger-dominated online contexts such as online dating platforms can encourage idealized self-presentation, in RoastMe, only one participant directly strove to hide what they deemed to be physical flaws (in this case, the participant stated he had purposely occluded his teeth). Seven participants chose to self-present authentically, many with the hopes of garnering honest feedback. In addition, I also saw a third type of photographic self-presentation that previous literature on online photographic self presentation has not yet discussed. Several participants actually *amplified* ways in which they might be considered flawed or aberrant in order to provide more roasting fodder, such as p12, who stated he tried to look dead inside in his photo in order to elicit better posts, or p3, who advertised in his posting title that he is Jewish, an artist, and bisexual in order to elicit funnier comments. In RoastMe, value is not just created for the roasters who want to engage in humor or unleash behaviors that might otherwise be considered taboo. Roastees, too, discussed how they benefited from RoastMe, gaining resilience and useful insights about themselves.

In this respect, roastee participants that I spoke to did see the RoastMe as embodying a circle of play, but we note that they viewed the circle as especially porous, or as situated within the larger context of the identities *outside* that circle of play. They did not interpret the feedback they elicited by engaging in RoastMe play as restricted to their ludic selves on RoastMe, for the selves they portrayed on RoastMe were highly tied to their "ordinary" selves, either meant to be directly reflective of their authentic selves, or reflective of flawed, conflicted, or aberrant aspects of their "ordinary" selves.

From the interview study, I learned that RoastMe participants abide by a set of play rules that suggest a different lens through which to view and design for subversive humor in online communities. In RoastMe, perspective-taking rather than disso-ciation from targets guides roasting behaviors; concerns about harm coexist with desires to engage in offensive humor; and benefits accrue to those who are the butt of the joke. In the next section, I present the quantitative results that extend these qualitative findings.

4.4 Quantitative Results

4.4.1 Descriptive Findings

In the qualitative portion of this study I just discussed, saturation rather than representation was my aim, but I will note that my sample ended up being fairly representative of the RoastMe community at large. Of the 4,710 posts in the cleaned quantitative data set, the Microsoft Face API classified 79.9% as male by the Microsoft Face API, and estimated the average age as 26.2 (instances with unknown age and/or gender values were omitted from the data set). As several of the interviewees spoke of taking a plain, authentic photo, we might expect neutral expressions to be common; the data confirms this. Over half of the participants in the data set displayed an expression that was classified as neutral (using a threshold of 0.5/1 or above).

The interviewees also discussed the value of reciprocity in the community (participating as both a roastee and a roaster). However, the data set suggests that although reciprocity may be a perceived shared value and implied rule of play in the community, it may not be quite as common in practice. Whereas a high proportion—87%—of the 3,964 unique roastees in the data set also contributed a roast at some point, a much lower 9.6% of roasters (5,183 of 53,885) had proffered photos of themselves for the community's ridicule, indicating that many RoastMe participants are only experiencing one side of the RoastMe experience. Meanwhile, a relatively small but far from invisible proportion of roastees—5.7%, or 224 roastees—had submitted content to mental health oriented subreddits that might indicate cause for concern, validating some of the participants' fears about whether the humorously-intended violations enacted on RoastMe are truly benign, and by extension whether their own participation in RoastMe as roasters could truly be deemed as safely restricted to a magic circle of play.

4.4.2 Topic Modeling of Post Titles

Iteratively tweaking the parameters to arrive at semantically meaningful results, I used Latent Dirichlet Allocation (LDA) to conduct topic modeling on the unfiltered data set of RoastMe posts (12,046). The most semantically meaningful model was built using 20 passes, 3 topics, and 5 words per topic. As I expected from informally browsing through the forum, RoastMe post titles are very homogeneous; the term "roast," for example, emerged as the highest or second-highest weighted term in all three models. Nonetheless, my model allowed me to generate three broad categories of titles: 1) titles stressing the age of the roastee (e.g. "I just turned 19 years old, roast me!"), 2) titles asking roasters to do or give "their best" (e.g. "Go ahead, roast me. Give me your best roast") and 3) titles indicating that the post is made on behalf of a friend, and that they want roasters to "do their worst" (e.g. "My friend thinks he

score	mental_health	age	female	sadness
0.24	9.41	0.49	17.95	-14.97

Tab. 4.1: Independent Variable Coefficients

corr coeff	MAE	RMSE	RAE	RRSE
0.86	22.49	41.94	66.26%	51.41%

Tab. 4.2: Other Model Metrics

can't be roasted. Do your worst, RoastMe"). At least on the surface level, such titles give the appearance of confidence and "thick skin," as the interview participants discussed.

4.4.3 Linear Regression Model

Lastly, I iteratively trained a linear regression model in WEKA to predict the number of comments a given post can expect to receive using the filtered post data set of 4,710 posts (only including posts for which I was able to derive numeric and binary estimates of age and gender, respectively). Given the homogeneity I witnessed in the post title topic modeling, I did not incorporate features extracted from text mining to the data set. I applied the WEKA unsupervised attribute removal filter to the data; this machine learning algorithm weighs and eliminates less useful attributes from the model. After iteratively training and testing, I evaluated the model using 10-fold cross validation. Results are shown in Tables 1 and 2.

The model accounts for 86% of the variation in the data, and predicts that female participants will receive about 18 more comments than their male counterparts, suggesting that the RoastMe community may have an especially high interest in roasting women. Activity in mental health related subreddits also positively predicts the number of comments (roasts), suggesting that those with potentially unstable mental conditions may attract more interest and roasts from RoastMe participants. Vulnerable and mentally unstable participants who may not have been fully consenting to the declared play of RoastMe actually inspired *more* roasts; these roasts and their roaster-authors, by extension and in keeping with my interview participants' fears, may therefore no longer be protected and circumscribed by RoastMe's magic circle of play. At the same time, I also note that sadness (derived from the Emotion API) has a high negative coefficient, implying that RoastMe participants may steer away from insulting those who appear more sad in their photos. This could possibly indicate that RoastMe participants view roastees appearing to be sad as not displaying lusory attitudes or fully consenting to RoastMe's rules of play, and therefore avoid engaging with such posts. Although "score" (downvotes subtracted from upvotes on the original post) is also present as a feature in this model, it is

difficult to assign any real world value to the finding given the black-box nature of the measure, as described previously.

4.5 Discussion

In accordance with previous literature, although the specific behaviors enacted on RoastMe might be considered non-normative, participants highly value and strictly abide by the community-specific norms that govern interactions in the subreddit. Through a lens of play, this dynamic indicates that participants in RoastMe accept and abide by the declared rules of RoastMe play, voluntarily entering into a magic circle of play on RoastMe and abiding by the rules prescribed by that circle. These rules of play dictate and guide a specific set of behaviors within the subreddit, and in so doing, highlight distinct, relatively unexplored needs and desires that people may seek from a niche online humor and self-presentation community. Where dissociation from targets (roastees) would be expected, RoastMe roasters instead relied heavily on perspective-taking. In photographic presentation, rather than self-present to elicit approval or affirmation, roastees reported learning and growing from the harsh judgment and ridicule they experienced on RoastMe, applying insights gained from their ludic selves to improve their "ordinary" selves. Unlike previously studied communities that encourage subversive humor or potentially offensive behaviors, participants in RoastMe are also quite sensitive to the emotional well-being of targets (roastees), desiring to mitigate harm. They were motivated to stay within the confines of the magic circle of play defined by RoastMe, and concerned that the presence of participants who did *not* fully understand or consent to RoastMe's rules of play might invalidate the magic of that circle, and by extension, the legitimacy and acceptability of their participation therein.

Viewing RoastMe not as a bizarre, niche community, but as (a) a community centered around a specific declaration of play, and (b) a part of a larger confederation of communities displaying similar desires to engage in play with behaviors that are not always playful, directs us to broader-reaching design implications. The specific desires expressed by participants in RoastMe, a fairly large and still-growing subreddit, implies that there may be other venues, settings, and audiences for which beneficial aspects of RoastMe's design could apply. Specifically, I discuss how the RoastMe paradigm led me to develop and use the *research through evocative play* approach in my work.

4.5.1 Implications for Reddit and RoastMe

First, I consider how the findings can inform Reddit's community leaders and administrators as they develop rules and guidelines for Reddit as a whole, and for particular subreddits. As prior work has shown, bans on Reddit can be effective in curtailing hate speech on the platform at large [33]. For example, when participants from banned subreddits migrated to relatively more innocuous subreddits like RoastMe, they did not sour the community by blasting it with hate speech or shifting the norms of the community towards pure hate or violence; rather, it appears that they conformed to existing rules of play. In light of the slew of Reddit bans and policy changes over the past few years [97, 101], it appears that Reddit leaders are shifting their stances and cracking down on subreddits with blatantly hateful or violent goals. They may lose support of some of their users as a result, calling some to question whether Reddit still holds true to its values of free and emboldened speech and off-color humor [21].

Amid these crackdowns, Reddit leaders would be wise to address not just what they will curtail, but what they will support. Although some people may still have qualms with the types of joke-insults bandied about on RoastMe— including current participants in the subreddit— "comedy, not hate" is an ostensible improvement on intentionally aggressive acts of hate speech. As this study's findings reveal, there are ways in which we can further improve existing subversive humor communities like RoastMe to support freedom of speech and expression while simultaneously protecting and caring for its participants. The high level of acceptance of and abidance by RoastMe's rules of play speaks to the feasibility of using rules on Reddit in a manner that both continues to support play on the platform, but also limits and bounds that play. In expressing their concerns about harm generation in the community, participants also displayed desires for a more clearly defined circle of play in RoastMe in which other participants enter only voluntarily, with a lusory attitude. In this manner, new guidelines and rules to delineate and more clearly circumscribe play appears to be a step that RoastMe participants actually *want*; moderation does not need to be presented as a form of punishment, but rather, as an improvement upon and clarification of rules of play.

As I discussed, participants in RoastMe struggled with certain aspects of the community's design. My interview participants expressed uncertainty and anxiety about the level of harm generated in RoastMe. The commenting patterns revealed by the quantitative analyses also suggest that certain types of users— namely, women and those who may struggle with mental health issues— are disproportionately targeted. Professed values and realities in the community do not always align. For example, among the participants who placed high value on the reciprocal ("roast and be

roasted") nature of RoastMe, not all had actually posted as roastees themselves. Thus, the community space could be redesigned to allow for different levels of self-exposition and exposure to critique. Such design modifications could permit more fluid and comfortable levels of reciprocity. For example, the current play design of Reddit does not encourage many one-to-one interactions; directly reciprocal exchanges (where a dyadic pair could serve as both roaster and roastee to one another) might have appeal to subsets of the community, and allow for currently exclusive roasters who are hesitant about submitting themselves to a roast to engage more fully in the community.

We might also consider the role and value of anonymity in the forum. Here, we can envision a RoastMe-identical space that differs in one key aspect: all participants, including roasters, are photographically identifiable. To be clear, I make no claims that such a design would be "successful." However, by exploring such a design as a digital probe, we could both deepen our understanding of the current role of anonymity and exposure in RoastMe, and inform alternative community designs. Another digital probe could explore the change in dynamics if the Reddit sorting algorithm for displaying new posts and popular posts were weighted and reconfigured to discourage excessive roasting of certain demographic subsets of users (in this case, women and those who had posted on mental health-related subreddits). I further advocate for studying these digital probes through a *research through evocative play* study design, in which the revised play design is not seen as an end in itself, but a tool by which to further explore participants' relationships to anonymity within RoastMe.

4.5.2 The Start of *Research through Evocative Play*

The brand of subversive, offensive humor celebrated on RoastMe is already integrated into Reddit and its culture of participatory play [148, 147]. However, the types of interactions we see in RoastMe could function similarly in other domains that don't currently embrace playful sensibilities. Here, I propose using a *research through evocative play* framework to consider how declaring play in different contexts can reveal nuances of power dynamics and participants' relationships to those contexts.

In my research on RoastMe, I learned about power dynamics as they relate to the (uncertain) play of RoastMe. A surface-level understanding of the community might suggest that roasters call the shots and hold a higher power status in the community, while their roastee targets are the powerless butts of the jokes. Now, such an interpretation is not patently *wrong*, per se, as my quantitative analyses suggested that vulnerable populations such as those suffering from mental health issues may be disproportionately targeted in the subreddit. Still, the power dynamics

in RoastMe are more complex than this surface-level understanding would imply. Roastees reported gaining resilience as well perspective on themselves, benefitting from the harsh and honest judgment of the subreddit. Roasters also identified with and took on the perspectives of roastees in the process of roasting them, suggesting a more reciprocal relationship between the two groups. Implicit rules of RoastMe play further support reciprocity by encouraging participants to take part as both roasters *and* roastees. Then again, as my quantitative analyses showed, in practice, the "two-way" dynamic of RoastMe (participating as both a roaster and roastee) is not actually very common.

In their reflections on the play engendered by RoastMe, participants also revealed their relationship to sets of behaviors as they exist *outside* of the circle of declared play. For example, with regards to the use of harsh language and humor, participants expressed feeling inhibited in other offline and online settings, and saw RoastMe as an outlet or safe haven for expressing themselves in ways that are not considered acceptable elsewhere. Roastee participants contrasted posting a photo on RoastMe with self-presentation expectations in other spaces, where there is pressure to take oneself seriously, or where they are unable to gain honest, harsh feedback that might ultimately enable self-improvement.

In RoastMe, participants enter into a circle of play with a defined set of rules. In this declared space of play, participants are then confronted with sets of behaviors that are not always presented or viewed as playful in other contexts— namely, online photographic self-presentation, and harsh language— and told to interact with the behaviors as forms of play. We might presume that because RoastMe is a pre-existing place of play, participants enter the circle voluntarily and with a lusory attitude, rendering RoastMe a magic circle defined by the rules of play the community declares. However, as my mixed methods study of RoastMe showed, even when situated within a pre-existing playful context (RoastMe and Reddit), play declarations do not guarantee that the inhabitants of the space have fully consented to the declared play or will necessarily maintain lusory attitudes, rendering the circle circumscribed by the play declarations ambiguously magical. The uncertainty these play declarations introduce also calls into question the ethics of engaging in play with sets of behaviors in the context.

Following my research on RoastMe, I saw opportunities to investigate other contexts and behaviors through the use of play declarations. With RoastMe, I saw how declared play can be uncertain and ambiguous even when contextualized within a pre-established space of play. RoastMe participants specifically go to RoastMe in order to engage in the play of RoastMe, but even so, boundaries of consent and ethical play are unclear. The declared play of RoastMe impelled participants to reflect on their ambiguous roles as players in a context where the "magic" circle of

play is not clearly circumscribed or perceived universally as magical. As they used RoastMe's declaration of play to poke fun at other player-participants and criticize sets of behaviors such as photographic self-presentation, RoastMe participants simultaneously worried about harm generation, and the implications of playing in a space where it is not clear for whom the circle of play is magical.

Moreover, the rules of play declared by RoastMe draws to the fore the nature of power dynamics in the context, and by introducing further ambiguity into a set of behaviors, forces participants to grapple with their attitudes towards and roles within those sets of behaviors. However, by engaging in RoastMe play, participants are *not* commenting or reflecting on power dynamics or their relationship with aspects of a specific context (say, online photographic self-presentation on Facebook). In order to gain access to such types of reflection, we would need to situate the play declarations more directly in the original context. Rather than taking an existing space of play (Reddit) and drawing sets of behaviors into that context and declaring them as play (online photographic self-presentation and harsh humor), I contemplated what a study design that overlaid play declarations *atop* a context might reveal about power dynamics in the context. Such a study design would introduce further uncertainty into the circle circumscribed by the declared play in that participants would not be seeking play; instead, they would be asked to engage with play within a context that they might not normally view as playful, and wherein play might introduce additional ethical concerns or ambiguities. Thus began my formulation of the *research through evocative play* approach.

In the chapters that follow, and inspired by my study of RoastMe, I implement two *research through evocative play* study designs that declare play in power-ambiguous contexts for the purposes of better understanding the nuances, limits, and possibilities of those contexts. In Chapter 5, I begin with the context of Amazon Mechanical Turk and my work with the *research through evocative play* study design of Turker Tales.

Turker Tales: Revealing Workers and their Relationships to Amazon Mechanical Turk through Play Instructions

<div style="text-align: right">5</div>

[1]I first began more formally implementing *research through evocative play* by designing, deploying and studying Turker Tales, a Google Chrome that uses tangential play to encourage crowd workers to write, share, and view short tales as a side activity to their main job on Amazon Mechanical Turk (MTurk). Turker Tales introduces a layer of playful narrativization atop typical crowd work tasks in order to alter workers' experiences of those tasks without aiming to improve work efficiency or quality. This system of tangential play brought to light underlying conflicts (such as unfair working conditions), and provided a space for participants to reveal aspects of themselves and their shared experiences. Turker Tales allowed for critical reflection on the role of researchers, designers, and requesters, as well as the ethics of incorporating play into crowd work.

In Turker Tales, I thus simultaneously considered the potential for and implications of play in a crowd work context to support crowd workers, and using the *research through evocative play* methodology, introduced the use of play declarations as a research tool to draw out and study power dynamics in a context.

5.1 Introduction

In brick-and-mortar workplace settings, leadership often implements structures and activities to engage workers in non-work activities and build community [213]. Examples include office birthday parties, happy hours, free donuts on Thursdays, casual Fridays, "break rooms" designed to foster informal workplace social interactions, such as "water cooler banter" [64], and the ubiquitous ping-pong tables and game rooms of tech companies. When workers voluntarily participate in the activities, play and leisure at work have been shown to improve general employee happiness and well-being, including increasing positive affect and general job satis-

---

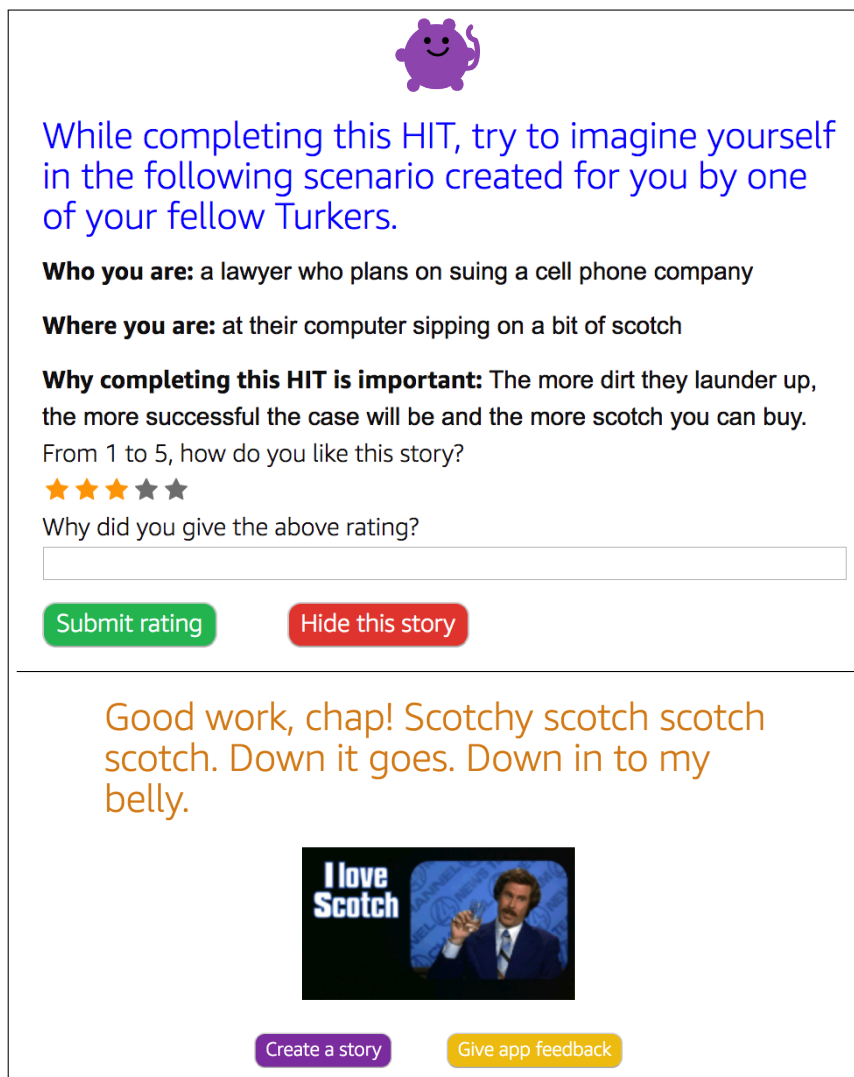[1]This chapter is based off work published at DIS 2019 [117].

Fig. 5.1: Turker Tales interface: Top panel shows a participant-created scenario another crowd worker could see during a categorization Human Intelligence Task (HIT). Lower panel shows the message and GIF they see upon completing the HIT.

faction, and decreasing the occurrence of emotional disturbances [177, 153, 112, 185]. Our existing models of work settings include notions and norms of structured, leadership-supported play and enjoyment.

By contrast, the tools we have developed as a research community for crowd workers, even those that incorporate elements of play, tend to focus on improving efficiency and productivity, building work-related skills, and addressing other "serious work matters" [35, 139, 133, 47]. Some crowd workers might choose to imbue play and entertainment throughout their workday in a variety of different forms, or engage in grassroots activities like forums that support varying levels of social connection [215, 132, 89]. However, as in other work settings, employers can greatly impact organizational culture and employees' overall work satisfaction through their management strategies (which can include mechanisms of play) [1]. Especially in light of broader cultural norms that may encourage overworking [25, 27], workers may sometimes require additional support in achieving work-life balance, realizing ways in which their own self-worth extends beyond their work performance or roles, and connecting to other people at work.

Research on teleworking has found that remote workers may be more prone to stress and depression [209], suggesting that crowd workers may need additional support structures and interventions. For example, survey research on crowd workers has suggested that at least a subset of crowd workers may suffer from depression and other mental health issues [116]. Moreover, crowd work platforms such as Amazon Mechanical Turk perpetuate unfair working conditions, wherein workers do not receive proper compensation for their work and have few opportunities to advocate for themselves (e.g. [96, 76, 105]). Incorporating play into crowd work does not solve such injustices, but, as I will show, could help surface workers' critical reflections on the ethics of crowd work, and proffers novel paths through which we can redesign crowd work.

As many academic researchers employ crowd workers or make recommendations as to how to structure crowd work, we are essentially part of the distributed management team for crowd work. As such, we might be inclined to study and design for crowd work in domains such as worker productivity, skill levels, and working conditions, all of which a solid body of research has explored, e.g. [30, 187, 56, 139, 105, 3]. In contrast to these prior approaches, I draw from literature in play theory to consider how designing for and declaring *tangential play* in crowd work can provide insights into the crowd work context and direct us towards new ways of designing to support crowd workers. I refer to this play as "tangential" because it is not structured with a goal of supporting work productivity or efficiency. Much like the break room setting in an office [185], the goal of play here is decoupled from the productive tasks. Given that storytelling is a natural inclination of humans, and a primary way

in which we process the world, embed our lives with meaning, and connect with one another [109, 49, 42, 54], I choose to ground my approach specifically in narrative play. To this end, using primarily qualitative approaches, I developed and studied a system called Turker Tales to explore how crowd workers respond to embedded declared play and anonymous social interactions via storytelling as part of their daily crowd work tasks.

To understand the design space and validate user needs, I employed the user research method of speed dating [50] (N=12) with Turkers (crowd workers on Amazon Mechanical Turk), and performed a pilot test of a storytelling play concept using receipt transcription tasks (on MTurk, tasks are called Human Intelligence Tasks, or HITs) (N=150). Based on these insights, I developed Turker Tales, a Google Chrome extension that allows workers on MTurk to create and share short tales with one another while completing HITs. I released Turker Tales in a one-week field deployment (N=171), collecting 1,096 tales and 1,527 ratings. I also performed analyses of the story content, both in the pilot and the field deployments, and gained insights into how declaring play within the crowd work context through Turker Tales allowed workers to share aspects of themselves with one another while maintaining anonymity, share critiques of working conditions on MTurk, and communicate about cultural and political phenomena. The insights revealed through Turker Tales allow us to explore tensions in the crowd work design space, including researchers' responsibilities as part of the crowd work system, and the ethics of designing for paid or unpaid play on crowd work platforms.

5.2 Initial Exploration: Speed Dating

I used speed dating [50], a design method that lies between sketching and proto-typing, and allows for rapid exploration of design concepts and their contextual dimensions, to better understand whether my design ideas fit MTurkers' needs, and where I might have inaccurate understandings and misplaced assumptions. I recruited 12 participants on MTurk to participate in a storyboarding feedback session, conducted over Skype via either voice or video chat (per the participant's choice). Each session, which I audio-recorded and later transcribed, lasted from 16-40 minutes, and I compensated each participant $8 (average of $21.65/hour). Here, and in all stages of the study when working with participants, I required that participants be 18 years or older and complete an IRB[2]-approved online consent form.

---

[2]Institutional Review Board

I presented participants with six different storyboard scenarios to explore my design concepts; four dealt with different dimensions of a playful intervention, and two explored public-facing profiles and anonymity. I wanted to better understand the extent to which crowd workers value anonymity, which I conjectured may benefit requesters more than workers, allowing requesters to maintain workers as faceless masses rather than individual employees [106]. For example, I asked participants to imagine different scenarios in which they might automatically share their crowd work accomplishments or activity with requesters and/or other Turkers, and probed around how much they would be willing to reveal about their personal identity or career path in such scenarios. Additionally, I asked participants to imagine themselves in different narrative and non-narrative play scenarios in order to assess the potential value of play, including storytelling play in crowd work, as a mechanism to escape from daily realities and improve overall user experiences while using MTurk [19, 31, 34, 171].

In keeping with the goals of speed dating to understand the contextual space— validating needs rather than design concepts— I displayed potential scenarios, but not actual user interfaces, and directly incorporated and probed around my original understandings. I conducted a thinkaloud with each scenario, asking participants to consider whether they could see themselves in the scenario, and what aspects resonated or didn't resonate. See Figure 5.2 for the storyboard scenario that most closely relates to the further steps I did. I qualitatively annotated the transcripts using inductive, iterative coding to extract key takeaways.

I found that my assumptions about anonymity on MTurk were off-target; participants highly valued their anonymity and privacy on MTurk, and did not want to engage in any activities that displayed their accomplishments on MTurk. The value crowd workers place on anonymity is underscored by the fact that the speed dating sessions were conducted over Skype, thereby revealing participants' voices and/or faces. As such, I may already have been speaking with a subset of Turkers that could be relatively less concerned about anonymity, and yet these participants still held strong opinions about maintaining anonymity. Explaining that the structure of MTurk lends itself to certain "dog-eat-dog" mentalities, participants said they didn't want to risk exposing information to other Turkers that might help them take over their slots in well-paying HITs. Even forums may not be entirely transparent or altruistic, as many forum members safeguard information about especially desirable HITs. Rather than seeing other Turkers as familiar colleagues or peers, it appears that some Turkers may view them more as competition. For example, p10 lamented how so many Turkers use scripts to grab HITs, exclaiming, "I don't know how they grab the stuff so fast, I really don't!" She said that although many Turkers boast in the forums about bonuses or large volumes of completed HITs, they don't reveal "who they did the work for, and how they got the work." This suggests that gamified approaches

You can transform HITs to make them more playful for others who are also using the AMT Play plug-in. For example, you can add story elements to specific HITs, or create HIT challenges.
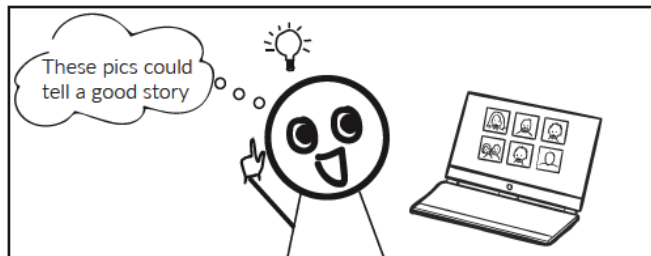
Other Turkers use what you created to complete HITs in a more enjoyable way, and feel more motivated to complete more work on AMT.
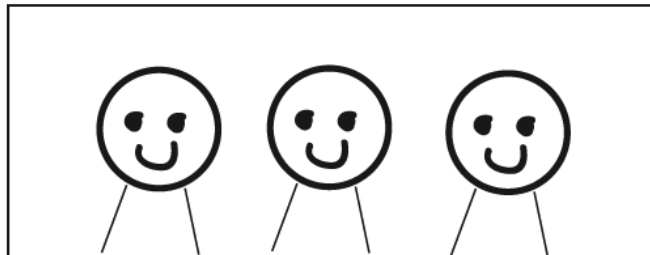
You get positive feedback from other Turkers. You feel proud of the creative work you did on AMT, and find that colleagues and employers outside AMT also find this work interesting and impressive.
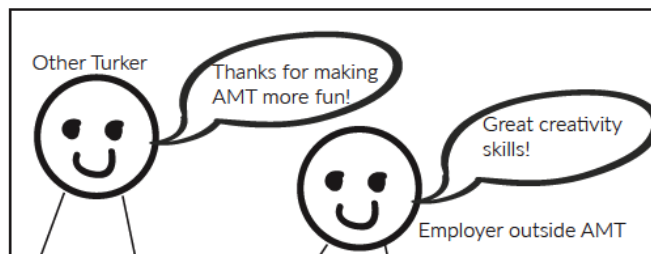
Fig. 5.2: A storyboard discussion prompt for speed-dating.

hinging on social comparison and rivalry may not be ideal for MTurk settings, as competition may already be a divisive factor in the community.

Meanwhile, the four play scenarios confirmed that Turkers may be open to "play-at-crowd-work." For all the participants, feelings of boredom resonated; I discovered that "dead time," or periods in which well-paying or otherwise interesting HITs become few and far between, can be an issue. For example, p02 said he would "without a doubt" be interested in a storytelling based plug-in, saying, "Sometimes on the weekends, it's kinda slow. And I'll be trying to get on sometimes, and they'll have some pretty bad jobs, that I kinda just do because I wanna make a couple of dollars." Moreover, when I pushed on participants' boundaries by asking whether they could still see themselves engaging if the play actually *decreased* their efficiency, participants often answered ambiguously, not ruling it out entirely, but rather, indicating that they make tradeoffs between enjoyment and efficiency. As p01 explained, "Depends how much longer, because I am very calculative with my time... If it floated down too much, to where you'd make $6 an hour, I wouldn't do it." In this way, for the Turkers I spoke to, efficiency is a relative, not an absolute, concern.

Social connection also emerged as a key desire. Several participants initially responded that they could not picture themselves engaging in play on MTurk because they preferred to treat "work as work." However, they changed their tune when I re-framed play artifacts as being created and shared by other crowd workers. For example, p06 initially said she preferred to work only, without any element of play incorporated. But then when I prompted her to consider sharing the tales she had created with others, she became very interested, and said she didn't care about the compensation, explaining, "It doesn't really matter... To put something out for somebody else; I think it'd be fun." She said she had even been considering posting as a requester (though wasn't sure about the costs) because she wanted to learn more about other Turkers and whether they had shared experiences outside of Turking.

In sum, through speed dating, I learned that crowd workers value anonymity, may be wary of sharing information with other crowd workers, make tradeoffs between earnings and enjoyment (especially during dead time), and value social connections with other crowd workers. These findings informed the subsequent design of Turker Tales, leading me to develop a storytelling play declaration design to allow for social connections while simultaneously maintaining worker anonymity.

5.3 Methods: Pilot Test, Design, and Deployment

5.3.1 Pilot Testing Method

I next moved forward with a pilot study (N=150) of a storytelling application concept on MTurk using a mock batch of receipt transcription HITs (receipts obtained from [62]), which are widely available on MTurk and often low-paying (most paying pennies per receipt transcription), and thereby potentially a task workers might encounter during "dead time." I instructed participants to provide another (hypothetical) Turker completing the same set of HITs with a character, a reason for why they are doing the HITs, and closing "success" and "fail" messages. I paid participants a base compensation of $3 for participating in the approximately 10-minute study, and a bonus payment of $0.04 for each receipt transcription they completed (approximate hourly wage of $18, though self-reports on sites like TurkerView were higher). I analyzed the results of the pilot, which I will present in tandem with the full deployment results, using a primarily qualitative approach of open, inductive, iterative coding; I conducted any statistical analyses in R. As will be discussed in the results section, participant feedback from the pilot testing helped me refine and revise the *research through evocative play* design for the actual system I deployed.

5.3.2 Design of Turker Tales

After refining and iterating on the design, I next developed a Google Chrome Extension called Turker Tales that allows crowd workers on MTurk to see short scenarios or stories (tales) that other Turkers have created while doing the same or similar HITs, as well as to compose and share their own tales. Images of the interface can be seen in Figure 5.3. To store data on the backend, I used Google's Firebase Firestore [86]. Below, I highlight some of the main principles that guided the plug-in's design.

I maintained the general storytelling structure, but modified the prompts for brevity, and removed the dual nature of the closing message options ("success" versus "fail"). Based off participant feedback, I also included a more playful, visual feature so that the closing message would be automatically represented to users as a GIF (using the giphy API) after they had completed a HIT. I used a smiling, playful purple creature with a tail ("tale") as the extension's icon.

In Turker Tales, Turkers submit stories anonymously shared with other Turkers doing the same or similar HITs. Whenever a Turker who has the extension installed is working on a HIT, if there are any stories associated with that HIT, they will have the option to click to see a story. Turkers may submit their HIT-associated stories either while the HIT is in progress, or directly after the HIT has been submitted.
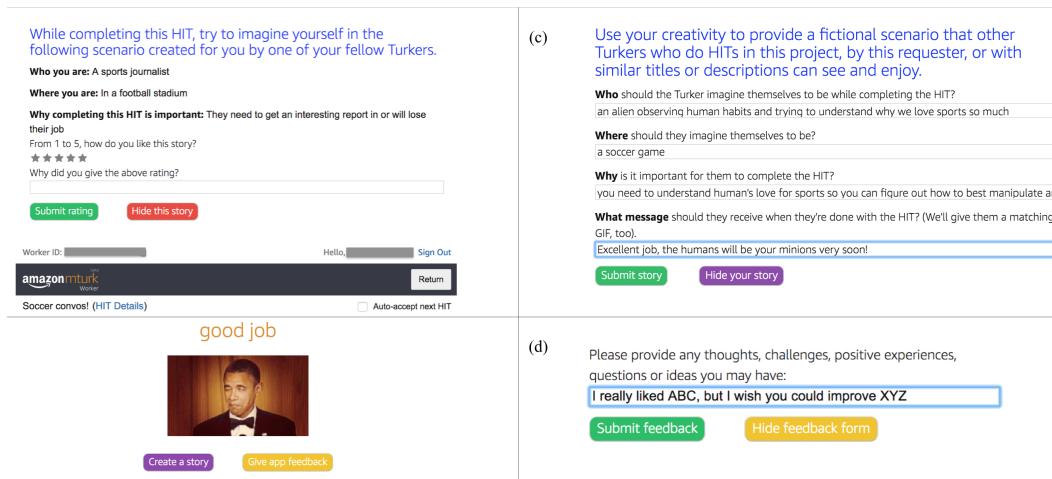
Fig. 5.3: Turker Tales interface. In (a), we see the screen for viewing and rating a tale. As with all aspects of the interface, this appears before all other elements in the HTML DOM; (b) shows a sample GIF received after completing a HIT; (c) is used to submit a tale; (d) is used to provide general feedback on the plug-in.

I maintained privacy where possible by, for example, using a content script in the extension, meaning I only accessed users' information when they were directly working on HITs. In keeping with general UI heuristics, I aimed to minimize distractions. For example, even if a story is available for view, if the participant does not choose to click on it, then I do not show them a closing message or a GIF upon completing the HIT, and all elements of the interface appear directly at the top of the current window rather than as a potentially distracting pop-up.

I took a loose, simplistic approach to story-HIT matching to maximize the likelihood of Turkers seeing a story, first prioritizing HIT match, then requester match (as many requesters may post HITs that are quite similar to one another), and lastly, using keyword match from titles (I used Python's Natural Language Toolkit to assist in removing stopwords and lemmatizing text). I gave highest priority to HIT (or HIT batch) match, and lowest priority to keyword match. If there was at least one story in a given group for a HIT, then that story would be displayed to participants; if more than one matched, a story was chosen at random (again, first choosing among HIT-matches, where available). To give an example from the dataset, one participant wrote a story about a magician while completing an audio transcription task posted by Requester X (maintaining requester anonymity) to determine whether the speaker is saying "Yanny" or "Laurel," a viral auditory illusion that has also been studied academically [226, 179]. Other users of Turker Tales that happened to complete the same Yanny or Laurel HIT could see this story. In addition, other HITs posted by Requester X, which are typically audio transcription HITs, could be matched with this story. Lastly, workers completing other audio transcription HITs not posted by Requester X might also be exposed to this story through the keyword matching process.

Because the newest users of the extension would be unable to see stories created by others, I also seeded the application with 93 stories selected at random from the pilot intervention. For the seed stories, I manually assigned keywords such as "receipt" and "transcription." Because this would mean that only those doing receipt transcription HITs (or similar) would see the seed stories, I slightly modified some stories to make them more generic in nature, and assigned them more generic keywords such as "categorize" and "pay." I also allowed feedback mechanisms, both for the stories themselves and the application, overall. Participants could rate the quality of any stories they viewed on MTurk on a scale from 0 to 5, and submit a comment about the story. They could also provide optional, open feedback on the extension.

5.3.3 Turker Tales as *Research through Evocative Play*: Play Cues

I designed Turker Tales in conversation with the four play design cues on which I have consistently focused while developing *research through evocative play*: (1) direct declarations of play; (2) perspective-taking; (3) harsh humor; and (4) the creation of shared artifacts.

1. **Direct play declarations.** In Turker Tales, I as the researcher declare play in the MTurk context. Participants, by contrast, are *told* to engage in a play intervention as part of (and while performing) other work on the platform. In my study of Turker Tales, I therefore had two goals. Using the *research through evocative play* approach, I studied what declared play reveals about power dynamics in a context. In addition, I used the declared play of *research through evocative play* to consider the possibilities and ethics of designing for play in a crowd work context.

2. **Perspective-taking.** I incorporated perspective-taking by designing 'imagine yourself as" template scenarios for workers to create and consider while completing HITs.

3. **Harsh humor.** Although I didn't design directly for critical humor, in my analysis of participants' responses to Turker Tales, I take special note of the use of critical humor, and how such humor is used to criticize power dynamics in the context.

4. **Creation of shared artifacts.** The tales created in Turker Tales are shared anonymously for other workers in the context to view while completing tasks on MTurk.

5.3.4 Deployment and Analysis Methods

I deployed Turker Tales by posting a HIT on MTurk, informing participants that they must use Google Chrome in order to qualify for the HIT and be willing to keep the extension installed for one week; 171 participated in the study by downloading Turker Tales. I paid participants $1 for downloading the extension, and bonuses of $0.10 for each story they submitted and $0.05 for each story rating they submitted, up to a maximum of $10. With approximate download times of 3 minutes, this translates to a roughly $20 hourly wage, though, as before, self-reports on sites like turkerhub were higher. Note that I made it clear that bonus activities were optional to receive the base payment. Originally framing Turker Tales as a precursor of a freely available Chrome Extension, I reasoned that in a true-in-the-wild experiment, participants would not receive any payment at all, but I still wanted to provide some compensation in keeping with the expectations of the MTurk platform. (Note: I further discuss the ethics and tensions around such unpaid play in this chapter's Discussion section). I encouraged feedback via the extension, but did not explicitly compensate for feedback.

I collected data through the application for one week. To analyze the results, I used Google Cloud services to help transfer the data from Google Firestore to Google BigQuery [85]. I used open, iterative coding (in Excel) to classify the content in both the pilot and deployment stories. For the stories from Turker Tale's deployment, I also used feature extraction in LightSide [150] based on unigrams and lemmatization (word-stemming), and ignoring stopwords and punctuation, to help with initial categories and groupings, and then iteratively and manually coded themes from there. In categorizing tales, I ultimately found qualitative analyses to be more informative than topic modeling approaches such Latent Dirichlet Allocation [18] or biterm topic modeling [225], and will thus focus on presenting the qualitative results in what follows.

5.4 Results of Pilot and Field Deployments

Overview of Engagement

In the one-week period in which I officially deployed Turker Tales, I received a total of 1,096 stories produced by 110 participants, and 1,527 story ratings from 113 participants. A total of 132 participants "lurked" on the extension in some capacity by, at the very least, clicking on "See story" when the option was available. On average, active participants submitted 9.96 stories, with a standard error of 1.45, a

median of 3, a mode of 1, and a max of 62. Additionally, 32 participants voluntarily provided 56 notes of feedback on the extension.

I manually labeled all the HITs based off project titles, and found that 66% (728) of stories in the field deployment were submitted while participating in surveys, quizzes, or studies; 10.8% (118) in reading or writing tasks (such as summarizing, describing, and reading comprehension); 6.8% (75) dealt with stories or games; and 6.6% (72) involved tagging, labeling, categorizing, or transcribing. Although I initially conceptualized Turker Tales to be used for more repetitive, "boring" tasks, such as receipt transcription, I instead found that participants were heavily engaged in reading, writing, creativity-oriented, and higher-level processing tasks. It is likely that Turker Tales, given that it features playful interactions, involves reading and writing, and is part of a research project, appealed more to Turkers that choose similar HITs on MTurk. The difference in tasks chosen, as I will discuss, may also have impacted the content and tone of stories submitted in the field deployment (primarily surveys and reading/writing) versus the pilot (receipt transcription only).

Stories submitted in the pilot were longer than those from the field deployment, with pilot stories averaging more than three times the length of field stories, at 369.7 versus 110.8 characters. This difference was statistically significant as measured by a Welch 2-sample t-test, with a p-value of 0.00. It could be— as one participant suggested in feedback— that the pilot participants were generously compensated for submitting their stories and therefore spent more time and effort on them ($3 for participating in the study), whereas the deployment participants were submitting "10-cent stories." However, longer stories aren't necessarily more engaging, nor are they practical for types of play that can be done during the work day. Average ratings for stories in general were above the threshold of neutral, indicating general enjoyment of Turker Tales. Moreover, on average, the shorter stories produced in Turker Tales received higher ratings than the longer pilot study stories I used to seed Turker Tales, with field stories having an average rating of 3.6, and pilot stories, an average of 3.3 (though this difference is not statistically significant). Thus, if anything, participants appeared to prefer the shorter stories slightly more, giving more encouragement to the scalability and viability of play through Turker Tales.

I note that "spamming" Turker Tales (e.g., by typing in nonsense text) could have been profitable, as I did not place any stipulations on the content of stories submitted. Instead (as the topic categories and story examples show, see Table 5.1), workers engaged playfully even though their earnings were not contingent on playfulness or creativity. In addition, 32 people in the field deployment voluntarily provided un-compensated feedback on Turker Tales via the built-in feedback form (56 comments). Of these, 17/32 people (24 comments) showed direct enjoyment and flow, e.g. "I really like this. It makes you think and laugh alot [sic]" and "That was actually kind

Theme	Pilot	Field	Example (from field deployment of Turker Tales)
Occupations	21% (31)	30% (324)	"A world famous musician [at] Carnegie Hall" about to receive "the biggest award of their career!"
Cultural/famous references	5% (8)	16% (177)	Annie, who is with MJ [Michael Jackson] "to prove that yes, she is okay." *(song reference)*
Animals	5% (8)	12% (129)	A cat flying in the sky near dinner time, told to "Have a cookie."
Fantasy (not otherwise stated)	0	7% (75)	"The opposite gender in a bathtub full of water. If they don't answer they will forever be changed to the opposite gender."
Everyday	5% (8)	7% (72)	"A call of duty player//On the couch//To get to prestige 10."
(MTurk) Critical narratives	12% (18)	6% (71)	"A masochist//Inside an Iron Maiden// They really enjoy tedious activities that lead to dead surveys.// (message): Sorry pal. Better luck next time."
Political references	0	6% (70)	"Donald Trump//In Russia//Because it will garner him praise and admiration from Putin.//(message): Good job, comrade! You are my #1 puppet!"
Objects	0	5% (69)	A pen who "wants to be chosen even though they are green" and later told,"You're the best pen!"

Tab. 5.1: Eight most common themes in the field deployment.

of cathartic." Further, 12 people (17 comments) offered ways to improve Turker Tales, e.g. "a way to track how many stories you have written and rated." Only two participants directly criticized the "play at work" concept behind Turker Tales as impractical and unwanted, e.g., "Nobody has time for this." In addition, 21 workers enjoyed the extension so much as to remain active on the extension in the three weeks after the study period ended, contributing over 300 additional stories. This suggests that at least for a subset of workers, the concept of Turker Tales holds appeal as a purely voluntary playful intervention.

The top eight most prevalent themes in the field deployment of Turker Tales are shown in Table 5.1. As shown, critiques of MTurk working conditions weren't as prevalent in the field as in the pilot, and superhero/savior themes were rarer in the field than the pilot. Meanwhile, politics did not arise in the pilot at all, whereas the field deployment featured 13 separate references to Donald Trump alone. In the following discussion of emergent dynamics in Turker Tales, all examples of tales come from the field deployment, unless otherwise stated.

Peeks into Turkers' Personalities, Thoughts, and Interests

Stories grounded in the everyday were quite common, as Table 5.1 attests. Whether it was a "a middle aged man who is on his lunch break from his engineering job," "a dad trying to get his daughter to wash her own dishes," "a person in their underwear finishing an episode on the couch," or "a senior citizen at home massaging [their] knee with a cold compress," it was at times difficult to ascertain to what extent the characters were autobiographical. By sharing aspects of everyday life and work, some participants may have shared aspects of their lived realities. This may have held true for more personal or revealing stories, such as a "heartbroken divorcee//On facebook checking on their ex." Likewise, the person who submitted, "Depressed male//In the bedroom eating popcorns//So as to understand their depression//(message): It's so easy to be under stress these days" might have personal experience with depression.

A small subset of stories in Turker Tales (8) waxed reflective as they reached the closing message, indicating the train of thought or belief held by the Turker-creator. For example, a story about a fisherman hauling in shellfish ended with, "Yes, he may be imperfect but we humans are all in this together and no one is perfect. Keep improving every day." Similarly, a short tale about a commercial pilot "waiting on the runway for take off instructions" ended with, "How different we all are even though we may come from the same place; the experiences are all unique and rich in scope." These short words of wisdom provided insight into the mindset of Turkers.

Many stories also incorporated fantasy and absurdism; this became even more present in the field deployment as compared to the pilot, and I conjecture that the fast-paced nature of the tasks (done while or between completing other HITs) may have encouraged participants to write the first things that came to their minds, leading to more fantastical or surreal scenarios. For example, in addition to the many animal-based stories, I saw quirky examples of objects as characters, including a cup of coffee, a banana, a strawberry, a leaf, a zipper, pizza, and an ice cube. The presence of the purple creature icon ostensibly inspired several purple characters, such as a purple people-eater in a pool of marbles who receives permission to eat, or "a one-eyed, one-horned, flying purple people eater flying above the city," holding a phone and looking to buy "a second eye/horn and a nice salad." Trends of fantasy and surrealism may also have been self-reinforcing; as participants saw absurdist and fantasy tales, they may have been inspired to create similar scenarios. Regardless, examples of fantasy give small glimpses into Turkers' thoughts, mindsets and creativity. Again, I highlight that the HIT itself did not require creativity or playful engagement as a prequisite for payment; participants voluntarily engaged playfully in Turker Tales. Participants' assumptions of lusory attitudes suggests that

they viewed the rules of Turker Tales as defining a porous magic circle existing both within and apart from the crowd work context.

Criticism of MTurk Working Conditions

In the *MTurk criticism* category, I separated out stories that specifically criticized or satirized either Amazon Mechanical Turk, requesters, or specific HITs. I first note that critique themes were more common in the pilot than in the field deployment, perhaps because I tied the pilot activity to receipt transcriptions. For example, one participant in the pilot study called the Turker "a sucker" who is "locked in their parent's basement desperately trying to crawl out of debt, one underpaid Brelig at a time." Here, the participant refers to John Brelig of the company InfoScout, a requester notorious for posting low-paying receipt transcription HITs and using techniques to get his HITs to show up at the top of the page [6]. For the "success" message for this sarcastic setup, the participant wrote, "Congratulations! You did so well, your mom decided to get you a $5 fill up from KFC[3] instead of the usual storebought tendies [chicken tenders] and has added a bonus to your good boy point fund." By speaking to the (imagined) Turker in a tongue-in-cheek, patronizing tone, this participant makes a political statement: workers with self-respect should not participate in such low-paying HITs. I posit that in the pilot test, many participants assumed that I myself was planning to post receipt transcription tasks in the future, and was seeking ways to keep wages low while still retaining workers. In this way, they playfully subverted the (perceived) activity itself by interrupting the flow of power; the activity itself became a way to discourage other (hypothetical) workers from participating in such a low paying HIT rather than a task to increase other workers' enjoyment of the HIT. By contrast, in the field deployment of Turker Tales, the extension itself was clearly separate from the intentions of the specific HITs, which Turkers had chosen according to their own predilections and goals.

Still, 6% (71) of the tales in the field deployment focused on MTurk criticism; this is not a trivial percentage. Several participants used dark comedy to highlight the low wages on MTurk and wide income gap between Amazon leadership and Turkers. One participant told Turkers to, "Pretend you're Jeff Bezos//Sitting in your pile of cash.// You need to make your pile even larger. //(message) Woohoo I'm even richer!" Another worker, taking an unpaid qualification survey, ended their story with a sarcastic, "Congratulations, you earned $0.00!!!" Often, Turking (completing HITs on Amazon Mechanical Turk) for pennies is wryly linked to basic survival, as in, "A very poor person, scrounging for pennies//Digging through the 'couch' of this survey for pennies //You need those pennies to buy food. //(message) Success, you found

---

[3]Kentucky Fried Chicken, a fast-food chain

$.03!" Similarly, a Turker trying to buy new shoes for their child is sardonically told, "Good job, you have earned 1 cent." References to "slaving away on mturk" further draw parallels between MTurk and slave labor. Another participant encourages other workers to think of themselves as Satan, ruining the results of a survey "just because," which we could also see as a form of rebellion against requesters on MTurk. Other stories from the field deployment do not explicitly refer to MTurk, but implicitly critique MTurk working conditions by highlighting themes of poverty and mistreatment. For example, a story about a sweatshop worker in China that will be beaten and starved if they do not finish the job draws a dark parallel to MTurk work, and a story about a scam artist "in a scamming room to scam everyone out of their time and money" could be a response to unethical, low-paying requesters.

Political and Cultural Communication

The field deployment of Turker Tales also evoked engagement around cultural and political themes, such as stories featuring famous political figures; many of these incorporated humor or absurdism. For example, a North Korean diplomat goes to Disneyland, "because it is necessary for world peace and harmony.//Say no to war" and Russian president Vladimir Putin plays ping pong with Chinese president Xi Jinping at the gym, because "You need to know even presidents are simple humans and they can do what we do." The multiple stories centered around Donald Trump often poked fun, such as a story about Trump's spellchecker who has an important job to do because Trump "can't spell//LOL." While many of these stories were ostensibly unrelated to the HIT's content, some were in conversation with the political content of the HIT. For example, while completing a HIT entitled "NEWS," a participant wrote about "a hybrid immigrant in a car being chased by ICE" and commented, "Immigrants help our country so why is everyone hounding them." Turkers used stories to communicate their views on politics and world events, and to participate in a distributed sharing of ideas around politics, often from a humorous angle. I conjecture that the absence of political commentary in the pilot may stem from the difference in types of HITs. In the field deployment of Turker Tales, participants engaged in many surveys and questionnaires that may have led to more contemplation and reflection about themselves and society at large than receipt transcription tasks did.

Participants' tales also incorporated figures and themes from popular culture, such as television, movies, and the Internet. For example, I saw multiple Harry Potter characters, four Tom Cruises, a Jerry Seinfeld saying "Yadda yadda yadda," a character from the indie game "Papers, Please" and other characters from television and movies, video games and animé. In a story about "Brother Orange," a participant

references a 2014 Internet phenomenon in which a stolen iPhone and a series of mysterious orange-tree selfies sparked a strange friendship between a man in the U.S. and a man in China [206]. While waiting backstage before going on the Stephen Colbert show, Lady Gaga completes HITs on MTurk "to calm down" (naturally). Through these stories, participants not only reveal their personal interests in certain media, but also communicate with other Turkers by evoking popular culture, thereby engaging in broader cultural communication.

In sum, Turker Tales effectively engaged workers in (compensated) storytelling-based play; participants showed evidence of lusory attitudes in the ways in which they engaged in the extension. This play allowed them to reveal aspects of their mindsets, personalities, lived experiences, or sense of humor to one another. In this way, Turker Tales highlighted ways in which Turkers are unique, creative, funny, and quirky individuals— a far cry from invisible "cogs in the wheel." Moreover, some participants also playfully subverted and co-opted Turker Tales in ostensibly productive ways for grassroots activism and criticism of MTurk working conditions, using the declared play as a springboard to criticize power inequities on the platform. Lastly, Turker Tales allowed participants to engage in political and cultural communication.

5.5 Discussion

At a high level, the development and evaluation of Turker Tales show that crowd workers are amenable to engaging in forms of play that are not designed to improve the efficiency or quality of their work, given that participants showed evidence of adopting lusory attitudes in their engagement with Turker Tales. Workers did receive small payments for submitting tales, but I note again that compensation wasn't contingent on *playful* engagement, and yet the results indicate playfulness. Perhaps this is not altogether surprising given human beings' proclivities towards play and storytelling, but because it is not an assumption that the research community often makes when designing for crowd workers, it still bears noting. Crowd workers expressed themselves through and subverted the tangential play presented to them through Turker Tales in intriguing and unexpected ways. The results of the study reveal potentials for using play as a part of a research method— *research through evocative play*— to increase understandings of a particular context. It also provides insights into how we can support crowd workers as individuals and users of a digital work space by designing technology to serve playful and multi-layered purposes in crowd work. Below, I present the potentials and caveats of using tangential play (a) as a way to encourage shared presence and social curiosity in crowd work; (b) as a paid activity to support crowd workers; and (c) as a research tool to reveals

aspects of a context and the power dynamics it encompasses, especially pertaining to criticism of a context.

Shared Presence and Social Curiosity through Tangential Play

In a platform where distrust and concerns about fending for oneself may eclipse desires to connect with one's peers, as some of the speed dating sessions suggested, interactive play such as that which Turker Tales engenders could creatively allow workers to share aspects of themselves in a safe, anonymous setting. By using methods of shared presence or digital traces, Turkers can get to know other Turkers in spaces in which they would otherwise remain invisible to one another. For example, Turkers may be reluctant to communicate with other Turkers about especially well-paying or otherwise especially beneficial HITs, lest they lose their cherished spot to another Turker. Although the challenges of physical distance may inhibit connections in ways that no single intervention can likely address, anonymous, playful connections could be a way to start to break down those barriers and thereby protect workers from some of the isolating risk factors that remote work carries. Platforms that focus on activism, like the inspiring example of TurkOpticon [106, 104], can increase worker visibility, but may not be especially instrumental in revealing Turkers to one another as quirky, funny, original, thoughtful, and creative people, or in engaging them in (asynchronous, distributed) conversation about social and political happenings. Embedding play in crowd work may be one way to connect crowd workers and build community in new, mutually beneficial, and socially rewarding ways.

Paying for Tangential Play

We can also consider the value of autotelic play ("play for play's sake") HITs, compensated by requesters. In my design of Turker Tales, I envisioned a freestanding, freely available plug-in that Turkers could voluntarily access while working; in other words, unpaid play. However, in order to ensure research equity, the form of play I implemented in the study of Turker Tales was compensated. In analyzing the results of the study, I reflected that there could be value in viewing the compensated research model not as a proxy for a fully voluntary and unpaid system, but as a potential model for directly compensated, requester-driven play within crowd work. For example, we could envision compensated, tangential play as an opt-in service that requesters pay for; workers that complete tasks for a requester that opts into the service are then invited to participate in additional, fairly compensated HITs that engage them in play that is tangential to their other work on MTurk; the play should

also not directly benefit the requesters. For example, metrics of engagement in the play could even be kept completely occluded from requesters to solidify this divide and reduce the tendency towards capitalist or "productive" play and gamification. These "play HITs" should engage participants, strengthen identity, and promote social bonding, while explicitly *not* attempting to improve crowd workers' fluency or efficiency in performing HITs or otherwise functioning on MTurk, differentiating them from crowdsourcing games that perform productive work for companies or organizations, e.g. [66, 92]. This might initially seem impractical or even radical. However, I argue that compensated play can and should have a space in crowd work. Hourly jobs provide paid breaks, salaried jobs provide paid vacations, and both may provide other small forms of play and leisure (e.g. free snacks). Likewise, requesters could show their appreciation to their distributed employees by compensating them for play and leisure performed on MTurk, even if it serves no practical purpose in terms of the work the requesters are trying to accomplish.

Still, compensating tangential play might *reduce* a tendency towards capitalist play [53], but it does so within the confines of a capitalist system. In fact, compensated pay on MTurk could even further entrench capitalist ideologies. Here, we can draw parallels to co-located workers in corporations such as Facebook and Google, that offer on-site laundry and cafeteria services, free childcare, and free egg-freezing services; such "perks" may seem supportive of workers' well-being on the surface, but also serve to make it easier for workers to devote more of their available time to laboring for the corporation. Thus, any initiatives incorporating play— including compensated tangential play— should be considered critically within the larger context of ethics in crowd work. My work with Turker Tales initiates an exploration of the design space, considers the possibility of compensated tangential play, and encourages further exploration of the concept of such play in crowd work, but does *not* directly advocate for uses of compensated tangential play in crowd work.

Developing *Research through Evocative Play*: Using Play as a Research Tool

In addition, from research perspective, designing tangential play for crowd work could be useful in and of itself, regardless of the "success" of the play itself. With Turker Tales, I start to more formally lay down the foundations of *research through evocative play* as research approach. Through designing for play, I revealed opinions of crowd workers about crowd work and specific HITs without explicitly soliciting them. Although I did not specifically design Turker Tales to be paternalistic, I found that when participants appeared to interpret more paternalism in the form of play presented, the resulting forms of play provided richer insights into the crowd work

context. I thereby put forth another somewhat radical implication, which is that further drawing out and actively *encouraging* interpretations of play as paternalistic could be especially useful as a research method to better understand power dynamics in a context.

For example, in the pilot study, many participants viewed the intervention as a paternalistic action. They assumed our research team sought ways to engage workers through play in a task that some participants felt was demeaning and unfairly compensated. As a result, over 10% of the resulting pilot stories centered on themes of imprisonment, poor treatment of workers, and satirical scenarios suggesting that Turkers who engage in low paying tasks like receipt transcription lack self respect. Meanwhile, in the field deployment, where the play was viewed as more voluntary and in line with their own choices and predilections, such themes of criticism were less prevalent. In co-located work settings, we might imagine criticism of play-at-work as taking the form of, say, employees rolling their eyes to one another or snickering behind the manager's back about "mandatory fun" activities. In the digital space of Turker Tales, criticism appeared to take on a much more meaningful and radical form, with participants making implicit and explicit statements about requester ethics and norms of behavior that self-respecting Turkers should adopt.

In this way, play aligned with powerful stakeholders in crowd work settings could serve as a way for crowd workers to simultaneously engage in and subvert the interactions, co-opting it into bottom-up play that serves as a tool for activism and organization. This may be especially useful for newcomers or other workers that are not yet critically assessing the working conditions or actively engaged in existing platforms that enable criticism and activism, such as TurkOpticon. As I saw in my research, play could become an oppositional means through which to warn newcomers about requesters that devalue the market, and urge other Turkers to esteem their own self worth.

In *research through evocative play*, the declared play is already top-down in the sense that the researcher announces play within a context, instructing participants within a context that,"This is play, now" without asking permission or consent of the participants. In Turker Tales, interpretations of paternalistic play suggest that some participants also saw the play as declared by, driven by, and in support of specific powerful stakeholders in the context— namely, researcher-requesters seeking to maintain low wages. Although I do not recommend paternalistic play as a means of direct support for crowd workers, I conjecture that as a research tool to better understand power dynamics in a context, play presented as benefiting certain stakeholders in the context could have especial value. For this reason and in order to further explore whether power-aligned declared play might be especially useful for *research through evocative play*, in my subsequent work with YouMercials, I included

components that directly primed subsets of participants to view the declared play as aligned with different stakeholders having different power statuses in the context.

5.5.1 Limitations and Broader Implications

I caution against overgeneralizing the findings from Turker Tales. My findings are limited by the self-selecting sample of participants in the study activities, by the choice of play implemented, and by my own perspectives and areas of interest, as well as those of my collaborators, in how I interpreted emerging dynamics. Moreover, although the goal of this work was to explore play in crowd work decoupled from efficiency and productivity measures, I acknowledge that some of the findings of this research, such as crowd workers' willingness to engage in unpaid play during work if it makes that work more enjoyable, could be co-opted by requesters seeking to minimize payments. In addition, although I discuss the *potential* of play to increase social connections, the present study does not directly explore or evaluate how play might affect workers' feelings of belonging and connection with other crowd workers. Lastly, I acknowledge that I present this work primarily from *within* the crowd work context. In so doing, my work considers the redesign of certain aspects of crowd work without seeking to upend the system as a whole, which necessarily limits the scope and impact of the research.

5.6 Next Steps

With Turker Tales, I presented a Google Chrome extension for MTurk that allows Turkers to view, create, and share short stories with one another. I demonstrated that the system allows Turkers to share aspects of themselves with one another, critique unfair working conditions on MTurk, and playfully engage in cultural and political communication. Turker Tales shows that play through storytelling is one way in which we can support crowd workers' user experiences in ways not directly tied to work quality or efficiency, and opens up discussions about roles of requesters in providing more holistic support to crowd workers. This work also demonstrates how declared play can be used as a research tool to gain novel insights into power dynamics in a particular context, giving birth to the *research through evocative play* approach that I will continue to delineate and flesh out in subsequent chapter. In its application of *research through evocative play* to the crowd work context, Turker Tales also broadens discussions of ethics and responsibility in crowd work.

5.6.1 Future Work for Play At Crowd Work

From one perspective, I hope that this work serves as a starting point for exploring more play-based and holistic supports for crowd workers, and provides a starting template for designing for and deploying systems of play using the *research through evocative play* framework in order to better understand and reveal aspects of variegated contexts in the field of human-computer interaction.

Within crowd work, future work should consider a wider range of tangential play interventions, and study the ways in which Turkers engage in and potentially subvert that play. Examples might include non-narrative play such as puzzle games, activities involving creative expressions such as drawing or music production, and strategy-focused play. I especially encourage further exploration of using play in crowd work to encourage and evaluate feelings of belonging and connection among crowd workers. Moreover, I want to highlight that play is just one way in which we can consider how to better support crowd workers holistically. For example, more work is needed to understand the mental health struggles crowd workers may be experiencing. A range of interventions, playful and non-playful, may help to relieve the stress and psychological struggles that crowd workers may face. Additionally, my work with Turker Tales is but a small step in a direction that calls into question the status quo of Amazon Mechanical Turk and similar crowd work platforms. We need more provocative work that challenges aspects of crowd work systems and also designs for alternate futures of work that do *not* presuppose the dominance of such systems.

5.6.2 Advancing *Research through Evocative Play*

After my study of Turker Tales, in which I uncovered interesting empirical findings using a *research through evocative play*-based approach, I felt empowered to continue more formally down the path of *research through evocative play*, using declared play as a research tool to investigate power dynamics among different participant stakeholders in a given context. In the following chapter, I present my work with YouMercials. In YouMercials, I drew upon and synthesized my experiences and findings from both RoastMe and Turker Tales to create, deploy and study a declared play study design to better understand power dynamics in the YouTube advertising context.

YouMercials: Exploring Relationships to YouTube Video Advertising through Play Directives

<div style="text-align: right">

6

</div>

I spoke of the subreddit RoastMe as a place of irreverence and unbridled speech that (precariously and ambivalently) casts practices that are not always assumed to be playful in other offline and online social media contexts— harsh language and online photographic self-presentation— as play through its "comedy, not hate" directive. With Turker Tales, I sought to understand how integrating play instructions directly into Amazon Mechanical Turk, a platform where crowd workers labor for a large corporation, and are often maltreated and marginalized in the process, could reveal aspects of participants' relationship to and power dynamics within the platform. With Turker Tales, building off emergent behaviors from RoastMe, I specifically focused on a form of play directive that encouraged perspective-taking and identity experimentation within the platform.

Emergent behaviors in Turker Tales included satirical humor about and criticism of Amazon Mechancial Turk working conditions, and highlighted perceived imbalances of power both within the platform and within the research design itself, with participants often behaving in accordance to what side of the capitalist agenda they perceived me, the researcher, to represent. With the final work I'm presenting, YouMercials, I research the context of YouTube advertising, notable in that it straddles play and exploitation; as Scholtz would say, it operates as both "playground" and "factory" [192]. YouTube supports aspects of participatory culture and play [148], with parallels to RoastMe and Reddit, in that users create, upload, and rank content on the platform, and users often use YouTube for entertainment and recreation. At the same time, YouTube also contains capitalist and exploitative tendencies, with parallels to Amazon Mechanical Turk. Specifically, user-viewers generate revenue for the YouTube platform each time they are targeted with advertisements based off their viewing history and other online usage data.

Using both a controlled experiment and an in-the-wild deployment study for an extension called YouMercials, I overlay elements of play onto the context of YouTube video in order to elicit reflection and reactions. YouMercials users modify or manip-

ulate YouTube advertisements by either recording new audio-dubs for the orginal advertisement, or creating "imagine yourself as" identity imagination exercises for viewers to see in order to change the experiences of advertisements. With YouMercials, I draw from and directly manipulate the use of ridicule or "roasting style" humor as defined in RoastMe, and incorporate identity imagination exercises similar to those used in Turker Tales to study how specific forms of play directives might function differently in a new context. I study how the play directives might encourage either subversion of or acceptance of YouMercials' declared rules of play, and how they might affect participants' engagement with YouMercials (especially with respect to uses of criticism and humor, and displays of product or brand support). In addition, I observe and analyze participants' attitudes towards both (a) YouMercials as a play concept and (b) YouTube advertising (including brand recall and impressions).

My research on Turker Tales also suggested that paternalistic, top-down play designs that align the play declarations with power-holding stakeholders in a context might be especially useful in evoking criticism of a context. To further explore this notion, in YouMercials, I also directly prime participants (using randomized assignment) to consider the perspectives of peer users and advertisers, respectively. Through my development, deployment, and analysis of YouMercials, I gain more experience in conducting *research through evocative play*, and arrive at a more solid foundation for the research approach that I will present in this chapter and continue expounding upon in Chapter 7.

6.1 Introduction

With (as of June 2019) 1.4 million users (RoastMe) and roughly 500,000 users (Amazon Mechanical Turk), both RoastMe and Amazon Mechanical Turk are far from small. However, in comparison to YouTube, the two communities are quite niche. Founded by three former PayPal employees in 2005 as a way for users to upload, share, and view video content [98], YouTube was purchased by Google in 2006, relies on an advertising-based revenue model, and has quickly ballooned into an Internet behemoth, with (as of 2018) over one billion users worldwide, and thirty million visitors a day [201]. A full 73% of all US adults use the YouTube platform (with usage as high as 91% for 18-25 year-olds, and 85% for 30-49 year olds [220]); although relatively less widely studied, especially within the field of Human-Computer Interaction, YouTube usage far exceeds that of Facebook (68% of US adults), Instagram (35%), and Twitter (24%) [143].

6.1.1 YouTube Advertising as Complex and Power-Ambiguous

For YouTube user-viewers,[1], watching advertisements on YouTube allows them to participate in the platform "for free," but means they are forced to leak their personal information to advertisers. Similar to workers on Amazon Mechanical Turk, YouTube ad viewers are subject to the dynamics of an algorithm-driven platform, including targeted advertising, which often evokes negative consumer reactions [194]. They are thereby likely to have a complicated relationship with ads. For example, they are making the choice to "pay" for YouTube through ads rather than through a monetary-based subscription, and may accordingly prefer watching advertisements to directly paying for services.

In addition, user-viewers may feel loyalty towards the small subset of users who are content creators. These user-creators also may benefit financially from targeted advertising, such that user-viewers that want to support the user-creators they subscribe to may have conflicted feeling about advertising (even though ultimately, the YouTube platform, not content creators, receives the benefits, as most content creators receive little to no payment [197, 146]). Where viewers of YouTube videos have at least some way to communicate with one another about the content through comments, video advertisements prevent such communication directly, as viewers of the same main video content may see different advertisements. Instead, viewers must passively consume the targeted advertisements, and may as a result experience feelings of frustration or entrapment.

6.1.2 YouMercials: Recasting YouTube Advertising as a Space for Play

To study the dynamics of YouTube users' relationship to advertising on YouTube, and to understand how varying the design of play directives can affect reflection and behaviors elicited in a specific platform, I created a concept and functioning prototype for YouMercials. YouMercials is an application (presented as a Google Chrome extension) that asks users to "play" with YouTube ads by directing manipulating those ads, either by recording new audio to replace the original audio, or by creating "imagine yourself as" identity imagination exercises to accompany the original advertisements.

---

[1]Referring to user-viewers that do not pay directly for YouTube Premium or use unsanctioned ad blockers

As part of my experimental study design, I assigned participants to different conditions in which they received different priming directions for the use of YouMercials (using a 3 X 2 + 1 study design, discussed in more detail in the Methods section). Participants were instructed to think of YouMercials in terms of how it could either (a) be used to help advertisers or (b) take back power from advertisers, and some participants were also explicitly told to use a roasting style of humor in creating YouMercials.

Like Reddit participants, YouTube participants may already be more primed and willing to engage in forms of play given the expectations and dynamics of the platform. However, by zeroing in on YouTube advertising, a space of ambiguity and possible disempowerment for users [168, 176], I posit that such an exercise does not clearly shift the power towards either viewers or advertisers. Instead, through users' engagement with the extension, YouMercials could reveal users' relationship with YouTube advertising, encourage critical reflection on YouTube advertising, and illuminate how specific play cues might vary in a new context. From participants' engagement in YouMercials, I seek to address the following four research questions:

RQ1: To what extent do participants choose to either subvert or accept rules of declared play in YouMercials?   Here, I focus on how priming participants to consider YouMercials as a way to (a) help advertisers versus (b) take back the power from advertisers might affect their acceptance or subversion of the declared ruled of play in YouMercials. As I saw in Turker Tales, when participants viewed the declared play as more paternalistic, driven by a requester seeking to maintain unethically low wages while still engaging Turkers, participants tended to engage in more criticism and resistance to the MTurk platform and the imbalanced power dynamics it encompasses. In YouMercials, I aim to understand whether explicitly aligning the declared play with a power-holding stakeholder— advertisers— might inspire resistance to the declared play instructions. In my analysis, I will focus on self-report responses in the controlled experiment and the post-survey reflection.

RQ2: How do participants engage in YouMercials, and how might that vary depending on how they were primed to engage?   I use a mixed methods approach with an emphasis on qualitative coding to assess the types of YouMercials users submit, and whether those YouMercials align with or diverge from how participants were instructed to engage in YouMercials (namely, with regards to whether they were primed to think of YouMercials as taking power away from versus helping advertisers, and using roasting humor or not). Here, following up on RQ1, I am curious to understand, for example, if participants assigned to *roast* choose to accordingly use humor and roasting, and if say, (following Turker Tales findings), those who are

told that YouMercials can help advertisers similarly choose to use criticism or satire. Can we expect that those told to help advertisers will accordingly be more likely to create advertisements that support brands or products than those who are told that YouMercials can instead take back power away from advertisers? To address this question, I observe participants' use of humor, criticism (of the brand/product and/or other people, concepts, situations or entities), relevance to the original product/brand or advertisement, support for the product or brand, and meta-references to advertising practices or strategies.

RQ3: What are participants' attitudes towards YouMercials created by others, and YouMercials as a play concept? I want to understand how participants experience the viewing of YouMercials created by others, as well as how this might vary depending on how participants were first primed to consider YouMercials. In order to understand participants' attitudes, I use enjoyment of YouMercials as measured through Likert scales in both the controlled experiment and the in-the-wild deployment, as well as ratings submitted of others' YouMercials (again, for both the controlled experiment and the in-the-wild deployment). In addition, I use participants post-survey reflections to understand their overall experiences with YouMercials and any self-reported effects YouMercials had on participants.

RQ4: How might using YouMercials influence participants' opinions of and engagement with YouTube advertisements? With RQ4, I consider the ethical implications of declaring play in a capitalist context, regardless of the intents or motivations behind that play. For example, even if some participants using YouMercials engage in play and feel empowered by the play, or use the play to express criticism of the capitalist context, could the play declared by YouMercials bring users attention, time and efforts to advertising, thereby further ingraining product messages in their minds and ultimately further supporting the capitalist system? I discuss the specific metrics used (e.g., empowerment, and brand recall impressions) in the results.

6.2 Methods

6.2.1 Initial Exploration Phase

As with any study, there should always be a period of initial exploration as part of the *research through evocative play* study design. In *research through evocative play*, this initial exploration should also involve self-reflection and defining one's own interests and biases in the area of study. In the case of YouMercials, the initial exploration stage took place over the course of nearly two years, and was aided by the input

and ideas of other students and colleagues. My initial exploration stage began with a self-reflective, introspective exercise, then expanded to include outward-facing design and testing, and lastly included an informal pilot test with the YouMercials concept and design.

Design Ethos & Action: Personal Reflection on and Exploration of Reactions to Online Advertisements

In the Spring of 2017, I had the opportunity to take the class Design Ethos & Action, a course in Carnegie Mellon's School of Design co-taught by the Human-Computer Interaction Institute's alumnus Peter Scupelli of the School of Design, and Kate Hamilton of the Eberly Center of Teaching Excellence.[2] The course had a dual focus on design and sustainability topics as well as feminism and gender theory, and asked students to consider how to link their personal values with the design projects and practices they pursue, and assess the impacts (both positive and negative) and embedded values of their work, such as those related to sustainability, gender inclusivity, or race relations.

One of the projects I undertook as part of the course was a daily measures project, in which I explored my daily online behavior for a period of two weeks. The timing of the project coincided with a period of personally undesirable online behavior on my part. Although I consider myself a feminist, I had been finding myself inhabiting online spaces and consuming messages on the Internet that conflict with that identity. What's more, I had been finding myself in such places without having a clear understanding of how I had gotten there; there was a sense of being trapped, and associated guilt at being so docile and gullible as to allow myself to enter the traps.

Thus, for the daily measures project, I focused on cases where I passively consumed what I term "un-feminist" media and messages online; media that actively conflict with gender and female empowerment and encourage restrictive gender expectations. Using an introspective method of tracking and analyzing my own online use over the course of two weeks, I discovered commonalities (1) in the paths through which I arrived at un-feminist online spaces, with starting points often in directly feminist or neutral online spaces; (2) in the body and appearance-centered nature of the un-feminist spaces; and (3) in the larger, systems-level influences on my own behavior, particularly in relation to online advertising and Internet revenue models. I used these findings as introspective prompts to consider how my own value system

---

[2]See: https://designethosaction2017.wordpress.com/about/ For any CMU students reading this, I highly recommend taking this course if it is still offered!

interacts with technology, and what personal behavior changes might affect my experiences with large-scale technology systems.

To observe my own experiences (without completely biasing the paths I took in the moment), I waded through my browser search history after the fact in order to view the paths. In analyzing my online experiences, I found that in almost every case, my paths began in either specifically feminist online spaces— spaces where I was actively seeking out discussions of feminism or messages of female or gender empowerment— or neutral spaces, such as social media platforms or online shopping experiences where I was seeking either unisex or male-oriented products.

The influence of advertisements and other alogirthmic recommendations. In almost every case, the turning point in the path—- that is, when I shifted from a feminist or neutral online space to an un-feminist space— was the result of an advertisement or "clickbait" (content whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page or video), both of which are algorithmically driven by the given platform (e.g., YouTube or Faceback). See Table 6.1 for examples of paths I observed myself taking.

Starting point	Turning point	End point
Searching for men's sweatpants on Amazon	One of the results was, perplexingly, for a corset (arguably, an ad)	After clicking on the corset to understand, ended up clicking on what is essentially a "butt corset" for "booty lifting"
Looking up biographies of Michele Obama on goodreads.com	Ad/clickbait	Article on all the times celebrity Blac Chyna dressed too sexy for her own good
Reading articles (originating from Facebook) on the women's march, intersectional feminism, and transgender experiences	Ad/clickbait	Watching YouTube videos on how to minimize negative facial appearance through makeup contouring, and women's experiences with plastic surgery such as rhinoplasty
Went to YouTube to find a fitness video to do	YouTube suggestion (similar to an ad)	Today Show video about celebrity socialite Khloe Kardashian's new show, "Revenge Body"

Tab. 6.1: Concrete examples of paths I took to "un-feminist" online consumption, explored in the context of a class project in Professor Peter Scupelli's Design Ethos & Action course.

Considering the implications of active engagement with algorithmic content that is normally passively consumed. I should note that in approaching this daily measures project, I was very much influenced by the work of designer Nathan Shedroff

[198] and feminist scholar bell hooks [100]. Drawing from Shedroff's systems-driven perspective of design, I considered what systems-levels forces such as economic systems and social systems influenced my own decisions and the shape of the Internet as a distributed institution. From bell hooks, I drew from the concept of the "oppositional gaze." In her 1992 essay "The Oppositional Gaze: Black Female Spectators," hooks starts with the premise that not only are black women under-represented, mis-represented, and negatively represented in film, but they are also discouraged from "looking," that is, critically reflecting on their absence and misrepresentation in film. hooks argues that by applying an "oppositional gaze" to mainstream film, or viewing mainstream film through a lens of criticism that analyzes and resists the social construction of ideals portrayed in media, black female spectators can experience "pleasure in resistance." In so doing, they effectively take back some of the power from white and/or male filmmakers, and fight gender and racial inequality.

I found it interesting but also perplexing that although I would often start off in feminist or neutral online spaces, marketing forces (ads/clickbaits) directed me towards what I considered highly un-feminist online spaces. If my other online behaviors suggested an interest in feminist messages, then why was I being sold un-feminist messages? Is it simply because Internet markers had identified me as female that sites were encouraging me to engage in un-feminist media? Is feminism unprofitable?

Much of my personal dilemmas regarding my interactions in these un-feminist online spaces is not simply that I engaged in them, but that I engaged passively; there was very little of the "pleasure in resistance" (in my case, resistance towards "un-feminist" portrayals and messages) that bell hooks described in the oppositional gaze. Further, I found that for many of the clickbaits I engaged in, there was a marked absence of a comment section, which suggested an interesting (and problematic) dynamic. Where many spaces on the Internet are social in nature (e.g., "social" media), I experienced un-feminist spaces in isolation. Because of the lack of comment thread options, I was unable to interact with or hear from other consumers of the content. I considered the implications of the anti-social nature of these spaces, and concluded that cultivating an oppositional gaze may be especially difficult to do in isolation, where criticism and resistance must remain internal, without external support or confirmation. Thus, I began to consider design potentials and implications for engaging actively with others in a shared space rather than consuming passively messages alone (in virtual isolation) that might conflict with my values.

Focusing on YouTube as a context.   Moreover, I found my experiences of "un-feminism" on YouTube to be more emotionally-laden than those on other platforms, such as Facebook or Amazon; I felt particularly passive and helpless during my

YouTube viewing experiences. On YouTube, I found that both advertisements as well as videos suggested by YouTube's recommendation algorithm influenced the turning points.

Ideas from the Design Ethos &Action Project that Impacted YouMercials. Thus, my introspective reflection on my own online behaviors in relation to feminist values and "un-feminist" messages consumed online led me to three prominent ideas that later influenced my design of the YouMercials *research through evocative play* study. First, I viewed myself (as an Internet user) as subject to algorithmic forces, and wanted to further study how individual users react to and reflect on their own individuality and identity in relation to such algorithmic forces. Second, I wanted to better understand the impacts that a lack of ability to communicate about or in response to mis-targeted, passively consumed media can have on individuals. I considered in tandem the potential implications of flipping the table by encouraging active engagement with algorithmically-targeted messages and media. Third, I was particularly interested in the further exploration of YouTube, and YouTube algorithmic forces (especially advertising) as a context (which further reading and research revealed has not been as traversed a research context in human-computer interaction as other social media platforms, such as Facebook and Twitter).

My introspective exploration was highly personal in nature, and revealed my own biased perspectives towards algorithmic recommendation systems, and online advertising and marketing. For example, I found that my own relation to advertising is negative in that I feel mis-targeted by online advertising in ways that conflict with my identity and values. However, I would also characterize my relationship with online advertising as ambivalent in that I do not necessarily want to do away with online advertising altogether. I found my internal conflict with my engagement in "un-feminist" online messages and spaces was not that I clicked on the ads or clickbaits in the first place, but rather, that I did so without resisting (to use bell hooks term). If the issue were only that I clicked, then the solution is slightly more clear-cut. I could, for example, pay for an ad blocker. I could also consider how companies are tracking me, and choose ways to peruse the Internet in less detectable manners (e.g., by using the Tor browser).

However, I realized through the process of reflection that I don't want to hide myself from what I view as distasteful, mis-targeted, offensive media messages, because I think it's important to critically engage with existing cultural media. To simply avoid negative, un-feminist online spaces would be to engage in online media in a highly limited and ultimately unfulfilling manner, and in addition, would conflict with other values I hold, such as the openness of the Internet. I thus concluded it would be unethical of me to impose limits on certain individuals and entities in

terms of information tracking, and not others. For example, as a person who is very curious about other humans, I value that the Internet allows me to explore different facets of human interactions and behaviors in ways that previously would have been very difficult, if not impossible. I also value the relatively low financial barriers to entry the Internet encompasses; although I'm not a huge fan of ads and marketing, (1) I prefer this state to a state where I would pay more monetarily and (2) I am also curious about the ads and marketing themselves. After all, marketing is a human-created economic force, and as such offers interesting insights into humans, human interactions, and human behaviors.

With Turker Tales, my initial exploration phase was very user-focused, and less introspective in nature. I have dabbled in MTurk as both a worker and a requester, but am not a regular or particularly active user of the platform. Thus with Turker Tales, I took stock of my own opinions and biases largely in relation to research I had read. Instead, my initial exploration phase for Turker Tales focused more on *other* users of the system, such as through the speed-dating sessions and formal pilot study with Turkers. It was only *through* the *research through evocative play* study of Turker Tales that I came to more fully realize the implications of my role as a researcher-requester on the platform. In retrospect, including more introspection in the initial exploration stage of Turker Tales could have benefitted me and my research.

Thus, by contrast, with YouMercials, I began with an introspective approach. I considered myself as an active user of the YouTube platform, with the attendant biases informed by a set of personal experiences that I knew would have a direct impact on how I designed the *research through evocative play* study. Despite the user experience/user design adage, "you are not the user," with YouTube, I *am* a user of the system. Only after I had analyzed and reflected on my own experiences and biases related to marketing and YouTube did I move to outward inquiry about others' experiences and perspectives as part of the initial exploration phase.

"RoastMe, Part 2:" Specific Considerations for Encouraging Engagement in the YouTube Advertising Context

In Fall 2018, I co-instructed six undergraduate and master's students in an independent study course alongside my advisor, Geoff Kaufman.[3] In the course and building off my prior work with RoastMe, we used digital probes to explore what behaviors and reactions the RoastMe "harsh humor as play" paradigm might elicit in other

---

[3]Shout out to the awesome and creative students in the course: Lily (Hsin Yu) Pai, Xi Jin, Jerry (Jee-Uk) Fu, Sarah Kim, Karen Kim, and Max (Yutong) Zhu.
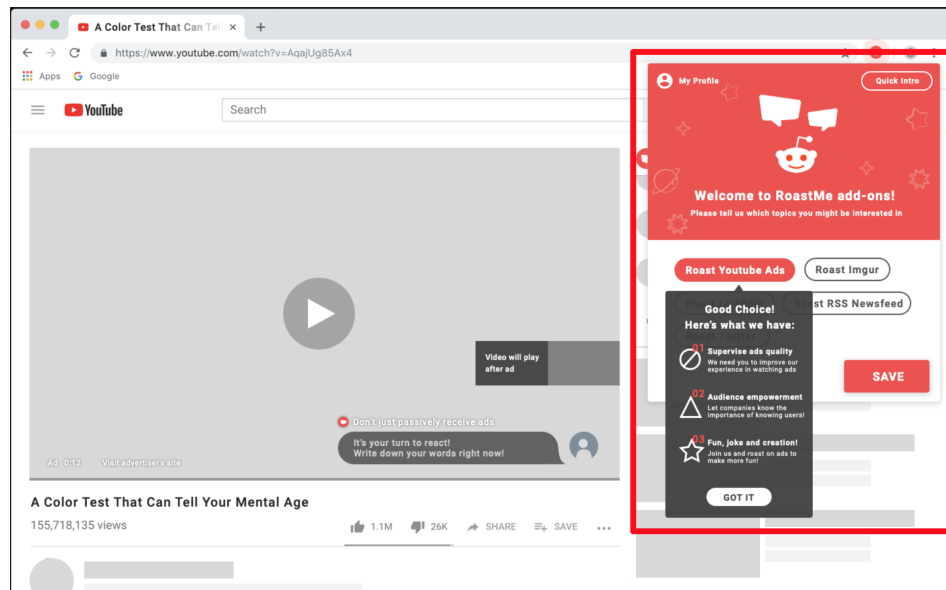
Fig. 6.1: Mock-up of introduction screen to the YouTube RoastMe add-on prototype (created by Lily Pai).

contexts and situations. The course centered around discussion, group brainstorming activities, and rapid iteration and user-testing of lo-fi (paper) and medium-fi (e.g. InVision and Keynote) prototypes. An ongoing, recurring discussion (and often a point of healthy disagreement and conversation among individuals in the group) was how and whether elements of harsh humor could be translated to other contexts in a way that created a safer, less harmful space than that engendered by and observed in RoastMe. For example, we considered that any form of harsh humor or ridicule aimed at an individual holds potential for injury, even if the individual being ridiculed appears to consent to the process, or if there are also beneficial components to the engagement. We discussed alternative designs that might mitigate harm while still allowing for expressions of frustration, such as roasting animals or brands instead of directly roasting people. This led to Human-Computer Interaction Institute (HCII) master's student Lily Pai designing and piloting a mock-up of a browser add-on in which users could "roast" YouTube advertisements or brands while watching YouTube ads (with comments uploaded to a subreddit). See Figures 6.1 and 6.2 for examples of the mock-ups that Lily created.

The feedback I gave and the influence I exerted on the prototypes were heavily biased by my own experiences, including the introspective daily measures project I had completed related to online marketing and algorithm targeting. For example, we discussed as a group how a YouTube ad-roasting prototype could potentially benefit either (a) advertisers, by providing insights into ways to make their advertisements more relevant and compelling, or conversely, (b) users, by giving them a potential means of creative empowerment and agency in that *they*, not advertisers, would have a chance to create the content that other users would view. In these discussions,
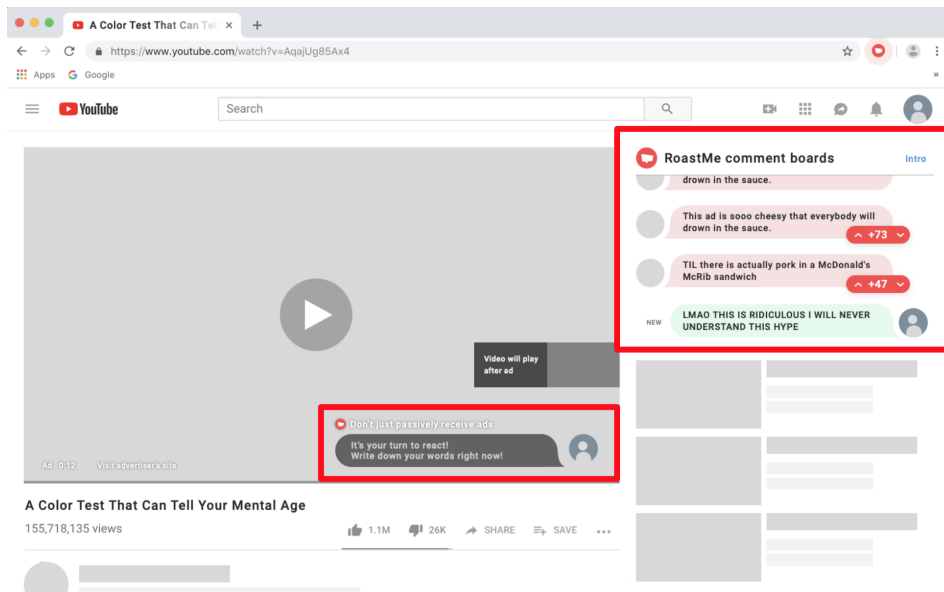
Fig. 6.2: Mock-up of roasting interactions using the RoastMe add-on prototype (created by Lily Pai).

I found myself fully biased towards the empowerment side, and resistant to designs with dual purposes to help both viewers and advertisers, or designs that were intended to help advertisers, primarily. Thus, the group discussions around the prototyping process within the context of the class helped further reveal my own biases and highlight alternative interpretations of play and engagement in the YouTube advertising context that I later integrated into the *research through evocative play* design for YouMercials.

In addition, students in the "RoastMe, Part 2" course also recruited participants to user-test and give feedback on the lo- to mid-fi prototypes and mock-ups they designed. One of the recurrent findings across students and prototypes concerned the focus on the UX design and usability of the prototypes. More polished medium-fi prototypes, especially (e.g. InVision-based prototypes[4]), tended to distract participants from roasting concepts as applied in the given context. Students had trouble directing participants away from discussions of the look and feel of the prototypes. Especially given that later, I wanted to explore YouMercials (a) as a *research through evocative play* design, not a play design in itself, and (b) as an HCI researcher, not a UX designer, this signified the importance of keeping the interface itself bare bones and relatively "rough" in nature to encourage users to focus on the concept and functionality of YouMercials rather than its aesthetic design.

---

[4]https://www.invisionapp.com/

Designing and Piloting YouMercials

These introspective and outward-facing explorations related to online advertising, algorithmic targeting, and engaging playfully in the YouTube advertising context, as well as my prior research of RoastMe and Turker Tales, and my nascent concept of *research through evocative play* as a research methodology, all influenced the development and design of YouMercials. Largely informed by the introspective daily measures project I had undertaken, I went into the design of the YouMercials *research through evocative play* study with clear biases; feelings of being targeted and powerless in the face of online advertising skewed me towards more negative perspectives on the power dynamics embedded in YouTube advertisements, and I tended to see advertisements and advertisers as an external power to be "fought" against. I reviewed existing literature on online advertising and YouTube, but read these sources already wearing a biased set of lenses.

From Turker Tales, I wanted to follow up on and further study the emergent notion that "paternalistic," top-down play declarations that seem to favor a power-holding party in the context could potentially encourage more unbridled (perhaps snarky or "roasting-style") criticism of that power. Where harsh humor arose spontaneously and unprompted in Turker Tales, I also wanted to explore how directly prompting harsh humor, as in RoastMe, could function as a declared play mechanism in a new context. Pulling from the "RoastMe, Part 2" course's prototyping work, I sought to create a "good enough" design for YouMercials, rather than a polished product, so as to encourage users to focus on the concepts embedded in the design rather than the design aesthetics.

The concept for YouMercials, as well as the specifics of the controlled experiment's design and flow, the study instructions and flow, and the Chrome extension functionality and interface were informally piloted by a small subset of users in a convenience sample (N=7), and revised and iterated on accordingly. For example, I improved the clarity of the study and extension instructions, and changed elements of the user interface for the extension. With Turker Tales, I ran a more controlled pilot test before doing the field deployment. Similarly, with YouMercials, I ran a controlled survey-based experiment before conducting an in-the-wild field deployment.

6.2.2 Overview of YouMercials Concept

YouMercials is a concept for an application, implemented in this study via (a) a controlled experiment and (b) a follow-up in-the-wild deployment via a Google Chrome extension. In YouMercials, users "play" with YouTube video advertisements by over-

laying their own content over the advertising content. Accordingly, a YouMercial is a video advertisement from YouTube that users modify, and that other YouMercial users can later see. There are two types of YouMercials that users can create:

1. "Imagine yourself as" YouMercials, wherein users create a character or persona for viewers to imagine themselves to be while watching the original advertisement. Users must fill in three prompts: (a) Imagine you are: *(user types here)*; (b) You are seeing this advertisement because: *(user types here)*; (c) Next you will: *(user types here)*.

2. Audio-dubbed YouMercials, where users record over the original advertisement using their built-in computer microphones.

Both "Imagine yourself as" and audio-dubbed YouMercials are later shown to other YouMercials users (details in subsections to follow). Accordingly, in addition to creating YouMercials, users can also view and rate others' YouMercials (both "Imagine Yourself As" and audio-dubbed) on a scale from 1-5, and (optionally) provide a comment on the YouMercials.

6.2.3 Participant Recruitment

Eligible participants in the YouMercials full study (controlled experiment + in-the-wild + post-survey) were individuals 18+ who were regular users of YouTube (i.e., those who reported using the platform at least a few times a week), and who were also able and willing to download and use a Google Chrome extension for a week. I recruited for the study through a combination of posting on Facebook groups and pages associated with popular YouTube channels, snowball sampling, email recruitment of university students through professors and colleagues at my local university, making general posts on social media sites including Twitter and Facebook, posting to Craigslist, posting on the research recruitment site Call for Participants, and contacting past study participants that had expressed interest in learning of new study opportunities (excluding RoastMe and Turker Tales participants, as I felt these two studies shared too many similarities with the current study). I compensated participants in the full study (which includes the controlled experiment + in-the-wild study + post-survey) with a $20 Amazon gift card.

Because the controlled experiment required larger numbers to achieve adequate statistical power for quantitative analyses, I boosted participation in the controlled experiment by allowing for controlled-experiment-only participation; for the controlled experiment-only portion, participants had to be 18+ and users of YouTube, but did not need to be Google Chrome users, nor did they need to download the

YouMercials extension. For the controlled-experiment-only portion, I recruited specifically from Amazon Mechanical Turk (approval rate of 97%+ and completed HITs over 10,000 to ensure high quality), and also allowed participants who signed up for the original study to discontinue participation after completing the controlled experiment, if desired. Participants in the controlled experiment were compensated $7.

After accounting for data tidying, the final sample counts were N = 156 for the controlled experiment, N = 57 of whom completed the full study (including the Google Chrome extension in-the-wild deployment and post-survey).

6.2.4 Controlled Experiment: 3 X 2 + 1 Experimental Design

The primary controlled experiment portion of the study was implemented as a survey via Qualtrics. Using a 3 (stated purpose of YouMercials: taking back the power from advertisers, helping advertisers, or no purpose given) X 2 (roasting instructions provided: yes, no) + 1 (YouTube advertisement watching-only control) experimental design, all participants were assigned to one of seven different conditions. To do so, I used a built-in Qualtrics randomizer (set to make the number of participants per group roughly equivalent). When explaining the concept of YouMercials, I varied the text shown to participants.

Factor 1: Purpose (3 conditions).   All participants were first shown introductory text about YouMercials: "The following exercises will prepare you for using YouMercials. When you normally watch YouTube, you will view targeted video advertisements. YouMercials lets you play with ads by overlaying your own content over the advertisers' content." Next, depending on their condition, participants saw the following additional text:

- *Condition = Take back power.* "In this way, peers see your creations instead of the original advertisements, and you take back some of the power away from advertisers."

- *Condition = Help advertisers.* "In this way, advertisers can gain insights into how they can make their advertisements more relevant, interesting, effective, and/or enjoyable for viewers."

- *Condition = No purpose.* [No additional text shown; no explicit purpose was stated].

Factor 2: Roasting (2 conditions). After I explained the general concept of audio-dubbed and "imagine yourself as" YouMercials, some users were also given instructions to use a roasting style of humor.

- *Condition = Roast.*"For both types of YouMercials, you should use a 'roasting' or ridiculing style of humor. To roast means to humorously mock or humiliate someone with a well-timed joke, diss or comeback. For example, you could choose to make fun of the advertisement, the advertisers, the actors in the ad, or the viewers of the advertisement."

- *Condition = No roast.* [No additional text shown].

Note that moving forward, I will use the shorthands of (factor 1) *take back power*, *help advertisers*, and *no purpose*, and (factor 2) *roast* and *no roast* when I refer to the 3 X 2 experimental conditions. To ensure that participants had read and understood their assigned study prompts, I required all participants to describe the instructions they received in their own words.

A final group of users was randomly assigned to a "hanging control" condition, in which participants were given no additional information about YouMercials and, as described below, were instead asked to watch a set of YouTube video advertisements (they neither watched nor created any YouMercials as part of the controlled experiment).

For the next portion of the survey, I asked all participants in the six experimental conditions represented by the crossing of audience and roasting instructions to create five "imagine yourself YouMercials" and describe how they would create five audio-dubbed YouMercials; before creating the YouMercials, I included text to remind participants of the conditions to which they had been assigned. For the controlled experiment, I did not ask users to actually record any audio for the audio-dubbed YouMercials, but instead instructed, "Please briefly describe how you would replace the audio for this advertisement. For example, you can type what you might say or record in place of the original audio." All participants created YouMercials for the same ten advertisements. As stated above, participants in the hanging control condition did not create any YouMercials, but instead were asked to watch at least five seconds of each of the ten advertisements, simulating the common five-second skip allowance on YouTube. I chose the ten advertisements based on Internet searches discussing commonly viewed advertisements, and sought to make the advertisements potentially applicable to wide broad audiences. I also strove to make the five advertisements used for creating and viewing "imagine yourself as" and audio-dubbed YouMercials coincide in terms of the products/brands being advertised. The brands featured in the YouTube advertisements I selected were:

Thumbnail	Brand or Product	Content of Original Ad	YouMercial (provided by pilot users)
	Grammarly (online application to automatically check one's grammar)	Stressed student writing a paper achieves success through the use of Grammarly.	Imagine you are: A student who aspires to be an author, leader, change agent in society. You're seeing this ad because: You are under an immense amount of pressure in your academic and personal life. In the midst of many projects, you want to hold onto you integrity, you want to be great. You are considering short-cuts. You want to take the proper path. Next you will: You are inspired to be moral and ethical, like the hero in the commercial story, but you will also consider using Grammarly as a tool which improves your arsenal without bending any rules.
	Google Home (smart home device)	Family shows multiple uses of Google Home, from father waking children through music, to mother setting dinner reminders and confirming package	Imagine you are: A head of household deciding which home assistant to buy You're seeing this ad because: To determine if Google Home is what you want to purchase Next you will: Read 1984 and be appalled, not realizing the irony
	Nissan Kicks (car)	Man dancing in headphones and sneakers constructs a Nissan Kicks out of thin air.	Imagine you are: your significant other You're seeing this ad because: you're homeless Next you will: yell at someone
	Olive garden (Italian-style chain restaurant)	An advertisement for "discovering Italy's lighter side."	Imagine you are: hungry You're seeing this ad because: you're thinking about going out for dinner Next you will: go to olive garden
	Walmart (big box retail store)	A celebration of International Women's Day emphasizing Walmart's support for female empowerment.	Imagine you are: A men's rights' activist who used to frequent r/incels and r/taketheredpill You're seeing this ad because: You absolutely hate International Women's Day and are making a list of all companies who have made advertisements in support Next you will: Break down, crying, repeating: "et tu, Walmart?"

Tab. 6.2: "Imagine yourself as" YouMercials shown to participants as part of the controlled experiment.

- *For "imagine yourself as":* Grammarly (web application to assist with grammar while writing), Google Home (smart home device), Nissan Kicks (car), Olive Garden (chain restaurant), Walmart (big box retailer).

- *For audio-dubbed:* Wix.com (website creation tool), YouTube music (premium music service), Geico (car insurance), Red Lobster (chain restaurant), Target (big box retailer).

6.2.5 Controlled Experiment: Survey Design

After seeing the initial priming text according to their assigned conditions, users in the six experimental conditions were asked to rate (using a Likert scale of 1 - 5 to mimic "stars", with 1 indicating worst and 5 indicating best) and comment on five "Imagine yourself as" YouMercials and five audio-dubbed YouMercials (showing text descriptions rather than actual audio, in the same manner that participants created their audio-dubs), again using the same ten advertisements they had encountered when creating the YouMercials. (Note: participants were not permitted to return to previous pages in the survey so as to prevent the viewed YouMercials from influencing their created YouMercials). I selected these ten YouMercials from among

Thumbnail	Brand or Product	Content of Original Ad	YouMercial (provided by pilot users)
	Wix.com (easy-to-use website creation tool)	An actor shows how he used Wix.com to make an acting portfolio.	Another YouMercial user would replace the audio for the following ad with: "As a member of the creative economy, I'm a ginger."
	YouTube Music (premium music service)	Voiceover asks "what is music" while showing music video clips and playing the song "Come Together."	Another YouMercial user would replace the audio for the following ad with:"I'd probably just sing a really poor version of `come together' and occasionally make fart noises"
	Geico (auto insurance)	Sports broadcaster talks with Geico Gecko mascot about car insurance during an NCAA game.	Another YouMercial user would replace the audio for the following ad with: MAN: Everybody loves you, so much that we can put you in any commercial, any situation and people will watch. GECKO: I know, let's me just say "it's couldn't be easier to save money with Geico" and now I'm just going to watch the game and ignore you to the finish. MAN: Can you at least tell us how we can save money with Geico. GECKO: I just pulled something, I just pulled something. BLEEP BLEEP BLEEP
	Red Lobster (seafood-based chain restaurant)	An advertisement for the endless shrimp special showing and discussing examples of available shrimp dishes.	Another YouMercial user would replace the audio for the following ad with: "This would be a direct conversation I have with my audience -- `breaking the fourth wall', so to say. I would most likely talk about food porn / food videography and how, when analyzed in isolation, individual food shots like shrimp rotating in the air do not seem like they should be appetizing in any way, and yet here we are."
	Target (big box retail store)	An advertisement for same day delivery showing different people using Target products and using a red-and-white color scheme.	Another YouMercial user would replace the audio for the following ad with: "Welcome to target's prison, where we hold and generate ambiguous people of color for our commercials. We have Z bars."

Tab. 6.3: (Hypothetical) audio-dubbed YouMercials shown to participants as part of the controlled experiment.

those submitted by informal pilot testers (N=7); I purposefully included a range of tones and assigned conditions in selecting the YouMercials. See Tables 6.2 and 6.3 for an overview of the ads displayed as well as the YouMercials (provided by pilot participants) that controlled experiment participants were asked to view and rate.

Next, in order to measure brand recall, I asked participants in all seven conditions to list as many of the ten brands or companies to which they had been exposed that they could remember, and to report how favorable their impression of the recalled brand or product was (using a five-point Likert scale, from Strongly Disagree to Strongly Agree with the statement "I have a favorable impression of this brand."). I also asked participants in the six experimental conditions (excluding the control) about who they viewed as the audience when creating YouMercials, what their goals were in creating YouMercials, and how much they enjoyed viewing and watching the two types of YouMercials (using five-point Likert scales).

The subset of participants who completed the controlled experiment portion only then answered a short series of demographics questions, modeled after the US Census standards, and including age, gender (write-in), race and ethnicity, country of residence, native or primary language(s), and level of education. Participants who completed both portions of the study (the controlled experiment and the in-

the-wild deployment of the Google Chrome extension) answered these demographic questions in a final post-reflection survey, so as not to prime participants with their demographic identities and thereby potentially bias the in-the-wild deployment results. The controlled experiment alone took approximately 20-35 minutes to complete. For the in-the-wild deployment, participants were required to spend 30 minutes using the extension, but may have chosen to spend longer; the final post-reflection survey took approximately 10-20 minutes to complete, such that those who participated in the controlled experiment, in-the-wild deployment, and post-survey spent approximately 90 minutes in total. To reiterate, I paid participants either a $20 Amazon gift card for the full study (controlled + in-the-wild + post-survey), or a $7 Amazon gift card for the controlled experiment, only.

6.2.6 In-the-Wild Chrome Extension Deployment: Study Design and Flow

Participants in the controlled experiment simulated the experience of creating YouMercials with the same pre-selected set of ten YouTube video advertisements. As a follow-up, I deployed a beta in-the-wild Google Chrome extension version of YouMercials to observe how participants would use YouMercials in a more naturalistic fashion.

I developed YouMercials as a Google Chrome extension, coded using the Google Chrome extension developer framework, Javascript, HTML and CSS. I used Google's Firebase Firestore and Firebase storage to serve and store data on the backend, and also used frameworks and APIs such as Bulma CSS and the YouTube Player API; source code for YouMercials can be made available upon request. I also maintained a simple companion site to YouMercials using Google sites. I used this site as a "home base" for YouMercials, in that it served as a repository for study instructions for users (including videos demonstrating how to use YouMercials and my contact information, should users encounter any issues). When manipulated via the YouMercials extension, the companion site became a locus for YouMercials creation and a portal through which to view others' YouMercials.

Participants who completed both portions of the study (both the controlled experiment and the in-the-wild deployment) were instructed to download the YouMercials extension at the start of the study. Immediately after downloading YouMercials, these users were instructed to click on the YouMercials icon in their browser bar, which would bring up the Institutional Review Board-approved consent form. Once they had completed the consent form, users were shown another modal leading them to the Qualtrics survey for the controlled experiment, discussed above. Upon completing the Qualtrics survey, these users were given a unique completion code
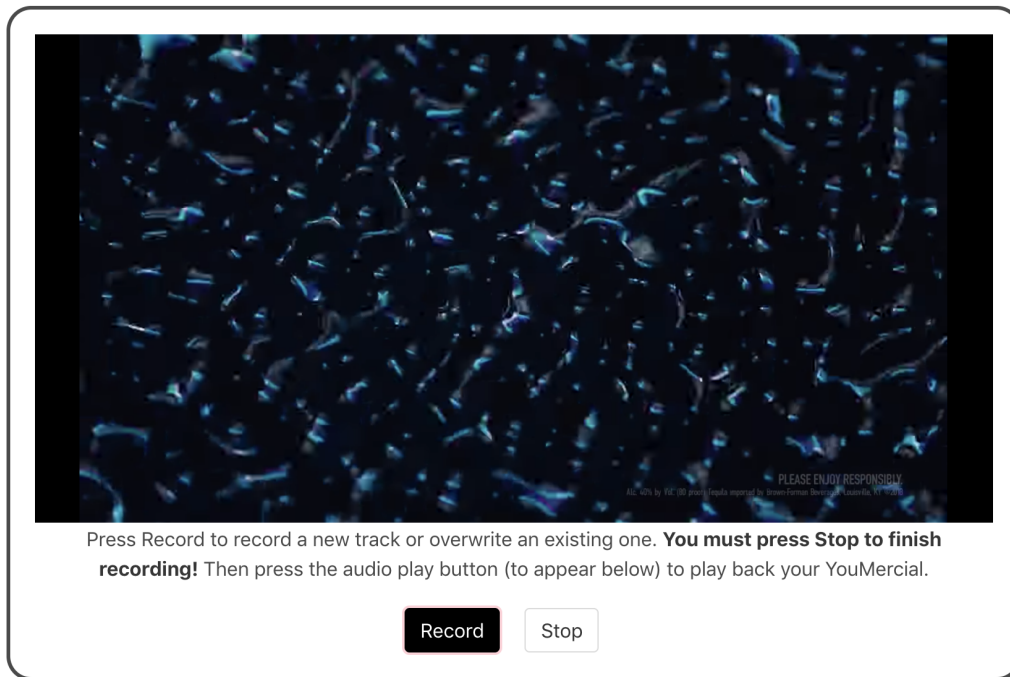
Fig. 6.3: Interface for making one's own audio-dubbed YouMercials. Users click the "Record" button to begin their audio dubs. The videos play automatically as they record their own audio in sync.

that they could copy and paste into the modal page. Only once they had completed first the consent form and then the Qualtrics survey could users partake in the full functionality of YouMercials. If a user had not yet completed the two forms, visits to any branch pages of the YouMercials website other than the main page that featured the study overview, instructions and tips would remain inaccessible, and no other functionality of YouMercials on YouTube was accessible. In order to restrict third party access and enhance privacy where possible, I opted against OAuth, and instead verified users by assigning them a random study id. To develop the Google Chrome extension, I also used content scripts so as to limit the extension's access to user information when using YouTube and the YouMercials companion site.

After a user had unlocked the extension's full functionality, YouMercials would begin collecting (via Firestore) any advertisements shown to the user while they were watching videos on YouTube. Once collected, these advertisements were then made available to users via the extension (accessed from the Google Chrome bookmarks icon or through the YouMercials companion site, directly) to create audio-dubbed and/or "imagine yourself as" YouMercials. In this way, and in contrast to the controlled experiment version of YouMercials, users created YouMercials using advertisements that they had been specifically targeted to them by YouTube's algorithms. Also in contrast to the controlled experiment, for the audio-dubbed YouMercials, users were required to record actual audio via their computer's internal
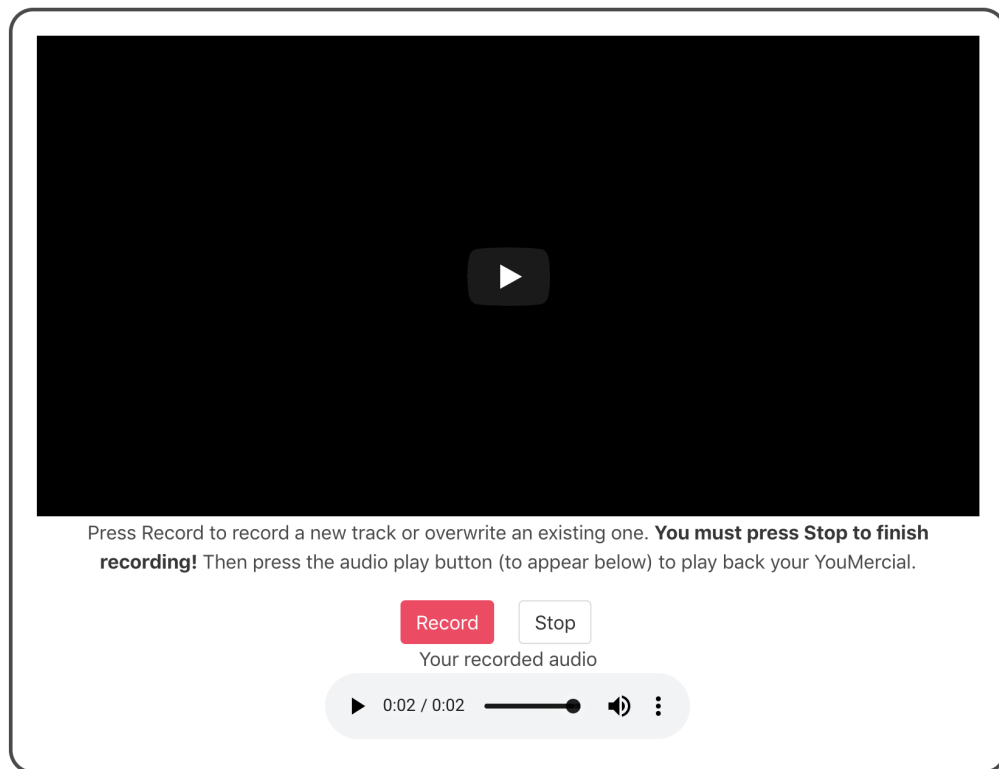
Fig. 6.4: Interface after finishing the creation of an audio YouMercial. Once completed, participants see a slider interface through which they can play back their recordings.

microphone. For both types of YouMercials, users could choose to overwrite their created YouMercials as many times as they pleased.

Users could also view YouMercials created by others, either directly through the website or serendipitously while watching YouTube. The website allowed users to access dynamically generated YouTube playlists of all created "imagine Yourself As" and audio-dubbed YouMercials (excluding the user's own), which users could then view and rate/comment on (using 1-5 stars) via a button that appeared atop the video content. For "imagine Yourself As" YouMercials, users were shown a brief pop-up at the start of the advertisement (pausing the advertisement in the process) presenting them with the identity imagination exercise another YouMercials user created for them to do while watching the video advertisement (Imagine you are... You are seeing this ad because... Next you will...). For audio dubs, the original audio was muted, and instead replaced with the YouMercials audio dub.

Users would only be able to view a YouMercial serendipitously if they happened to be shown an advertisement on YouTube for which another study participant had already created a YouMercial. In my pilot testing, I found this could potentially be quite rare. Although an individual YouTube user may feel they are seeing the same advertisements over and over again, some brands and products create a plethora

Fig. 6.5: Interface for making one's own "imagine Yourself As..." YouMercials. Participants create these YouMercials by filling in responses to the following three prompts: (1) "Imagine you are:"; (2) "You're seeing this ad because:"; and, (3) "Next you will:".

of very similar but ultimately unique ads (with unique associated URLs and IDs). Moreover, given the diversity of YouTube content and YouTube advertising, it was unclear the extent to which users might encounter YouMercials serendipitously. For this reason, study participants were encouraged to make use of the companion site's playlist feature, instead.

6.2.7 Post-Survey

I asked participants in the full study to complete the consent form and initial survey (controlled experiment portion), then use the YouMercials extension for at least thirty minutes over the next several days; I encouraged but did not require that participants try out all four aspects of the extension (creating audio-dub YouMercials, creating "imagine yourself as" YouMercials, rating audio-dub YouMercials, and rating "imagine yourself as" YouMercials), but did verify that they participated in at least one active way in order to be eligible for compensation. I also conducted preliminary data checks to monitor usage, and contacted participants via email to remind them of
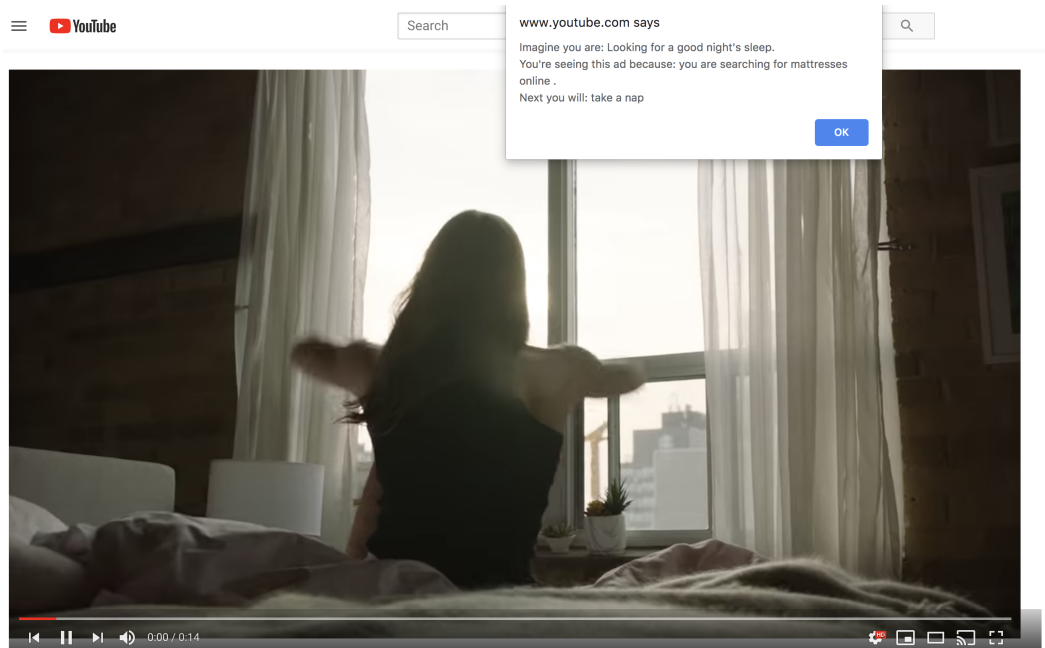
Fig. 6.6: Interface people see when viewing an "imagine Yourself As..." YouMercial on YouTube. Before the commercial plays, viewers encounter a Javascript alert that presents them with the responses to the three prompts that another YouMercials user entered.

study expectations and protocols. I then emailed the final, approximately 15-minute post-reflection survey to those participants that had actively participated in at least one way. The post-survey contained a rehash of some of the controlled experiment measures, including the Likert scales regarding enjoyment, as well a subset of the Empowerment Scale, modified to increase relevance to the YouTube platform, an open response section to capture attitudes about and experiences with YouMercials, and a demographic section.

6.2.8 Methods of Analysis

I analyzed the data collected through the Qualtrics survey using a combination of qualitative analysis methods (performed in Excel), and quantitative analysis methods (performed in R, with the help of installed packages, e.g., dplyr). To qualitatively code the content of YouMercials produced in both the controlled experiment and the in-the-wild deployment of the study, I used an open, iterative coding approach, ultimately honing in on six, non-mutually-exclusive binary categories. Working with a second coder with whom I did not show my own ratings, and after multiple rounds of iteration and discussion of disagreement (sharing new subsets of the data each time so as not to skew our agreement, using previously unseen instances comprising 20% of the original data set size for the final measurement), we arrived at agreement percentages hovering near or exceeding 90% for each of the categories.
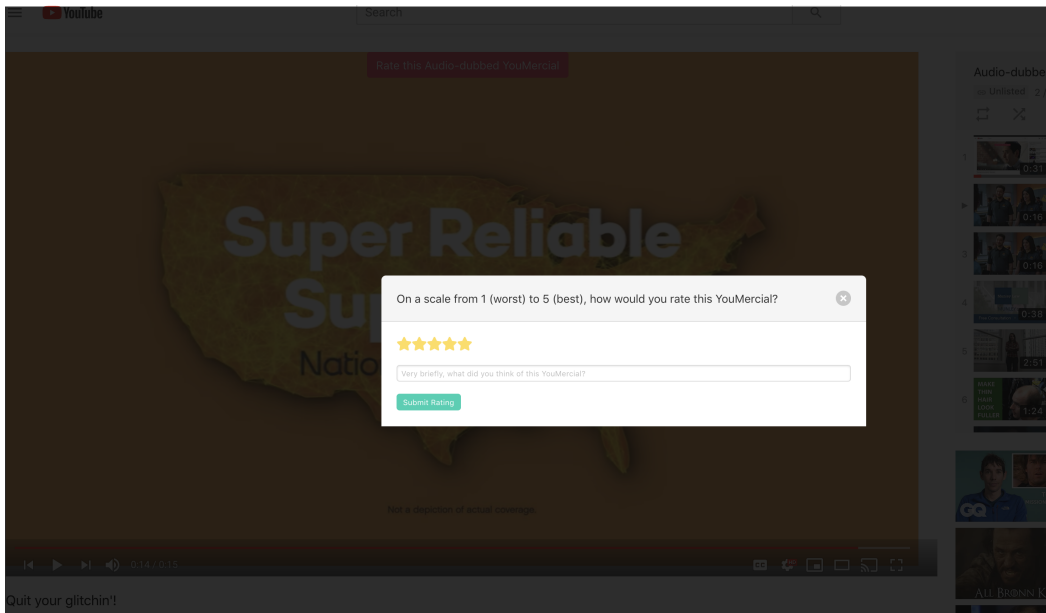
Fig. 6.7: Viewers can rate other users' YouMercials on a scale from 1-5 and can enter comments to explain their scoring.

For the in-the-wild analysis, I transferred data from Firestore to Google Big Query, and used SQL to create appropriate views for analysis, which I later exported as comma separated value (csv) files to analyze using R and Excel. In addition, I manually transcribed all audio-dubbed YouMercials, and qualitatively coded all submitted YouMercials from the in-the-wild deployment (audio-dubbed and "imagine yourself as") using the same schema as the controlled experiment. In order to do so, I also reviewed the original YouTube advertisements to which the YouMercials pertained. I approached qualitative responses in the post-survey with a similar open-coding approach, but without the use of a second coder given the very small size of the data set (57 responses total per question, given that N=57).

6.3 Results

6.3.1 Overview of Participants

156 individuals in total participated in the YouMercials study, with all of these participating in the controlled experiment, and 57 of these participating in the full study, which includes the controlled experiment + the in-the-wild deployment of the Google Chrome extension + the post-survey reflection.

Not all participants chose to complete the demographic questions, with 151/156 overall completing the demographics (all from the controlled experiment portion), and thus I report for only that subset of 151 in what follows. Overall, participants

were fairly evenly split by gender, with 76 males and 74 females (25 males and 31 females in the in-the-wild deployment). On average, participants were aged 31, with a median of 29 and a range of 18-65. Those that chose to participate in the in-the-wild portion were slightly younger on average (range of 18-55; mean of 27; median of 26). Most participants (133/151) resided in the U.S., with remaining participants coming from the UK (8), India (5), as well as Canada, Nepal, New Zealand, and Spain. All but 14 spoke English as a primary or native language. Most participants (90) held a bachelor's or above, along with 24 having some college, 16 a high school degree or equivalent, and 18 an associate's or technical degree. Of the 151 participants, taking the five most commonly appearing race responses in the data set, a majority (102) were White only, and of the other 49 participants, 14 were Chinese, 13 were Black or African American, and 7 were Indian. Of those that continued to the in-the-wild deployment, 30/57 were White only, 10 were Chinese, 7 were Black or African American, and 4 were Indian.

As far as YouTube usage goes, most participants (67) reported spending 30 minutes to an hour each day watching YouTube (based on yesterday's usage, and adjusted using follow-up questions if participants reported that yesterday had not been a typical YouTube viewing day). Additionally, 40 spent longer, at around 1.5-2 hours per day, and an additional 10 spent 2.5-4 hours, whereas 36 reported only 0-15 minutes of use a day. Only 3 participants reported spending more than 4.5 hours daily on YouTube. For the categories they watch (not mutually exclusive), comedy YouTube channels and videos were most popular (106), followed by entertainment (102), music (93), education (64), food (56), science (53), and gaming (51). Given that comedy channels were so popular among participants, I also ran an ANOVA (with the binary comedy category as the independent variable, and the assigned roast condition as the independent variable) to test whether comedy aficionados might be unequally distributed by condition, thereby skewing results, but found that there were no significant differences.

In the controlled experiment, all 156 participants in the study were assigned to 1 of 7 conditions through the Qualtrics survey, according to the 3 X 2 + 1 study design. A breakdown of the number assigned by condition in the controlled experiment as well as for the subset that participated in the in-the-wild portion can be seen in Table 6.4. I note that the numbers aren't *perfectly* balanced due to attrition (e.g. some participants started but did not complete the controlled experiment portion of the study). Nevertheless, the distributions still stayed relatively equally balanced in both the controlled experiment and the in-the-wild deployment. As mentioned earlier, I will use the shorthand of (factor 1) *take back power*, *help advertisers*, and *no purpose*, and (factor 2) *roast* and *no roast* when I refer to the 3 X 2 experimental conditions. Note that aside from comparing the control (+1) condition to any one of the six

Recalling that *Controlled = all (N = 156)* *In-the-Wild = subset (N = 57)*	TAKE BACK POWER (away from advertisers)	HELP ADVERTISERS (help advertisers improve relevance)	NO PURPOSE (no explicit message about purpose)
ROAST (told to roast)	Controlled: 23 In-the-Wild: 9	Controlled: 22 In-the-Wild: 6	Controlled: 18 In-the-Wild: 5
NO ROAST (no explicit instructions about humor & roasting)	Controlled: 20 In-the-Wild: 10	Controlled: 20 In-the-Wild: 8	Controlled: 25 In-the-Wild: 9

Tab. 6.4: Overview of assigned conditions

assigned play conditions in the controlled experiment portion only, I will largely focus on the 3 X 2 study conditions.

6.3.2 RQ1: To what extent do participants choose to either subvert or accept rules of declared play in YouMercials?

In both the controlled experiment (N=156) and the in-the-wild (N = 57) portions of the study, I asked participants questions about who they saw as their audience when creating YouMercials to understand how participants in the *take back power* and *help advertisers* conditions had interpreted the conditions they had been assigned (or whether participants chose different audiences in spite of the conditions they had been assigned) in both portions of the study. For example, following the finding in Turker Tales that perceiving a play intervention as paternalistic might encourage resistance to and subversion of that play, I explored the possibility that the instructions might have led participants in the *help advertisers* condition to focus on peers and empowerment despite their assigned condition.

Expressed Goals in Using YouMercials

First, I found that those assigned to the *help advertisers* condition reported their goals accordingly. As measured by a 2-way ANOVA, they more strongly agreed with the statement that a goal was to help advertisers, and less strongly agreed to having a goal of entertaining peers than their *take back power* counterparts.

Those in the *help advertisers* condition, both in the full study and the in-the-wild deployment, reported acting in accordance with their assigned conditions, with participants in the *help advertisers* condition consistently reporting higher agreement with the statement "One of my goals in creating YouMercials was to help advertisers"

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Roast condition (roast vs. no roast)	1	0.69	0.35	0.262	0.77
Purpose condition (help advertisers vs. take back power vs. no purpose)	2	8.44	4.22	3.18	0.05*
Interaction effect (roast condition X purpose condition)	2	2.61	1.31	0.99	0.38
Residuals	50	66.29	1.33		

Tab. 6.5: ANOVA results with expressed goal of helping advertisers (Likert scale response) as independent variable. '*' indicates significance at the 95% confidence level.

		Factor 2		
		Take back power	Help advertisers	No Purpose
Factor 1	Roast	Mean: 1.85 SE: 0.17	Mean: 3.41 SE: 0.28	Mean: 2.06 SE: 0.21
	No Roast	Mean: 2.09 SE: 0.23	Mean: 3.65 SE: 0.27	Mean: 2.64 SE: 0.25

Tab. 6.6: Mean and standard error breakdown for expressed goal of helping advertisers (Likert scale response) by assigned condition.

(as measured by by a 2-way ANOVA, with the two different condition axes and the interactions of these conditions as independent variables). Results were significant for factor 1 (*help advertisers* vs *take back power* vs *no purpose*) at the 0.05 confidence level (p = 0.000 for controlled experiment; p = 0.050 for in-the-wild). See Tables 6.5 and 6.6 for full ANOVA results and breakdown of means by condition.

I also found that those in the controlled experiment differed by their assigned conditions when it came to whether their goal was to entertain their peers, again measured according to agreement on a 1-5 Likert scale with the statement "One of my goals in creating YouMercials was to entertain my peers." I conducted an ANOVA with the entertainment goal as the dependent variable. I found that the interaction between participants' assignments in factor 1 (*help advertisers* vs *take back power* vs *no purpose*) and factor 2 (*roast* vs *no roast*) impacted the extent to which they saw their YouMercials goal as entertaining their peers (p = 0.009). See Tables 6.7 and 6.8 for ANOVA results and means by condition.

Here, I continued to see adherence to the assigned conditions. Whereas entertainment goals were highest overall for those in the *roast* condition, this only held true if participants were also in either the *take back power* or *no purpose* conditions. Those in the *help advertisers* condition, regardless of whether they were also assigned to *roast* (which we would expect to gear participants more towards entertainment

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Roast condition (roast vs. no roast)	1	2.87	2.87	3.45	0.07
Purpose condition (help advertisers vs. take back power vs. no purpose)	2	8.16	4.08	4.91	0.009*
Interaction effect (roast condition X purpose condition)	2	0.05	0.02	0.03	0.97
Residuals	122	101.39	0.83		

Tab. 6.7: ANOVA results with expressed goal of entertaining peers (Likert scale response) as independent variable. '*' indicates significance at the 95% confidence level.

		Factor 2		
		Take back power	Help advertisers	No Purpose
Factor 1	Roast	Mean: 4.20 SE: 0.21	Mean: 3.73 SE: 0.24	Mean: 4.28 SE: 0.18
	No Roast	Mean: 3.91 SE: 0.15	Mean: 3.35 SE: 0.23	Mean: 3.92 SE: 0.16

Tab. 6.8: Mean and standard error breakdown for expressed goal of entertaining peers (Likert scale response) by assigned condition.

purposes) consistently reported *lower* Likert scale ratings (see Table 6.8). A possible interpretation here is that participants in the *help advertisers* condition took their study assignment quite seriously, and as a result, the goal of helping advertisers to better target and appeal to customers figured as a more prominent goal than more more "frivolous" entertainment of those consumers, even for participants that had also been assigned the roasting humor condition.

Expressed Purpose of YouMercials

I also found participants to largely adhere to their assigned goals in expressing their own intentions while creating YouMercials. However, I found there was higher support for advertisers and ads, particularly among *take back power* participants, than I had expected.

To analyze perceived purposes of YouMercials, I used open responses from the in-the-wild deployment post-survey. Given findings about advertising from prior research and my preliminary insights from the initial exploration phase of the study, I would have expected those in the *take back power*, especially, as well as the *no purpose* conditions to express criticism of the YouTube platform or YouTube advertising in discussing the purpose of YouMercials. For example, participants might have spoken of YouMercials as a tool to fight back against advertisers or corporate power. However, within the *take back power* and *no purpose* conditions, the percentage of in-the-wild users that expressed criticism of YouTube ads in their descriptions of the purpose of YouMercials was quite low, at 5/33.[5]

What's more, some users, even though they were explicitly told that YouMercials could help take back power from advertisers, and even though their responses to the manipulation check question indicated understanding and acceptance of this assignment, expressed desires to support advertisers. For example, f15 (*take back power + roast*) explained the purpose of YouMercials as "To keep people entertained while watching commercials so they wouldn't want to skip them" and f18 (*take back power + no roast*) guessed at underlying, undisclosed study motivations, saying "I thought it was probably a way to target advertising towards a particular demographic." Despite their assigned conditions, f53 (*take back power + roast*) declared the purpose was, "Making the ads more relatable to the target audience." Similarly, some of those in the no audience condition, possibly influenced by the content they viewed from others, assumed the purpose as tilted towards the "help advertisers" side, as in f29's (*no purpose + roast*) response, "To make advertisements more memorable."

---

[5]33 represents the subset of 57 users that were in either the *take back power* or *no purpose* conditions.

Moreover, for those in the *take back power* condition that focused exclusively on peers in their discussion of YouMercials purpose, most emphasized making the advertisements more enjoyable or otherwise helping users to "get through" the advertisements. Examples of "getting through" include f7's "I thought of youmercials as a humor based tool to help people get through bad advertisements [sic]," f8's statement, "Turning the boring time you have to spend watching a commercial into something more amusing" and f20's, "Make ads more tolerable." Very few examples in the *take back power* condition explicitly spoke of a purpose of resistance, critique or empowerment; two exceptions are f45, who described the purpose as "To reduce the effectiveness of actual commercials," and f54, "To take some power back from the advertisers and amuse other YouMercial users."

Perceived Audience Targets in Creating YouMercials

Where results for expressed goals and purposes for using YouMercials indicate patterns of adherence overall, but higher ad support than expected among *take back power* participants, the results for perceived audience targets were less conclusive. In observing who participants viewed as their audience, I saw that in the controlled experiment, almost all participants saw peers, not advertisers, as their audience. I had been curious to understand whether those in the *help advertisers* condition, in particular, might be thinking about directly addressing or creating for advertisers instead of peers, but found that 28/42[6] of those in the *help advertisers* condition specified peers as their imagined audience (this also held true in the post-survey for the in-the-wild participants, with 11/14 in the advertising condition specifying peers as the audience).

Moreover, for those that specified "advertisers" as audience in the controlled experiment using the multiple choice question, in their optional open response descriptions of the audience, many still described peers or hybrid peer-advertiser audiences, e.g. "People around 20-30 years old," or "My immediate thought was people my age/young adults who frequently watch youtube." However, I note that the preponderance of peer-as-audience responses does not necessarily indicate a divergence from assigned conditions. Rather, it appears that in keeping with the overall description and presentation of YouMercials, users in all conditions primarily saw the viewers as peers, including those in the *help advertisers* condition.

Interestingly, in their open descriptions of the purpose of YouMercials in the in-the-wild deployment post-survey (N=57), participants displayed more variation in their uses and perceived audiences. For example, 4/14 of the participants in

---

[6]42 refers to the total number of participants assigned to *help advertisers* in the controlled experiment

the *help advertisers* condition displayed a measure of subversion in describing what they perceived as the purpose of YouMercials, given the priming text to which they had been exposed. For example, f2 (*help advertisers + no roast*) explicitly stated that his personal purpose in using YouMercials differed from his assigned purpose, explaining, "Your purpose was that we engage with the commercial so that, by adding our own creativity, we create a sense of ownership and familiarize ourselves with the brand. My purpose was to show the inherent absurdity of that proposition, as these brands do not have our best interests at heart and never will." Others saw hybrid purposes, possibly influenced by content that they viewed from others. As f11 (*help advertisers + no roast*) noted, "I thought the purpose was to provide critiques of the original commercial (in a way) through humour or what was perceived to be a better delivery of a commercial, which was more for the benefit of other viewers, and indirectly for the knowledge of advertisers." Likewise, f16 (also *help advertisers + no roast*) used YouMercials to voice concern about the product or brand, explaining, "I wanted to point out the main characters of this produce or service. I pointed out what I concerned about this product [sic]" and f37 (*help advertisers + no roast*) felt the purpose of YouMercials was to "rethink the influence of advertising."

Summarizing RQ1: To what extent do participants choose to either subvert or accept rules of declared play in YouMercials? Although there were some exceptional cases of subverting the assigned rules of play, for the most part, participants accepted the rules of play declared by YouMercials in terms of their expressed goals, purposes, and audience targets. For example, those in the *help advertisers* condition *did* report that their goal was to help advertisers, and likewise, those in the *roast* condition rated high on goals of entertaining their peers. When I did see instances of subversion of the assigned condition, it was not always in the direction I anticipated. Specifically, participants in the *take back power* condition sometimes strove to help, rather than fight against advertisers. As we will see with RQ2, this dual trend of (1) general adherence to the declared rules of YouMercials for the *roast* vs *no-roast* factor and (2) higher than expected support for brands and products among those in the *take back power* condition also stayed fairly consistent in terms of how participants engaged in YouMercials (both in the controlled experiment and in-the-wild).

6.3.3 RQ2: How do participants engage in YouMercials, and how might that vary depending on how they were primed to engage?

Creation of YouMercials: Controlled

Qualitative codes. As mentioned in the Methods section, for each of the audio-dubbed and "imagine yourself as" YouMercials created (both in the controlled experiment, and later, in the Chrome extension), I qualitatively coded the content into six non-exclusive binary categories, using an iterative, open-coding technique, in collaboration with a second coder that allowed us to assess interrater reliability (via percentage match, given that the categories were non-exclusive; all exceeded 80% match). Those six categories were:

- *Humor*. YouMercials in this category must include some form of humor, be it dark or light. I was somewhat lenient on what constitutes humor, attempting to capture intent rather than personal and subjective judgments of what is humor. In determining whether humor was present, I leaned on Benign Violation Theory's definition for what constitutes humor [216]. This category could include humor that is outlandish, surreal, silly, sarcastic, or caustic, for example.

- *Critique*. YouMercials designated as critical must include some form of critique of something or someone (often done in a humorous or snarky manner, though the tone needn't be humorous). Descriptions of negative qualities that state the facts but don't necessarily judge or criticize do not count, nor do sarcasm or irreverence that isn't necessarily critical.

- *Roasts*. YouMercials in this category must clearly mock, ridicule, or harshly criticize (humorously or non-humorously) the product or brand being advertised.

- *Related.* This code captures whether or not an "imagine yourself as" scenario or audio-dub is related to or inspired by the content of the ad. YouMercials that are completely off-topic and unrelated to the original ad content are marked as "0."

- *Supportive.* YouMercials marked as 1 for this category must support the advertisement's original message and/or support use of the product or brand.

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Roast condition (roast vs. no roast)	1	97.6	97.65	36.28	1.92E-08***
Purpose condition (help advertisers vs. take back power vs. no purpose)	2	0.1	0.06	0.024	0.976
Interaction effect (roast condition X purpose condition)	2	4.2	2.10	0.872	0.460
Residuals	120	322.9	2.69		

Tab. 6.9: ANOVA results with number of YouMercials created (out of a total of 10) that roasted the product or brand as the dependent variable. '*' indicates significance at the 95% confidence level.

- *Advertisement.* This code refers to direct references to or reflections on the processes of advertising through the content in the YouMercial. For example, a YouMercial might directly address the phenomena of advertisement mistargeting in its content, or directly discuss the act of watching and/or being influenced by advertisements.

Roast participants used more humor, criticism, and brand-roasting. After finalizing the qualitative coding, I integrated these binary codes into quantitative analyses. I found that being in the *roast* condition did indeed predict higher levels of roasting across the YouMercials created, as well higher use of critique and humor. Conversely, as we might expect, being in *roast* negatively predicted support for a product or brand in the YouMercial (see Tables 6.9, 6.10, 6.11). However, when we look at some of the breakdowns in more detail, there are some interesting dynamics that emerge (see Tables 6.12, 6.13, 6.14). For example, although *roast* participants engaged in significantly more roasting behaviors in creating YouMercials in the controlled experiment, they still were not engaging in roasting all that much; on average, *roast* participants only created 2.3/10 YouMercials that roasted the brand. Instead, *roast* participants tended to use criticism in ways that didn't directly lampoon the product or brand being advertised (mean of 5.6/10 YouMercials created that used criticism, as compared to *no roast* of 1.6/10). For example, an "imagine yourself as" YouMercial with the conditions *take back power + roast* for the company Grammarly read: "Imagine you are: A college student with awful spelling and grammar skills. You are seeing this ad because: Your English Lit. teacher has hacked into your PC. Next you will: Become obsessed with letters and start writing them all over your bedroom walls." Such a YouMercial displays a silly, surreal sense of humor that does not show support for Grammarly, but also doesn't directly mock or ridicule

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Roast condition (roast vs. no roast)	1	493.1	493.1	73.76	4.69E-14***
Purpose condition (help advertisers vs. take back power vs. no purpose)	2	17.4	8.7	1.302	0.276
Interaction effect (roast condition X purpose condition)	2	2.4	1.2	0.183	0.833
Residuals	116	775.4	6.7		

Tab. 6.10: ANOVA results with number of YouMercials created (out of a total of 10) that employed critique (regardless of target) as the dependent variable. '*' indicates significance at the 95% confidence level.

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Roast condition (roast vs. no roast)	1	591.9	591.9	79.03	7.0e-15*
Purpose condition (help advertisers vs. take back power vs. no purpose)	2	30.7	15.4	2.05	0.13
Interaction effect (roast condition X purpose condition)	2	0.8	0.4	0.05	0.95
Residuals	119	891.3	7.5		

Tab. 6.11: ANOVA results with number of YouMercials created (out of a total of 10) that employed some element of humor as the dependent variable. '*' indicates significance at the 95% confidence level.

		Factor 2		
		Take back power	**Help advertisers**	**No Purpose**
Factor 1	**Roast**	Mean: 2.05 SE: 0.42	Mean: 2.29 SE: 0.57	Mean: 2.61 SE: 0.45
	No Roast	Mean: 0.74 SE: 0.25	Mean: 0.50 SE: 0.22	Mean: 0.40 SE: 0.14

Tab. 6.12: Mean and standard error breakdown for number of YouMercials created that directly roast or criticize the product or brand (out of 10) by assigned condition.

		Factor 2		
		Take back power	**Help advertisers**	**No Purpose**
Factor 1	**Roast**	Mean: 5.95 SE: 0.60	Mean: 5.3 SE: 0.79	Mean: 5.61 SE: 0.54
	No Roast	Mean: 2.27 SE: 0.62	Mean: 1.21 SE: 0.52	Mean: 1.25 SE: 0.32

Tab. 6.13: Mean and standard error breakdown for number of criticism-containing YouMercials created (out of 10) by assigned condition.

the product, instead critiquing the college student for having awful spelling and grammar.

Support for brands/products: Some unexpected patterns. Although *roast* participants displayed less product support (See Table 6.15), and there were not significant interaction effects between factor 1 (*take back power* vs *help advertisers* vs *no purpose*) and factor 2 (*roast* vs *no roast*), we might expect that those in the *roast + help advertisers* condition would display more support for advertisers than say, *no roast + take back power*. However, as can be seen in Table 6.16, those assigned to *roast* had (on average) lower means for brand/product support than *no roast* participants,

		Factor 2		
		Take back power	**Help advertisers**	**No Purpose**
Factor 1	**Roast**	Mean: 7.33 SE: 0.56	Mean: 6.10 SE: 0.68	Mean: 6.89 SE: 0.69
	No Roast	Mean: 3.00 SE: 0.56	Mean: 1.80 SE: 0.58	Mean: 2.24 SE: 0.53

Tab. 6.14: Mean and standard error breakdown for number of humor-containing YouMercials created (out of 10) by assigned condition.

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Roast condition (roast vs. no roast)	1	429.3	429.3	63.02	1.27e-12*
Purpose condition (help advertisers vs. take back power vs. no purpose)	2	32.4	16.2	2.34	0.10
Interaction effect (roast condition X purpose condition)	2	0.7	0.4	0.05	0.95
Residuals	119	810.7	6.8		

Tab. 6.15: Mean and standard error breakdown for number of humor-containing YouMercials created (out of 10) by assigned condition.

		Factor 2 Take back power	Help advertisers	No Purpose
Factor 1	Roast	Mean: 2.47 SE: 0.53	Mean: 3.90 SE: 0.67	Mean: 3.22 SE: 0.63
	No Roast	Mean: 6.45 SE: 0.49	Mean: 7.55 SE: 0.63	Mean: 6.88 SE: 0.48

Tab. 6.16: Mean and standard error breakdown for number of brand or product-supporting YouMercials created (out of 10) by assigned condition.

regardless of their factor 1 assignments (*take back power* vs *help advertisers* vs *no purpose*).

Moreover, those in the *take back power + no roast* condition actually showed a great deal of direct support for advertisers' products and brands, despite being primed to think of YouMercials as a way to take back the power from advertisers. For example, a description of an audio-dub for a Geico commercial (featuring the car insurance's gecko at a basketball game) appeared to work directly on the side of the advertisers by offering constructive advice for improvement: "I think what the gecko says works. I think I'd change up the background cheers. It doesn't sound like a college basketball game in the background, make it more authentic. The humor is trying to imagine that the little lizard is actually at a March Madness game so make the setting as realistic as possible."

In addition, those in the *take back power + roast* condition, while having the lowest mean support for products and brands overall, still do not come in at 0. They might instead produce sentiments in support of the product or brand. For example, a *take*

back power + roast user created an "imagine yourself as" scenario for a Walmart ad celebrating International Women's Day, saying, "Imagine you are: An adult woman. You are seeing this ad because: You feel a gender bias against you at your job for being female. Next you will: Look at Walmart job openings or google companies with women in positions of power." In this way, even when YouMercials users were primed to take power back from advertisers, it appears they did so in ambivalent conversation with and not necessarily in opposition to advertising forces.

Creation of YouMercials: In-the-Wild

General patterns of engagement.   The 57 participants in the in-the-wild deployment collectively were exposed to 549 video advertisements while watching YouTube during the time period of the study (approximately 3-6 days, depending on when participants joined the study and completed the final survey), with an average of 9.6 ad exposures (SE 1.3) per participant. Nearly all participants (54/57) chose to create at least one "imagine yourself as" YouMercial, with an average of 4.3 per participant, whereas 35/57 created audio-dubbed YouMercials, with an average of 2.1 per participant. On average, 40 of the participants viewed 6 audio-dubbed YouMercials each, and 53 of the participants viewed, on average, 5.5 "imagine yourself as" YouMericals each, though they did not rate the YouMercials during each viewing.[7] Similar to distributions of created YouMercials, users engaged less heavily with the audio-dubbed YouMercials. Thirty-five users rated on average 2.2 audio-dubbed YouMercials each, whereas 47 users rated on average 3.7 YouMercials each. A summary of usage can be seen in Table 6.17.

Demographic effects on engagement.   As noted, YouMercials extension participants were more active in their creation of "imagine yourself as" YouMercials. I ran linear regressions for both the number of "imagine yourself as" and audio-dubbed YouMercials created, including the two factors representing the 3 X 2, as well as key demographic variables (gender, age, and race, using White as a proxy to distinguish between white-identifying and other racial orientations, including mixed race identifications). For the linear regression model of the number of "imagine yourself as" YouMercials created, I found that age negatively predicts the number of such YouMercials, with each year of age increase decreasing the predicted number of "imagine yourself as" YouMercials by 0.17. This suggests that younger participants (regardless of condition, gender, or race) may have engaged more heavily in the creation of "imagine yourself as" YouMercials. For audio-dubbed YouMercials, it appears that the those in the *take back power* condition were *less* likely to produce

---

[7]Note that 40 and 53 refers to the number that viewed at least one audio-dubbed and "imagine yourself as" YouMercials, respectively.

	# of YouTube ads viewed	# of audio-dubbed YMs created	# of 'imagine yourself as' YMs created	# of audio-dubbed YMs rated	# of 'imagine yourself as' YMs rated
Toal across all participants	549	118	244	123	210
Mean per participant	9.63	2.07	4.28	2.16	3.68
Standard error of mean	1.28	0.38	0.51	0.50	0.60
Max per participant	43	8	20	18	21
# of active participants per YouMercial feature (out of 57 participants, total)	55	35	54	35	47

Tab. 6.17: Summary of in-the-wild YouMercials usage. Note that "# of active participants per YouMercials feature' means that to "count," a participant must have contributed at least one instance (e.g., watched at least one YouTube ad, or created/rated at least one audio-dubbed or "imagine yourself as" YouMercials, according to the corresponding column).

more audio-dubs, whereas those in the *no roast* condition were significantly *more* likely to produce more audio-dubs. This may partially explain why, as we will see in following sections, the content and tone of many of the audio-dubs produced through the extension were somewhat more serious and less humorous than I had anticipated. However, as we will see, it does not explain why those in *take back power* or *roast* might produce still non-humorous audio-dubbed YouMercials.

In the post-survey, I asked participants who had not created any audio-dubbed YouMercials to explain why they hadn't, and many expressed self-consciousness and concerns about privacy such as f3 (*no roast + no purpose)*, "English its not my native language and i'm not confortable audio-dubbing adds [sic]" or f14 (*control)*, "I was not comfortable having my voice recorded. Plus, people would be able to recognize me potentially" and f21 (*help advertisers + roast)* "I started and was unsatisfied. I felt self conscious. I'm not a person to make videos or anything normally, so that might affect it." Those in the roasting condition categorized audio-dubs as especially difficult, such as f43 (*take back power + roast*), "It was a lot harder to be funny than the imagine yourself as commercials. "

Overview of quantitative analyses. In analyzing created YouMercials from the in-the-wild deployment, I further filtered the dataset to remove audio-dubs that contained empty, seemingly accidental, or otherwise inaudible or indecipherable audio. I also removed "imagine yourself as" YouMercials with incomplete (seemingly acci-

dental) text, resulting in a total of 96 audio-dubbed YouMercials and 243 "imagine yourself as" YouMericals. These datasets served as the bases for both quantitative and qualitative analyses.

Due to the much smaller nature of the data set, particularly when looking across the 7 different assigned conditions, I was much less likely to find instances of statistical significance. For example, among audio-dubbed YouMercials created in the in-the-wild deployment, 73/96 were created by users in the *no roast* condition, wheres only 15/96 were created by those in *roast* (note that remaining instances belonged to the *control* condition). In what follows, I use quantitative summary measures as a way to guide the qualitative inquiry, and focus on overall patterns rather than differences by condition. I'll remind readers that participants weren't explicitly instructed to adhere to their originally assigned conditions, and that exposure to a variety of others' YouMercials would likely dilute or confound those conditions, regardless.

Lack of humor especially common in-the-wild.   Over half of both the audio-dubbed and 'imagine yourself as' Youmercials submitted in-the-wild lacked humor, with 51/96 audio-dubs and 138/243 of the "imagine yourself as" YouMericals lacking a humorous tone. Of these, I separate out what I term "serious, earnest, supportive YouMercials." These YouMercials did not use humor, general criticism or snark, or roasting (criticism) of the product or brand, were directly related to the advertisement or product, and supported the product or brand. These "serious, earnest, supportive YouMercials" constituted nearly half (42/96) of the created audio-dubbed YouMercials and "imagine yourself as" YouMercials (117/243), respectively. Given that YouMercials is declared as a space for play, and that a subset of users were directly encouraged to either take back the power away from advertisers, and/or to use a roasting style of humor through their creations, the preponderance of these "serious, earnest, supportive YouMercials" was striking. For example, in response to an advertisement for YouTube ads, a user (*take back power + no roast*) wrote the following "imagine yourself as" YouMercial: "Imagine you are: businessman. You are seeing this ad because: I want to put my ads on Youtube. And I would like to know whether Youtube is a better way to broadcast. Next you will: Contact to Youtube about the ads details." Similarly, in response to a cruise ship ad, a *take back power + roast* user wrote "Image you are: bored at home. You are seeing this ad because: you really want to travel on a huge ship. Next you will: immediately book a cruise."

Take back power participants display more support than expected for brands and products.   As seen in the YouTube ad/businessman example above, and in keeping with observed trends thus far, those in the *take back power* condition displayed more support for ads then expected. Sixty-four percent in the *take back power + no roast* and 37% in the *take back power + roast* displayed support for brands and products

in the YouMercials they created (regardless of whether humor was also present). For example, one participant in the *take back power + no roast* condition took the task so seriously that he created an audio-dubbed YouMercial that lasted the entirety of the advertisement (nearly six minutes long) in support of a Chinese dance company performance, Shen Yun. A small excerpt from his transcribed audio reads, "Do you like Chinese culture? Are you a fan of Chinese theatre? Do you like Chinese dance? Are you a fan of some of the most acrobatic, the most electrifying dance performances ever seen?" Despite assignment to the *take back power* condition, when participating in-the-wild using the extension, many participants took the task seriously and did not appear to make attempts to take power away from YouTube advertisements through their YouMercials. Given that we would expect those in the *take back power* condition, if acting in accordance to their assigned conditions, to display virtually zero support for brands and products, such examples thus indicate a deviation from the assigned conditions.

Presence of humor in-the-wild: Not a majority, but still prevalent. Although non-humorous YouMercials outnumbered humorous YouMercials in-the-wild, I want to point out that humorous YouMercials still constituted nearly half of the audio-dubbed and "imagine yourself as" YouMercials, respectively. Overall, humorous YouMercials accounted for 45/96 of the audio-dubbed YouMercials, and 105/243 of the "imagine yourself as" YouMercials. All those YouMercials created by users in the *roast + take back power* and *roast + no purpose* used humor, as did 50% of the *roast + help advertisers* YouMercials.

Participants in-the-wild commonly incorporated critical humor, often with seemingly little regard to their originally assigned condition. For example, 98/243 of the "imagine yourself as" YouMercials and 35/96 of the audio-dubbed YouMercials used some form of criticism, 33 and 14 of which respectively directly roasted the brand or product. For example, in response to a YouTube advertisement urging viewers to support impeaching Donald Trump, one user created the following audio-dubbed YouMercial (*take back power + no roast*), employing a critical, roasting style of humor though they were not explicitly assigned to that condition. "So, here's a random white guy lecturing us about how we need to sign a petition to impeach Donald Trump. But, the alternative to that, is that Mike Pence would be in office, and that's really not any better. But, if you want this guy to stop lecturing you, go ahead and sign that petition." Here, the participant used a roasting style as a means towards taking back power, though she had not been explicitly assigned to *roast*. Likewise, another participant (*roast + help advertisers*, in response to an ad for an Oprah-sponsored diet program, mockingly exclaims, "Hi! It's starving! You'll lose weight eventually..." In roasting, this participant appears to mock rather than support the advertisers and the product.

I should note that I also saw examples of critical humor in cases that aligned with participants' assigned conditions. For example, as one user's (*roast + take back power*) scenario for Grammarly goes, "Imagine you are: Someone who questions the price of 'free' software. You are seeing this ad because: They think they can bamboozle you. Next you will: Realize that everything you type, and where you type it, is stored in a giant database and sold to the highest bidder." Another *take back power + roast* participant created a YouMercial for Call of Duty that directly critiques gun violence, saying, "Imagine you are: Very sleepy. You are seeing this ad because: Everyone plays violent video games. Next you will: Request that guns be banned." In all these examples, the participant use humor to display at least some level of opposition to the ad's message. In these latter examples, the use of humor to display opposition to targeted messages is in line with their assigned condition in YouMercials (e.g., *take back power + roast* participants), but as the earlier examples showed, this was not always the case.

Although less common, the data set also included also examples of non-critical, non-oppositional humor. In some cases, humor might take the form of silly humor only tangentially related to the original advertisement. For example, a participant in *control* created an audio-dubbed YouMercial for Wikibuy, an extension that facilitates finding products through Amazon, by simply intoning in a high-pitched voice "COFFEE MAKER!" (the ad briefly showed a coffee maker product on an Amazon page). Another participant, also *control*, sang Star Wars' Imperial March for a full minute to replace the audio for a hair growth advertisement for balding men. Such examples show a measure of playfulness that is ambiguous about the creator's attitudinal stance in relation to YouTube advertisements.

Summarizing RQ2: How do participants engage in YouMercials, and how might that vary depending on how they were primed to engage? In the controlled experiment, I saw some of the same trends of general adherence to assigned condition among those in the *roast* condition. Those in the *roast* condition did accordingly produce more roasting-style YouMercials, with higher rates of humor and criticism used, and lower levels of product support. However, counter to my expectations, but also in line with other findings regarding higher support for brands and products than expected, brand-supportive and serious and earnest YouMercials were quite common in-the-wild. Especially in-the-wild, where it is also not clear the extent to which participants felt their original instructions remained integral to YouMercials' rules of declared play, those in the *take back power* condition sometimes chose to specifically support the brands and products featured in advertisements.

Next, with RQ3, I observe and analyze whether individuals' level of enjoyment of participating in YouMercials might have varied by assigned condition.

6.3.4 RQ3: What are participants' attitudes towards YouMercials created by others, and YouMercials as a play concept?

Enjoyment of YouMercials

In observing levels of enjoyment of YouMercials, I did not find significant or otherwise notable difference across or between conditions for the controlled experiment. I also note that enjoyment ratings (as measured by self-reported Likert scales) were fairly high overall for the play conditions (participants assigned to any of the six categories in the 3 X 2 experimental design) and, unexpectedly, the control condition (the +1 condition, who were asked to rate their enjoyment of YouTube advertisements). Although play conditions were rated somewhat higher for different forms of enjoyment (e.g. watching audio-dubbed and "imagine yourself as" YouMercials, respectively, and creating the two forms of YouMercials), those in the control condition (who were only asked to view five seconds each of the advertisements) reported relatively high rates of enjoyment, averaging above 3 (where 3 indicates neutrality). Note that to facilitate comparisons between the control ("+1") condition and the 3 X 2 conditions (factors 1 and 2), I looked at each enjoyment scale corresponding to watching or viewing audio-dubbed or "imagine yourself as" YouMercials, and used the "I enjoyed viewing the YouTube advertisements" rating from the control condition for each scale as a proxy data point.

It thus appears that I overestimated the amount of distaste for and annoyance towards viewing advertisements that participants might experience. I find it notable and unexpected that those who engaged with YouMercials actually *didn't* enjoy playing with YouMercials significantly more than those who had simply watched video advertisements, and that watching video advertisements overall was not viewed as a negative activity, given average ratings above 3. However, the setting likely matters. In the controlled experiment, participants were viewing the YouTube advertisements in a separate sphere from their normal YouTube activities, such that the advertisements did not interrupt their flow of video consumption. Participants in the control condition may have thereby experienced the controlled experiment advertisements differently than they would experience YouTube advertisements encountered in their normal, everyday use of YouTube.

Ratings of YouMercials: Controlled

In the controlled experiment, participants were exposed to a mixture of YouMercials (see Tables 6.2 and 6.3 for a description of all the controlled experiment YouMercials presented), from uncritical and supportive of the product/brand (e.g., the Olive Garden YouMercial); to humorous and critical of the product/brand (e.g., the Google Home YouMercial); to humorous and tangential to the product, brand or advertisement (e.g., the Nissan Kicks YouMercial). I was especially interested in how reactions to critical, humorous, and supportive YouMercials might vary by condition, and thus focus my attention on these dimensions.

In analyzing participants' ratings of others' YouMercials in the controlled experiment, I did not find clear patterns in the use of criticism in general or of the product/brand specifically, but I did find patterns in how they rated humor. Those assigned to the *roast* condition, regardless of assigned purpose, rated humorous YouMercials significantly higher than those in the *no roast* condition ($p = 0.005$), with means of 3.2 (*roast*) and 2.8 (*no roast*). Likewise, *roast* participants rated non-humorous YouMercials significantly lower than those in the *non-roast* condition ($p = 0.003$), with means of 2.8 (*roast*) and 3.3 (*no roast*). Moreover, those assigned to *help advertisers* also rated **non**-humorous (serious) YouMercials significantly higher than those assigned to *take back power* or *no purpose* ($p = 0.02$), with mean ratings for non-humorous YouMercials of 3.33 (*help advertisers*), 3.10 (*no purpose*) and 2.80 (*take back power*).

On the one hand, it appears (building off results from RQ1) that participants took the rules of play declared by YouMercials seriously in that *non-roast* participants and *roast* participants viewed the use of humor in others' YouMercials differently, as did those in the *help advertisers* condition as compared to the *take back power* and *no purpose* conditions. However, note that the magnitude of these differences is not very large; participants' ratings in the aforementioned examples still hovered around 3, regardless of condition (indicating neutrality).

Moreover, when I looked at the average of ratings overall, I did not observe significant difference by assigned conditions; across all participants and conditions, the mean average rating was 3.0 (median 3.01; SE 0.05), with 3.0 indicating a neutral reaction (using a five-star rating scale, with 5 indicating best and 1 indicating worst). Thus, participants overall and in specific conditions had relatively neutral experiences of viewing others' YouMercials. Note that these ratings also match up with participants' Likert scale enjoyment responses, which also hovered around the 3.0 mark (on a Likert scale of 1-5). It appeared thus that participants felt indifferent or ambivalent about the declared play of YouMercials.

Ratings of YouMercials: In-the-Wild

In analyzing ratings of YouMercials in the in-the-wild deployment, I first exclude any ratings that rate YouMercials produced by individuals not in the study.[8] I also exclude ratings of audio-dubs that feature blank, indiscernible, or otherwise ostensibly "accidental" audio. For the final dataset I used for the in-the-wild deployment, this leaves me with a total of 100 ratings for audio-dubbed YouMercials, and 201 ratings for "imagine yourself as" YouMercials.

In-the-wild "imagine yourself as" ratings: Ratings of critical YouMercials. Given the much smaller dataset, the variable quality of audio produced in the audio-dubbed YouMercials, and comments that focused more on the nature of the audio produced than the content or tone of that audio, I choose to focus on and exclusively present the ratings for "imagine yourself as" YouMercials. I created a linear regression model with ratings of "imagine yourself as" YouMercials as the dependent variable. For the independent variables, I included both the rater's and the creator's assigned conditions (and the interactions between them), as well as the binary qualitative codes applied to the YouMercials. First, I found that critical YouMercials were more likely to receive higher ratings ($p = 0.04$; coeff $= 1.38$) regardless of the rater's assigned conditions. This could suggest that using YouMercials encouraged criticism, skepticism, and cynicism in relation to watching YouTube advertisements, though it could also be a reflection of a general predilection for dark or critical humor, in general.

Ratings of YouMercials and the *take back power* condition. In addition, there were significant effects related to the *take back power* condition. If either the creator of the YouMercial OR the rater was in the *take back power* condition, then all else being equal, ratings were lower (for creator as *take back power*: $p = 0.013$ and coeff $= -1.23$; for rater as *take back power*: $p = 0.001$ and coeff $= -1.41$). However, if *both* the rater and the creator had been assigned to the *take back power* condition, then ratings were significantly *higher* ($p = 0.019$ and coeff $= 1.43$). This suggests that those instructed to think of YouMercials as a way to take power away from YouTube advertisers identified YouMercials created by "like-minded" participants as higher quality, possibly evidencing a shared approach in accordance with their assigned conditions to YouMercials. However, recalling that a substantial portion of "imagine yourself as" YouMercials submitted by *take back power* participants actually seemed to favor and support the brand/product (64% of *take back power + no roast* and 37% of *take back power + roast*), the results are ultimately inconclusive.

---

[8]Certain individuals downloaded the extension, but provided nonsensical responses on the controlled experiment survey and were thus excluded from the study results, even if they were active in the YouMercials extension.

Post-Survey: Reflections on In-the-Wild Deployment Participation

Reflections on others.   In the post-survey, I included several open-ended questions in order to better understand participants' experiences with YouMercials in-the-wild. I found that when discussing aspects of other study participants that were revealed to them through using YouMercials, about 22/57 participants referred to a sense of community in some sense, often expressed through shared distaste for and annoyance towards YouTube advertisements. As f6 explained, "I realized how funny people are, and that we're all in the same boat— tired of watching the same ads over and over." Participants across conditions expressed sentiments of shared distaste.

However, a subset of participants noted how other users approached the process of creating YouMercials differently than they had. Here, assigned condition becomes more relevant. For example, f27 (*help advertisers + no roast*) said, "Many other users didn't seem to take it seriously and provided nonsense, irrelevant, or joke responses. Of the ones that completed the tasks seriously, the responses were good and sometimes challenged my view of an advert." In this example, we see a participant that took the rules of declared play in YouMercials very seriously, not attempting to resist the notion of helping advertisers. Other participants in different conditions also revealed their commitment to the assigned rules of YouMercials in reflecting on other YouMercials users. For example, f54 (*take back power + no roast*) commented, "...I was surprised that not all YouMercial users used the extension for entertainment and instead seemed to mimic the advertisement message." Though these views reflect quite different perspectives, both participants are responding in accordance with their understandings of the rules of play in YouMercials, indicating a voluntary acceptance of YouMercials as play.

Reflections on the self.   In discussing what YouMercials revealed to them about themselves, 18/57 participants reflected positively on their own sense of humor, creativity, or imagination. For example, f23 (*help advertisers + no roast*) reflected, "I can be more creative than I thought of myself." For a smaller subset, however (8/57), YouMercials highlighted what they perceived as their own flaws or lack of creativity. For example, f14 (*control*) lamented, "I am very not creative and quite literal." Meanwhile, an additional 13 participants reflected on what YouMercials revealed about themselves *in relation to* YouTube advertisements. This could be negative, e.g. f30's (*take back power + no roast*) simple and clear statement, "I hate ads." But for most of these 13, it was more nuanced, and often reflected either positively or neutrally/ambiguously on YouTube advertising. For example, f13 (*help advertisers + roast*) said YouMercials, "Forced me to reconsider how I feel about ads because usually I just skip over them or turn adblockers on," reflecting on YouTube advertising without offering a directly negative or positive judgment. Directly positive statements

could include recognition of positive aspects of advertisements, such as f46's (*control*) statement, "That ads aren't always bad." In an even more positive endorsement, f25 (*help advertisers + roast*) discussed how they paid more attention and directly bought into the messages promoted by YouTube advertisements through using YouMercials, saying, "Yeah. One of the advertisement is about loose weight and gain bonus. As the person who is trying to shape my body recently, I think it's not merely loose weight, but also encourage me to step forward [sic]."

In this way, it appears that YouMercials encouraged some degree of reflection not just on one's own identity, but on one's identity in relation to YouTube advertising. However, this reflection did not patently reflect resistance to assigned conditions or the assumptions embedded by them. If anything, participants appeared to reflect in accordance with their assigned conditions, with those in the *take back power* condition expressing more negative sentiments related to their identity and YouTube advertisements, and those in the *help advertising* condition expressing more positive sentiments with regards to that relationship.

Only one participant discussed how engaging in YouMercials had caused him conflicted thoughts and emotions about his personal identity. Participant f2 (*help advertisers + no roast*) wrote, "That despite my veneer of activism and social consciousness, I am willing to throw away my values for an easy $20, playing right into the hands of the people I swore to oppose." This participant with a moral dilemma indicated that although he had accepted the rules of YouMercials for the study, he felt extremely conflicted about that acceptance. This comment thereby draws to the fore the role of compensation and the researcher-designer in the declared play of YouMercials. Participants were compensated for their participation, and observed by me, the researcher-designer. In this way, I posit that accepting and abiding by the rules of YouMercials may not sufficient "proof" for assessing the presence of a lusory attitude. Participants like f2 might engage in the declared play in an ostensibly lusory fashion, while still feeling highly skeptical of and resistant to that play. Nonetheless, based on the participants' responses to open-ended questions, it appears that relatively few participants experienced internal resistance via moral and ethical dilemmas to the declared play of YouMercials.

Challenges of YouMercials: Ambivalence about engaging in declared play. I did, however, observe a few more examples that highlighted the *ambivalence* of YouMercials play in participants' discussion of the challenges of YouMercials. For example, f22 (*take back power + roast*) wrote, "It is freaky to think about participating more in a capitalist system for fun..." Assigned to conditions that would encourage resistance to and critique of YouTube advertisements through YouMercials play, f22 pointed out that such play is by nature ambivalent. Engaging in such play— even if done in

a manner that critiques the system— does so in a manner that derives enjoyment through and in conversation with that system, thereby further contributing to that capitalist system. For example, participant f48 (*take back power + roast*) said, "I thought this premise of this study was really interesting and now i pay more attention to ads when watching youtube [sic]." Even if f48 views these ads critically, the heightened attention they now pay to YouTube advertisements could again result in this participant contributing *more* to the capitalist system through YouMercials.

I specifically highlight three examples of comments from those in the *take back power + roast* conditions, as we would expect this interaction of conditions to be primed to reflect most critically on YouTube advertising, given that they were told to use critical or sarcastic humor, and also to consider YouMercials as a way to take back power away from advertisers. Playing within contexts in which power, ethics, and morality are ambiguous is by extension, also bound to be ambiguous. In this way, YouMercials first reveals participants' relationships to YouTube and YouTube advertising in the ways they engage with the system and reflect on that system, which as we have seen, are often layered, ambiguous and ambivalent. Secondly, it also highlights the ways in which play is inherently subjective and power-laden, especially in contexts that feature forms of marginalization, power ambiguities, or communication limitations. Any *research through evocative play* study design must call into question and critically reflect on the role of the researcher's power as the (declared) play designer, and the moral implications of designing for play in power-ambiguous or power-imbalanced contexts.

Summarizing RQ3: What are participants' attitudes towards YouMercials created by others, and YouMercials as a play concept?   I found that engaging in the declared play of YouMercials was not any more an ambivalent or neutral experience for participants than watching advertisements on YouTube. However, some of this ambivalence towards YouMercials centered around participants' general acceptance of the rules of declared play. For example, they reported dissatisfaction when they viewed other participants as "breaking" the rules of YouMercials, or not taking their tasks seriously (likely not realizing that other participants had been given different priming instructions when introduced to YouMercials).

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Control (+1) vs. any of the 6 (3X2) play conditions	1	33.2	33.19	9.08	0.003*
Residuals	154	563.3	3.66		

Tab. 6.18: ANOVA results. Dependent variable is the number of brands recalled, out of the total of 10 brands participants were exposed to in the controlled experiment.

6.3.5 RQ4: How might using YouMercials influence participants' opinions of and engagement with YouTube advertisements?

Brand Recall and Impressions

I had contemplated that one likely result of engaging in YouMercials might be better brand recall; because users are forced to engage more with ads through YouMercials than they might have chosen otherwise, the brands may become more salient to them. To measure brand recall, I counted the number of brands or products that participants correctly remembered in the survey (out of a total of ten), and used this as the dependent variable in an ANOVA. Those in any of the play conditions, as I suspected might happen, did indeed have better recall. Measured via ANOVA, the difference between *control* and *play* was significant, with p = 0.003. Average recall did not differ greatly in sheer numbers, however (means of 5.57 and 6.77, and SE of 0.37 and 0.17, respectively); see Table 6.18. I did not find significant differences according to factors 1 (*take back power* vs *help advertisers* vs *no purpose*) and 2 (*roast* vs *no roast*).

Here, I can't necessarily separate play from time spent watching advertisements, but note that likely because participants ended up spending more time engaging with and considering advertisements, participating in the play of YouMercials increased brand recall. Given that advertising manipulations often seek to tap into subconscious associations and thought patterns, engaging in YouMercials, even if critically or with resistance to advertising persuasions, might make users even more susceptible to advertising persuasions.

To measure brand impressions, I had asked participants to respond to a five-point Likert scale, rating their agreement with the statement "I have a favorable impression

Variable	Degrees of freedom	Sum of Squared Errors	Mean Squared Error	F-value	Prob (>F)
Control (+1) vs. any of the 6 (3X2) play conditions	1	2.53	2.53	4.60	0.034*
Residuals	154	84.68	0.55		

Tab. 6.19: ANOVA results with average Likert scale rating of impressions of the 10 brands participants were exposed to in the controlled experiment (with 1 as least favorable, 5 as most favorable) as the dependent variable.

of this product or brand," with 5 indicating most favorable, and 1 indicating least favorable. Similar to brand recall, I found a significant difference (p=0.034) in ratings of brand impressions according to whether participants were in the *control* or one of the six *play* condition combinations (mean of 3.76, SE of 0.10 for *control*; mean of 3.42, SE of 0.07 for *play*), but did not see significant differences along either factors 1 or 2. Those in the control condition rated the products and brands they were exposed to in the controlled experiment more favorably than those in any of the play conditions; see Table 6.19 for full results. Taken together, the results for both brand recall and brand impressions suggest that although participating in the YouMercials exercises strengthened brand recall for participants, doing so also might have led to more critical reflection on those brands, and thus less favorable impressions.

6.3.6 Opinions of YouTube Advertisements

Negative-leaning opinions.   With regards to their opinions of YouTube advertising, participants in the controlled experiment and in the post-survey (after having engaged with YouMercials in-the-wild) expressed largely negative to negative-to-neutral sentiments about YouTube advertisements. Only 16/156 participants expressed positive sentiments related to YouTube video advertisements. Several participants expressed especially negative or hostile attitudes towards YouTube video advertisements, with seven participants directly expressing hatred for YouTube advertisements. For example, s43 (*control*) stated, "Generally, viewing advertisements causing me to reflect mournfully upon the reality of late-stage capitalism and watching these felt no different. I thought about how many of the products were unnecessary (new cars), leading to technological dystopia that is tearing apart the fabric of society (Google, YouTube), or unhealthy (Olive Garden) [participant refers directly to advertisements featured in the controlled experiment]. It made we wish that at least some of the advertisements I encountered regularly were for positive things. I thought about

the ways in which I have tried and will continue to try to limit my exposure to advertisements in everyday life."

YouTube advertising as a "necessary evil."   However, the preponderance of participants saw YouTube advertisements as an annoyance, nuisance, or "necessary evil" to be endured in order to watch YouTube. As f18 (*take back power + no roast*) explained, "I try to ignore them actually but occasionally I get sucked in to watching one. I watch it because it's part of what I have to do to get to what I want to watch..." Similarly, s66 (*take back power + roast*) expressed an ambivalent relationship with YouTube ads in that they wanted to support creators, saying, "I'm not an avid fan of ads on YouTube, but I understand that it's necessary for creators to make money." Others expressed a level of immunity to advertising (something that has also been reflected in recent research, e.g. [94]). For example, s5 (*take back power + roast*) explained, "I don't mind it much. It's nice that I can usually skip them after 5 seconds. I understand that the companies need to make money somehow, and I understand that by being bombarded with advertisements - I am the product. However, I'm used to it, and it doesn't bother me." Twenty-four participants said they normally use ad blockers or a premium service to avoid ads, and forty-one stated that they always ignore or skip the ads as quickly as possible.

Positive opinions towards YouTube ads.   Among those 16 expressing positive sentiments about YouTube advertisements, most of these were juxtaposed with ambivalent, neutral, or negative statements, e.g. s18 (*no purpose + no roast*), "They can sometimes be helpful, but other times they are frustrating because they interrupt what I'm watching." However, a very small minority expressed purely positive sentiments, such as f42 (*control*) "I feel very strongly about the advertisements that are shown on youtube. The reason being, that these types of things expose a person to a new product or service and in a way encompass the passions that the person already holds. They also structure and reinforce certain ways of thinking while introducing a little spice to everyone's life [sic]" and s12 (*no purpose + roast*), "I think they are good because they're short. They also let you find out about certain items or products you wouldn't have known about."

Shifts in opinions of YouTube ads after in-the-wild engagement: Positive.   In the post-survey for those that continued to participate via the in-the-wild deployment, I asked them the identical question about their opinions of YouTube advertisements, and compared their responses (after they had used YouMercials in the wild) with their original responses. While most statements showed no discernible shift in attitude, 10/57 were more positive in the post-survey, and only 2/57 were more negative. For example, f15 (*take back power + roast*) stated in the controlled experiment that they never watch YouTube ads because they use an ad blocker.

In the post-survey, f15's attitude shifted towards a slightly more positive attitude towards advertisements in that they expressed some regret for not watching ads. However, the participant still emphasized their aversion to YouTube ads, saying, "I felt bad that I don't normally watch commercials b/c they must be helping fund my favorite content creators. I didn't realize how much I was missing. But they are still the worst to watch." A participant (f17, *control*) who in the controlled experiment expressed a dislike of the continual interruptions stated in the post-survey, "Video advertisements on Youtube provide a good opportunity to consumers to explore different products and services that we otherwise don't come across." Still others tied their views of YouTube advertisements to the YouMercials extension. For example, f25 (*help advertisers + roast*), who had originally expressed annoyance towards YouTube ads in the controlled experiment, opined in the post-survey that YouTube advertisements could be "good" if participants are using the YouMercials plug-in.

These shifts in opinion towards positive sentiments do not follow clear patterns by condition, but given the much higher proportion of shifts to more positive attitudes views of YouTube advertisements (10/57 versus 2/57), it is worth nothing. For example, it could be that engaging in YouMercials may, by impelling participants to reflect more deeply on advertising, opens them up to ambivalent or even positive sentiments towards advertising rather than critiques of a capitalist structure.

Shifts in opinions of YouTube ads after in-the-wild engagement: Negative. Of the two negative shifts in opinion, one went from negative-neutral to more intensely negative. This participant (f1, *no purpose + roast*) originally simply stated that they used ad blockers, but in the post-survey (after having removed ad blockers for the purposes of the study and having been exposed to more YouTube advertisements as a result) said,"I lack sufficient knowledge of expletives to adequately express my contempt for intrusive advertising." Another participant (f26, *help advertisers + no roast*) had first (in the controlled experiment) spoken positively of YouTube advertisements, saying, "some are applicable and provide information at my finger tips without having to actively search for it" shifted to a negative opinion in the post-survey, saying, "They are sometimes too long and repetitive."

Time Spent Watching Advertisements while Using YouMercials

For the in-the-wild deployment, I asked participants (in the post-survey) whether they felt they had spent more, less, or about the same amount of time watching YouTube video advertisements while using YouMercials in the wild than they would normally, and found that 40/57 reported spending more time engaging with advertisements due to YouMercials; 11 said they spent about the same time, and only six said they

spent less time. Although I did not directly measure product recall or impression in the in-the-wild deployment due to the variegated nature of users' usage and engagement habits, self-reported time spent suggests that when given more freedom and flexibility in terms of how they choose to engage with YouMercials and in relation to the advertisements that were actually targeted towards them, users ended up engaging more heavily with advertisements and brands through YouMercials.

Empowerment Scale

For the post-survey, I also asked participants to answer a subset of the Empowerment Scale [138], focusing only on the questions pertaining to social-political empowerment, and modifying the statements to better fit the YouTube advertising context. Note that I specifically waited until the post-survey (after the in-the-wild deployment usage) of YouMercials rather than measure the empowerment directly in the original survey distributed through the controlled experiment portion. I distributed the empowerment scale as part of the post-survey because I wanted to learn about users' reactions *after* they had played with YouMercials using the advertisements that had actually been targeted towards them, which I posited could elicit different, potentially deeper critical reflection. For each of the ten statements in the scale, participants were asked to rate their agreement on a five-point Likert scale, from "Strongly disagree" to "Strongly agree." Example statements include "People would pay attention to what I have to say about YouTube video advertisements," "If I need to express myself to the YouTube corporation, I can find a way," "I can convince others to accept suggestions I might have about YouTube advertisements," and "If YouTube users cooperate, they can produce greater social forces." I conjectured that those in the *take back power* condition might, regardless of whether they had been assigned to *roast* or *no roast*, rate higher in empowerment, as they had been told that YouMercials could be a way to take back power from advertisers. However, I found that although the difference was not statistically significant, those in the *help advertisers* condition actually rated *higher* on the empowerment scale than those in either the *take back power* or *no purpose* conditions.

It appears that by priming *help advertisers* users to align themselves with those with perceived power in the context (advertisers), they actually felt more empowered and able to voice their opinions and enact change as a result. Conversely, those who had been explicitly assigned to *take back power* scored comparatively lower on the empowerment scale, suggesting that simply being told that the declared play could be a way to take back power was not sufficient for participants to feel empowered and feel as though their actions were balancing power structures.

At the same time, because empowerment and any specific goals and desires for the YouTube platform are self-defined, the forms of change that those in the *help advertisers* condition had in mind might not have aligned with the *take back power* form of empowerment that I had envisioned. Instead, *help advertisers* participants could have been more focused on improving the relevance or interest-level of advertisements as a form of empowerment, rather than resisting against the power status quo.

Summarizing RQ4: How might using YouMercials influence participants' opinions of and engagement with YouTube advertisements? Engaging with the declared play of YouMercials— as expected, and in keeping with the reported longer periods of time engaging with YouTube advertisements as a result of using YouMercials— enhances brand recall. However, it also simultaneously leads to less favorable impressions of those brands. This suggests that by using YouMercials, ads and their associated brands may have become more salient for participants, but participants may have simultaneously viewed those brands and ads through a more critical lens. Most participants' opinions of YouTube advertising didn't visibily shift after using YouMercials in-the-wild. For those that *did* shift opinions, they tended to shift in a more positive direction, suggesting that for a subset of participants, playing with a capitalist context has the potential to reinforce capitalist power structures and values. In addition, although I had anticipated those in the *take back the power* condition to have higher empowerment scores, those who were primed to align themselves with the advertisers— stakeholders holding relatively more power in the context— felt more emboldened to voice their opinions and enact change. When participants felt that those currently holding power (the advertisers, rather than other users) were the ultimate targets of their YouMercials play, it appears they may have viewed their play as more consequential and having more potential for impact, displaying a desire for not just playing *with* or *in response to* a capitalist power, but directly playing *on the side of* that source of power.

6.4 Discussion

With YouMercials, I found that while declaring play within a capitalist context (here, advertising on YouTube) may invite ambivalence and reflection, it does not necessarily invite subversion of that play to rebel against the power status quo. For the most part, participants accepted their assigned conditions and the rules of play declared by YouMercials. Although the opinions of YouTube advertising I measured were generally negative-leaning, those that did resist their assigned YouMercials conditions (and by extension, the rules of play) tended to act not in opposition to corporate powers, but rather, in support or service of them (e.g., those

in the *take back power* condition displaying support for brands and advertisements). Overall, there was far more ambivalence in relation both to the role and presence of advertising on YouTube than I had anticipated, with many reflecting mixed feelings. These mixed feelings may partially account for the general acceptance of YouMercials play and the assigned conditions; such indifference or ambivalence towards YouTube advertisers may have rendered participants more flexible and open to a play role in which they supported or lampooned, fought against or worked in service of the advertisers. In addition, although those who engaged in YouMercials play reflected on YouTube advertisements with more criticism than those who did not (based off the controlled experiment results), they also engaged more with YouTube ads and recalled brands more clearly, further reinforcing the ambivalence of declaring play within a capitalist context if the goal is to resist existing power structures or to call into question power imbalances in the status quo.

6.4.1 On Eliciting Critique of a Power-Ambiguous Context using *Research through Evocative Play*

One of my misplaced assumptions in conceptualizing YouMercials was the degree to which participants would be critical of and resistant to YouTube advertising. Based off my study of Turker Tales in the context of Amazon Mechanical Turk, I had posited that purposefully using a paternalistic *research through evocative play* study design, wherein participants are encouraged to view the researcher as aligned with an oppressive source of power, could be instrumental in eliciting candid criticism of that power. Such a design, I conjectured, could even inspire inklings of shared activism in terms of how participants choose to influence other users in the declared play system with the artifacts they create. I thereby directly incorporated elements of a paternalistic design in YouMercials; a subset of participants were assigned to the *help advertisers* condition, wherein they were primed to think about how creating YouMercials could help advertisers create content that is more relevant and interesting to users.

However, I found that the effects of a paternalistic design maybe be highly dependent on the context. Although opinions expressed about YouTube and YouTube advertising in the surveys conducted for YouMercials often leaned on the more negative side, for the most part, participants were not so hostile to YouTube advertising so as to completely ignore or actively resist their assigned study conditions. This becomes especially pertinent with respect to those assigned to directly help advertisers in their creation of YouMercials, as most participants in *help advertisers* followed the assigned rules of play declared by YouMercials and created YouMercials largely in support of the products and brands. Moreover, when it came to the in-the-wild deployment engagement, some participants who had been assigned to conditions encouraging

critique of or resistance to YouTube advertising (*take back power*) actually ended up directly supporting advertisers, instead.

Thus, when the rules of YouMercials *were* violated, it was not in the direction I expected. These violations may have indicated resistance to the assigned condition—for example, a *take back power* participant might have resisted the notion that YouTube and its advertisements are an oppressive or restrictive force, and thereby purposefully chosen to create YouMercials in favor of YouTube advertisements. However, based off the reflections in surveys and the goals that participants expressed they were pursuing in creating YouMercials, which for the most part did not suggest explicit resistance to assigned conditions, it seems more likely that either (a) participants inadvertently leaned towards expressing their true feelings while creating YouMercials in-the-wild, which sometimes operated in support of particular brands or products or (b) participants were influenced by viewing product-supporting YouMercials created by others. For example, participants who already feel neutral or positive about advertisements who had nevertheless been assigned the *take back power* condition might have viewed other advertisements in support of the products or brands positively and followed suit, or subconsciously absorbed the structures other participants were following in creating their own YouMercials.

6.4.2 On Demand Characteristics in *Research through Evocative Play*

Part of the discrepancy in my expectations and participants' behaviors may stem from issues with academic research and demand characteristics, wherein participants interpret the goals and intentions of the study, and shift their behavior in accordance with interpretation; this dynamic can also be characterized as the Hawthorne effect or the observer effect [151], wherein participants change their behaviors in response to being observed. In most academic studies, researchers strive to mitigate such effects. Inspired by critical design and abusive game design, in *research through evocative play*, I advocate for drawing out, analyzing and considering the broader implications of these effects.

As compared to RoastMe, where I was not at all involved in the design of the play declarations, or Turker Tales, where I had very limited contact with participants during the study period, as all activities were facilitated through Amazon Mechanical Turk, I had much more direct contact with participants in YouMercials. For example, I reminded participants via email to complete the initial survey and to engage in the YouMercials extension for at least half an hour. Participants could also reach out to me via email before, during, or after the study period ended if they had questions or concerns. Where similar contacts were also allowed in Turker Tales, the

MTurk platform encourages anonymity. Email messages are sent through Amazon Mechanical Turk, whereby actual email addresses are obscured, and workers typically do not sign their names via this anonymizing system (though they can choose to do so). Moreover, at least a subset participants I recruited for YouMercials may have felt more closely aligned with me, and thereby more inclined to act in accordance with my (perceived) research intents and goals. For example, for YouMercials, my recruitment strategies included recruiting members of my local university community as well as through social networks and also encouraging snowball sampling, which may have impelled participants to view me as a researcher-designer (directly or indirectly, if the study had been recommended by a friend or colleague) with similarities to them, even if they did not fully share my (perceived) research objectives.

Ultimately, because I am not formally affiliated with either YouTube or advertisers, nor with an organization or movement striving to limit YouTube advertising or "take down" YouTube in some fashion, participants are bound to see me through traditional research-participant lens, and act accordingly. However, this ostensible limitation has its advantages. With YouMercials, I was able to assign participants to different conditions and observe the effects of that manipulation precisely because I held a more neutral role, and could plausibly be conducting research either in support of or in opposition to YouTube and YouTube advertisers. Moreover, in my formulation of *research through evocative play*, I argue that role of the researcher and designer in *research through evocative play* should not be mitigated or ignored, but rather, brought to the fore and directly discussed and analyzed. This does not make for an easy or unambiguous analysis; like Research through Design, *research through evocative play* is more valuable in its ability to pose new questions than its capacity to answer existing ones.

6.4.3 On the Ethics of Designing for Play where Power is Ambiguous or Inequitable

I advocate for *research through evocative play* as a closed study design; although *research through evocative play* may point to opportunities for designing for play in a given context, the play declarations engendered by a *research through evocative play* study is not intended to be viable, long-term, or even appealing. Optimizing for enjoyment in *research through evocative play*, therefore, should not be the priority *unless* exploring the boundaries of the play design's potential to engage and distract from ambiguity is part of the specific study design.

Imagine for example, if we were to create an optimized version of YouMercials that participants genuinely enjoyed— a magic circle of play that participants entered into with with their ready, ludic selves. As I saw in my study, such a design could

be highly problematic if one of the aims were still to encourage users to critique the power status quo on YouTube. For example, even if actively engaged in playful criticism of YouTube or advertising practices, participants would be likely to engage more heavily with YouTube advertising as a result, thereby further contributing to the capitalist system. They might even adopt more positive attitudes towards YouTube advertisements through their engagement in play. Of course, this also brings back to the fore the role of the researcher-designer. If I enter into play design with a subjective goal of the play, such as say, encouraging activism against and critique of YouTube advertising, then I as the designer act as a power structure imposing my will on the players, which carries its own set of ethical concerns.

Moreover, is it ever possible to resist elements of a context while simultaneously engaging with and participating with those elements? *Research through evocative play* as I have presented it in these three studies positions the space of play as belonging to, in conversation with, or taking elements of either a context or set of behaviors. However, perhaps a truer form of resistance would be for that declared play to divorce itself from the original context or set of behaviors altogether; criticism of and attention to a context, as I saw with my research in YouMercials, could have the effect of giving more power and visibility to oppressive elements of that context.

Research through evocative play, then, allows researchers to explore play concepts while still including a measure of protection; *research through evocative play* is not play itself, but a research method built around play declarations to explore a context, participants' relationship to that context, and the implications of play within or in relation to that context. But it is far from perfect, and far from danger-free. I have far presented *research through evocative play* "in action," first showing its seedlings in RoastMe, then showing it beginning to burgeon in Turker Tales, and lastly presenting YouMercials as a more full-fledged model of the *research through evocative play* approach. In the following chapter, I provide a practical guide to using *research through evocative play* that synthesizes key features of the research method, lessons learned, and important steps to take in preparing for, designing, implementing, and analyzing the results of a *research through evocative play* study design. I hope any readers that might consider using *research through evocative play* in their own work, either in the proximate or distant future, will find the next chapter especially helpful.

A Guide to Applying *Research through Evocative Play* Methods to Your Own Projects

<div style="text-align: right">7</div>

It took me years of trial and error, circling, and searching before I began to define the work I do as *research through evocative play*, and to feel confident in presenting *research through evocative play* as a valid form of research. This chapter, then, is an effort to shorten the path for others who may want to use *research through evocative play* in their own work. This chapter is a mixture of practical advice, references to relevant resources that can help you, and a good deal of (well-intentioned, and hopefully somewhat helpful) cheerleading and encouragement largely based on my own experiences, perspectives, and struggles with conducting *research through evocative play*; focus on the aspects of this chapter that most serve you.

7.1 Whom *Research through Evocative Play* Can Benefit

7.1.1 *Research through Evocative Play* "For All"

The *research through evocative play* approach I lay out is meant to be inclusive and flexible, and I thus encourage researchers and practitioners from multiple disciplines to consider making use of *research through evocative play*. Design and play research and theory inspire and inform *research through evocative play*, but I consider myself neither a designer nor a play or games scholar, as I am not formally trained in these disciplines. My point is: you do not need to be an expert in design, play, or even human-computer interaction to use *research through evocative play*. At least some background in one or more of these three areas will definitely be helpful, so if you are well-versed in one or more of these disciplines, you're already ahead of the game. At the same time, I'll also caution that given my own background in HCI research, my presentation of and envisioned uses of *research through evocative play* are bound to be skewed towards HCI researchers. Those who study, work, or practice in domains outside of HCI and academia may find they need to adapt *research through evocative play* as I've presented it so that it better fits their needs.

For all readers, I'll provide you with what I see as some especially helpful background reading and skill-building activities to get you up to speed. In addition, I'm sure that your knowledge and experience *outside* the disciplines of HCI, design, and play/games will also enrich your work in *research through evocative play*, so I highly encourage for you to both adapt *research through evocative play* to your particular field or discipline, as well as to integrate your own experiences and expertise *outside* into the *research through evocative play* process. It can be easy to feel intimidated by the unknown, and to focus on what you are and know *not*, but please do not let perceived lack of knowledge hinder you from using *research through evocative play*. As an HCI researcher (with prior academic and work background in public policy, non-profits, and foreign language-learning), I believe strongly in the value of multi- and inter-disciplinary work and collaboration. My presentation and application of *research through evocative play* is necessarily biased by my education in human-computer interaction, and the methodology of *research through evocative play* is strongly influenced by design and play research (both of which are outside my areas of expertise), However: non-HCI researchers, non-designers, and non-games and play researchers and practitioners are highly welcome.

7.1.2 *Research through Evocative Play* for HCI Researchers

That said, given my background in HCI, I'd also like to add some additional thoughts for those HCI researcher-readers among you. Is *research through evocative play* right for you? In keeping with my overall inclusive vision of *research through evocative play*, I would again argue that *research through evocative play* can be of benefit regardless of your sub-domain in HCI, but that you should keep in mind the intended goals and projected outcomes of *research through evocative play* before you embark on a *research through evocative play* project.

If you are a qualitative or quantitative empirical researcher in HCI, meaning you primarily contribute findings about a context based on observation and data gathering [222], then *research through evocative play* can be considered another tool in your arsenal to better understand and report on a given context. (In the following section, I'll address in more detail how and when to choose between *research through evocative play* and other empirical methods).

However, this does not imply that *research through evocative play* is exclusively for empirical (HCI) researchers. For other HCI researchers seeking to achieve other types of contributions in the field, *research through evocative play* can be considered a useful method along your research path. For example, if you primarily work in the area of systems design (for example, your contributions tend to be artifacts that

"reveal new possibilities, enable new explorations, facilitate new insights, or compel us to consider new possible futures" [222]), *research through evocative play* may be useful to you as a method to better understand a context or population for which you want to propose and design an artifact contribution. However, the declared play itself of the *research through evocative play* project is not, in itself, such an artifact contribution. Just as one might conduct a series of empirical, qualitative, semi-structured interviews in order to inform the design of a novel system, HCI researchers seeking non-empirical contributions can make use of the empirical findings afforded by *research through evocative play* to ground, inform and inspire their research.

Similarly, if you are a designer in HCI, bear in mind that *research through evocative play* can help you better understand— and thus design for— a context, but again, the artifacts created through *research through evocative play* are not the end goal or main contribution of the work. *Research through evocative play* is not the design of play itself, but rather, the use of play declarations to study power dynamics within a given context. *Research through evocative play* shares similarities with critical design in its reflective qualities, but differs from much design research in HCI in that the artifact itself is not the contribution, the research does not directly seek to achieve an improved future state (though it may provide insights into future design possibilities), and the specifics of the methodology (namely, using declared play as a research tool) are unique to *research through evocative play*.

7.1.3 *Research through Evocative Play* for Those who Study Power Dynamics

As I've discussed, *research through evocative play* is designed to highlight and reveal the nuances of power dynamics in a context, and as such, it can be of especial benefit to those whose work or interests center around power dynamics and power inequities (be that within or outside of the field of HCI). This might include, for example, power dynamics related to race, class, gender, sexuality, nationality, language, political affiliations, education, capitalist or market forces, or context-specific hierarchies or precedents. Although it's hard for me to envision an area of research, design, or production that *wouldn't* benefit from a closer consideration of power dynamics in a context, *research through evocative play* should be especially appealing to individuals who are already invested in understanding contextual power dynamics.

In conducting *research through evocative play*, I should note that it's perfectly valid to enter the context with a pre-defined area of inquiry or interest, but to simultaneously remain open to findings related to other forms of power dynamics. For example, in my Turker Tales study, I paid special attention to class and labor relations within

a capitalist context. It's important to acknowledge that this likely biased (but did not invalidate) the nature of my findings. At the same time, the design of *research through evocative play* should allow for other power dynamics that the researcher had not previously considered to emerge. In my work with Turker Tales, I had not fully considered my personal role as a researcher— and thereby another requester— in that context, nor had I considered how participants' individual self-expressions in the declared play might open up new insights into Turkers that interact with but also extend beyond crowd work, itself. Likewise, in YouMercials, I was interested in how YouTube users would interact with and view their own actions in relation to YouTube advertisers, but had misplaced assumptions about the nature of how they viewed those relationships.

In short, then, *research through evocative play* can both reveal insights into the particular set of power dynamics that initially appears most relevant to you (the researcher) about a context, but can (and likely will) also reveal additional power complexities that extend beyond your initial grasp of a context. So long as you are (a) willing to admit and reflect on your own biases and (b) open to new revelations that might call into question your original stance on a subject or set of power dynamics (or even which set of power dynamics are most relevant or pertinent to a context), then *research through evocative play* is a methodology that can benefit your work.

7.1.4 *Research through Evocative Play* for Non-researchers

I also want to emphasize that *research through evocative play* need not be constrained to purely academic research. You may not label yourself as or personally identify as a researcher, but can still take on the role of a researcher within *research through evocative play*. As I mentioned in the Preface (Chapter 1), I am looking forward to exploring the potentials of *research through evocative play* outside formal research contexts, and encourage others to do the same. For example, if you are personally curious about a particular context or area, even if you have no concrete end goal (e.g., a publishable research paper) in mind, I'd encourage you to consider *research through evocative play* as a means to edification and self-reflection and revelation that can be rewarding and fulfilling on a personal, non-professional level, If you are experiencing a period of personal or professional flux and uncertainty, *research through evocative play* can be a way to help you define your own path and realize your own values and interests, as I have found in my own experiences with the methodology.

Research through evocative play can also have practical value to those in non-research based professions and pursuits. Perhaps you are, for example, a game or play designer with an end goal of designing commercially viable play experiences. You

may already be well acquainted with methods such as play-testing [36], but may not consider research as important to your design process. I would suggest that better understanding the context for which you are designing will challenge, enrich, and transform your approach to play design, and that designing for declared play that is ambiguously playable, laden with questions and uncertainties, or perhaps even specifically commercially *un*viable can reveal rich insights into how to design play for a given context or end that *is* viable.

7.2 When to Use *Research through Evocative Play*

But *when* exactly will applying the *research through evocative play* methodology prove most useful? I put forth that there are four elements that should be present when using *research through evocative play*. *Research through evocative play* is an appropriate methodology if the following apply to you and the context you are exploring:

1. If you care about understanding power dynamics in a context

2. If you are dealing with wicked problems

3. If you are outcome-ambiguous (or at least, willing to pivot)

4. If you are invested in your own well-being as a researcher

I'll now explain each of the four in a bit more detail.

7.2.1 If You Care about Understanding Power Dynamics in a Context

As already discussed, *research through evocative play* is a methodology specifically intended to help reveal the nuances of power dynamics in a context. Thus, if your research or field of practice centers around a certain kind of power dynamic and/or imbalance, then *research through evocative play* can be a way to deepen your understanding of how those power dynamics *as well as* other power dynamics you might not normally consider central to your work, function and interact. For example, if you are invested in understanding and fighting against racial injustices in particular contexts, then *research through evocative play* could be helpful in revealing complexities of race-related power dynamics in a context, and can simultaneously

bring to light how these power dynamics might intersect with or relate to, say, class, gender, or education dynamics in a given context. *Research through evocative play* can be especially helpful in revealing points and lines in which ethical boundaries related to a certain power dynamic are ambiguous or ambivalent.

Note that I continue to speak in terms of *context specificity*. *Research through evocative play* explorations are meant to be embedded within and thereby provide insights into specific contexts, not broadly applied to make claims about power dynamics that are divorced from context. That said, *research through evocative play* findings still have the potential to have broader implications beyond one given context. In the same way that studying a particular context within HCI may provide insights that could apply more broadly to other, similar contexts, findings from *research through evocative play* may often have relevance beyond the specific context at hand. For example, findings from YouMercials have relevance to other online contexts in which targeted advertising plays a prominent role.

7.2.2 If You are Dealing with Wicked Problems

As discussed in the Background chapter (Chapter 3), wicked problems [24, 182] is a concept from design and design thinking, and refers to problems to which there is not a clear scientific or engineering solution. In Rittel's definition, wicked problems are "a class of social system problems which are ill-formulated, where the information is confusing, where there are many clients and decision makers with conflicting values, and where the ramifications in the whole system are thoroughly confusing" [182]. In keeping with this definition, I argue that virtually any problem that involves complex, imbalanced, and/or ambiguous power dynamics is, by nature, a wicked problem, as there is (almost) never a computational solution to such issues.

Rittel [182] lays out the ten main characteristics of wicked problems, of which I'd like to highlight one: "For every wicked problem there is always more than one possible explanation, with explanations depending on the Weltanschauung of the designer." Note that "Weltanschauung" simply refers to one's perspective, philosophy, or worldview. In *research through evocative play*, contexts are chosen because the researcher is interested in better understanding how power dynamics operate in that context; the interplay of these dynamics and issues engendered by these dynamics constitute wicked problems. As such, there is not one, clear explanation of those dynamics. Although *research through evocative play* provides insights into the dynamics of the context, it does so through the subjective lens of the researcher-designer (although I should reiterate here that going through the process of *research through evocative play* often can alter this subjective lens, and draw to the fore previously unrecognized biases and limitations in the views of

the researcher-designer). The presence of subjectivity and bias thus is an integral component of *research through evocative play*, and should be acknowledged and analyzed by the researcher during the *research through evocative play* process (more on this later).

Essentially, you should never use *research through evocative play* if you are aiming for a clear-cut, unambiguous solution to a problem. *Research through evocative play* offers insights, *not* solutions. Note further that where design seeks to produce solutions to wicked problems, by contrast, *research through evocative play* seeks to provide empirical insights into wicked problems (that can, in turn, later contribute to and inform a solution).

7.2.3 If You are Outcome-Ambiguous (or at least, willing to pivot)

Those using *research through evocative play* should go into the process outcome-ambiguous, at least in theory. If you go into a context or problem space with a specific concrete objective in *how* you want play to be enacted in the space, then *research through evocative play* is not the appropriate method. For example, let's say you are positioned in the crowd work context. If in your research or practice, you are seeking a way to use play to better engage and improve retention among crowd workers, then *research through evocative play* is not the right method for you, at least not as an end in itself. *Research through evocative play* could, however, prove useful as an exploratory stage of your research, but you should be prepared for the findings from *research through evocative play* to alter your original mindset and goal orientations in the context. To give another example, if you are studying or designing for the context of Reddit or a similar platform and specifically want to prevent sexist comments, then again, *research through evocative play* as an end in itself will not meet your needs. Instead, *research through evocative play* could be used to help illuminate and better understand the gender and sexuality-related power dynamics in the context, which could then later be used to design in a way that addresses issues of sexism in the context with a more nuanced understanding of the power dynamics involved.

Of course, as I've alluded to at multiple points in this document, it is quite nearly impossible to be fully outcome-ambiguous; even when we strive to be so, we often are operating with preconceived notions and assumptions that may bias us in favor of certain outcomes. I saw this in myself in my study of YouMercials, and my inclination towards favoring empirical outcomes that would emphasize peer empowerment and critiques of advertisers. Having some level of bias, then, is unavoidable, and does not violate the "rules" of *research through evocative play*. However, it is "pivotal" for the

researcher-designer to be willing to pivot, self-analyze, and self-reflect throughout the research process, potentially coming to terms with aspects of their own biases that they had not fully recognized or acknowledged previously. If as a researcher, you will only consider your work successful if a certain set of findings apply, then *research through evocative play* is not the methodology for you. Instead, you should consider the gamut of possible outcomes, and consider a range of possible outcomes compelling and informative.

7.2.4 If You are Invested in Your Own Well-Being as a Researcher

This point might seem even a little silly at first blush, but I actually think that the personal well-being component is one of (or perhaps *the*) most critical advantage of *research through evocative play*. Recall that in my formulation of *research through evocative play*, it is perfectly fine if you are not conducting *research through evocative play* within the context of formal academic research. However, with this point about researcher well-being, I am especially speaking in terms of formal academic research and associated settings, though the ideas here may apply to "informal" researchers, as well.

Termed a "mental health crisis" [60], a number of recent studies have shown that graduate students and PhD students, especially, are at high risk for mental health issues like anxiety and depression [60, 141, 11, 223, 137]. For example, one study found that graduate students are more than six times as likely to experience depression and anxiety as compared to the general population [60], and another study (specifically focused on Harvard Economic PhD students) found higher than average reports of feelings of loneliness and suicidal ideation, and lower than average reports of feeling that one's work is useful (as compared to Economics faculty and the working age population), with mental health issues appearing to increase as time spent in the PhD program increases [11].

Although evidence suggests that mental health issues might be especially prevalent and concerning while researchers are in the process of going through a PhD program [137], with mental health probems more prevalent among PhD students than the highly educated general population or employees, these issues may not end once researchers graduate from PhD programs. First, although they vary by field, job prospects are not always promising for PhD students, with 40% of doctoral students not having secured a job at the time of graduation (as of the most recent national survey of PhD students conducted by the NSF) [69]. Moreover, researchers at any level (from student to faculty, for example) are at high risk of imposter syndrome, a psychological term describing an internal experience of intellectual phoniness

wherein those affected doubt their accomplishments and have a persistent, often internalized fear of being exposed as a fraud [38, 39, 102, 167]. As Hutchins puts forth, imposter traits, "... might be further heightened within the 'publish or perish' academic culture where performance targets are often vague, support can be inconsistent, and a highly competitive research and funding climate may inadvertently create a setting conducive to feelings of self-doubt and fraudulence" [102].

Re-framing the work of a PhD student through a lens of play may be one small, but potentially still significant means by which to combat mental health travails. Engaging in play itself has the potential to benefit mental health [88], and play can help create enjoyable experiences for those involved— "fun," in a nutshell [135]. Moreover, the creativity involved in engaging as a researcher in *research through evocative play*, which requires taking on attitudes of openness, flexibility, and playfulness, can also have positive affects on one's mental health [46, 149, 81]. Engaging in creative thought, exercises, and activities can in turn enhance one's work (in this case, research) [142, 130, 4, 218]. *Research through evocative play*'s emphasis on self-reflection as the researcher/designer/player/study co-participant could also be one means to combat imposter syndrome. In psychology, "self-distancing" [129], or increasing the psychological distance from your own perspective when assessing events, allows individuals to better cope with negative emotions such as stress. By considering the researcher-self as another study participant to be observed and analyzed, *research through evocative play* also allows for self-reflection and even self-criticism that other research methodologies do not directly encourage. Rather than encouraging researchers to occlude bias and study imperfections, which I posit may contribute to feelings of fraud (largely based on my own experiences and informal observations of my peers' experiences), *research through evocative play* asks researchers to directly address and acknowledge such biases as part of the research process; biases and criticisms of the study design are positioned as opportunities for reflection and further exploration rather than instances of failures. *Research through evocative play* encourages a form of self-distancing that may encourage curiosity about one's role and participation in the research rather than self-judgment that could lead to symptoms of the imposter syndrome [127, 102, 167, 39, 38].

I have not studied the benefits of *research through evocative play* to researchers beyond my own personal experience, but I can say that using *research through evocative play* has often helped me to re-frame and reconsider my work as a researcher from an angle that I have found to be more personally fulfilling than other research methods. Unlike most other research methodologies, *research through evocative play* engages me directly as the primary *player* in the scenario. By positioning the researcher as an active participant in the research rather than as an objective, outside observer, and by encouraging self-reflection and self-analysis in the research process, *research*

through evocative play encourages me to view myself from an angle of curiosity rather than of judgment. By bringing play to the fore, *research through evocative play* also encourages me to take myself, my role in academia, and my perceived level of success in research less seriously.[1]

Here, I am especially inspired by Bogost's work as articulated in his book *Play Anything* [20]. Bogost argues that rather than enduring or railing against the boredoms, frustrations, and limitations of everyday life, we can accept them and view them through a lens of play (e.g., the boundaries of a magic circle), and subsequently derive enjoyment and fulfillment from even the most mundane aspects of life. When viewed through a non-playful lens, research can often be daunting, overwhelming, and downright panic-inducing. When viewed playfully, our failures and frustrations in research can be seen as exciting challenges. Of course, as my own experience can also attest, *research through evocative play* is no panacea to issues of mental health as they relate to conducting research or completing a PhD. Moreover, this is not to say that other methodologies cannot be approached playfully, or are bound to result in disillusionment. However, unlike *research through evocative play*, other methodologies are not specifically designed to engage the researcher as a player. Thus, I firmly believe that *research through evocative play* is a step in a helpful direction when it comes to the mental well-being of researchers.

7.3 Comparing *Research through Evocative Play* to Other Research Methods

As an HCI researcher, I am most comfortable comparing *research through evocative play* with other research methods commonly employed in HCI, and thus will focus my attention there. I will ask that readers who are researchers and practitioners from other disciplines bridge my discussion of HCI research methodologies to apply more directly to their particular disciplines and methodologies. In considering the primary research, I draw primarily from Wobbrock and Kientz's overview, "Research contributions in human-computer interaction" [222] as well as Olson and Kellogg's *Ways of Knowing in HCI* [164]. Wobbrock and Kientz separate the types of contributions in HCI into seven types:

1. *Empirical.* Empirical contributions provide new knowledge based on observation and data gathering, can be qualitative or quantitative in nature, and can emerge by means of a variety of methods, including experiments, user tests,

---

[1]Readers might note a parallel here to the ways in which playing with harsh humor and self-presentation allowed RoastMe roastees to take themselves less seriously.

field interviews, surveys, focus groups, diaries, ethnographies, sensors, and log files.

2. *Artifacts.* In HCI, researchers can also contribute new knowledge by creating interactive artifacts and prototypes, including new systems, architectures, tools, toolkits, techniques, sketches, mockups and envisionments.

3. *Methodological.* Method-based contributions inform *how* researchers can carry out their work, improving how we discover, analyze, or create in research.

4. *Theoretical.* Theoretical contributions explain why certain phenomena exist, providing new concepts , definitions, or frameworks for understanding (can be qualitative or quantitative).

5. *Dataset.* Dataset contributions are new corpi that are useful to the research community.

6. *Survey.* "Survey" here refers to meta-analyses or literature reviews on work done (not survey instruments and attendant contributions, which are empirical contributions).

7. *Opinion.* Often presented in the form of essays or arguments, opinion contributions aim to persuade, not only inform.

7.3.1 *Research through Evocative Play* as a Method to Facilitate Empirical Contributions

In this dissertation, by introducing the *research through evocative play* methodology, I claim to make a methodological contribution, as through my introduction of *research through evocative play*, I provide a new way of *how* we can carry out research in HCI and related disciplines. However, to be clear, in *using research through evocative play*, researchers should expect to arrive at empirical contributions in that *research through evocative play* provides new knowledge and data about power dynamics in a given context. *Research through evocative play* also shares similarities with methods intended for artifact contributions, but differs in that the ultimate research contribution of *research through evocative play* lies in the understanding through creation, not the creation itself. When, then, should we choose *research through evocative play* over other empirical research methods? I've touched upon the situations and goals for which *research through evocative play* is especially suited, but in what follows, I more clearly define the value of *research through evocative play* in comparison to other research methodologies.

7.3.2 *Research through Evocative Play* versus other Empirical Contribution Methods

Before I dive in here, I want to clarify a couple of points. First, readers should keep in mind that there are not clear "rights" and "wrongs" when it comes to choosing research methods; instead, there are tradeoffs. Just as there is not a clear-cut answer to when you should choose, say, contextual inquiry over semi-structured interviews, or a large-scale survey over large-scale data analysis, I cannot provide definitive, blanket statements regarding when to apply *research through evocative play* over other methods. Instead, I can provide guidance and advice on the decision. I've already covered some of the ways in which *research through evocative play* can be advantageous as a research method. For example, I spoke at length about the unique benefits of *research through evocative play* for the well-being of the researcher, and the relevance of *research through evocative play* to understanding power dynamics in a context.

Second, readers should be aware that *research through evocative play* also *subsumes* other empirical research methods under the larger framework of the methodology. For example, *research through evocative play* involves designing and declaring play in a context, and then observing how participants choose to playfully engage, disengage, or oppositionally engage in the declared play. The researcher's observation of participants' engagement in the declared play could take the form of semi-structured interviews (similar to my study of RoastMe), or instead focus on collecting logs of engagement, as in my study of Turker Tales. The declared play could be set up as a controlled experiment, and could involve a post-survey in addition to data collected directly from play engagements, as in YouMercials. Here, we can draw parallels to research through design. Research through design defines itself as a methodology that involves generative processes of ideation and produces an artifact or prototype. However, the research process of research through design can involve the use of methods such as controlled experiments e.g., [217], interviews, e.g. [122], and field deployment user studies, e.g. [140]. The core of *research through evocative play*, by contrast, lies in the design, creation, and implementation of declared play in a context, and the subsequent study of behaviors and engagements that emerge in response to those play declarations.

Thus, a *research through evocative play* study design might very well involve the use of semi-structured interviews, but the methodology is quite different from a pure interview study in that it involves the design and implementation of declared play in a context. For example, although RoastMe was a precursor to *research through evocative play*, let's imagine for the moment that it was a fully *research through evocative play* study, and that I had designed and implemented RoastMe myself as a

way to (for example) understand power dynamics involved in online photographic self-presentation. I could have then interviewed participants after the declared play experience about their engagement in the declared play, which is similar to the work I actually did in my study of RoastMe.

But had I chosen to use semi-structured interviews *in lieu of* the *research through evocative play* approach, in that case, I might have asked participants to describe their existing experiences with online photographic presentation. I might have gained some useful insights into the context from such an interview approach, but it would have been a different set of insights than those obtained through *research through evocative play*. Recalling again Sicart's work on how play engagement can help reveal nuances of a context [200], *research through evocative play* can reveal aspects of power dynamics that other empirical methods miss.

Where there are always opportunities to engage participants in empirical methods like interviews and surveys *after* implementing a *research through evocative play* design, the *research through evocative play* design itself allows for the observation of engagement in declared play, and this engagement (regardless of any follow-up interviews or surveys, and regardless of whether or not a controlled experiment was used in tandem with *research through evocative play*) reveals aspects of context. *Research through evocative play* is especially useful in that it allows for the observation of power dynamics and interactions that would not otherwise be easily observable. For those who study power dynamics in a context, it can be especially difficult to fully observe and understand those dynamics when aspects of those power dynamics are occluded by the context. If I were to gain access to use logs from Amazon Mechanical Turk and user engagement or activity during YouTube advertising, I would learn very little about participants' views of the contexts, or their views of other inhabitants in the context, because the original contexts do not allow for such self-expressions. In the contexts of Amazon Mechanical Turk and YouTube advertising, participants are normally unable to directly communicate with one another. The ambiguity of the declared play in *research through evocative play* requires participants to make choices and decisions as to how they will engage, and by those choices reveal their own orientations towards the context and other participants in the context.

Circling back to the previous RoastMe example I discussed, you might wonder why I couldn't just have asked participants about their opinions about a context, or other inhabitants in a context. Here, I return to the concept of tradeoffs; although *research through evocative play* may be a less direct route to obtaining certain insights, it also has the potential to reveal more nuanced insights that other empirical methods might miss. Semi-structured interviews and surveys, for example, might well provide some insights into power dynamics. However, they suffer from both self-report bias (we are not always accurate judges of our own behaviors and motivations), as well

as the "hypothetical" problem; participants may *think* they know how they would behave in response to certain stimuli, or if certain aspects of a context were modified or manipulated. But only by actually implementing and observing such modifications can we truly observe participants' reactions and behaviors.

To give an example from YouMercials, as part of the study, I *did* directly ask participants in the controlled experiment survey about their opinions of YouTube advertisements. The opinions expressed tended to skew negatively, and emphasized the annoying and intrusive elements of YouTube advertisements, as I expected. However, when I actually observed participants in their use of YouMercials, a different, more nuanced view of user-advertiser power dynamics emerged. Contrary to my assumptions and expectations, participants were more pliable in their support of YouTube advertisements than their self-reported opinions implied. Moreover, by engaging in YouMercials, participants reflected on and gained new insights into their own relationships to YouTube advertisements and other viewers. For example, some users questioned their own practices in relationship to YouTube ads, wondering if they should actually watch the ads in the future rather than use ad blockers so as to better support YouTube content creators. Others were surprised by their own engagement in YouMercials, finding themselves easily persuaded to take a position (e.g., helping advertisers) that conflicted with their own purported set of values.

Similarly, with Turker Tales, I did probe participants about their willingness to engage in tangential play during crowd work in the initial exploration phase of the study as part of the speed dating process. I gained some insights into how participants viewed their role in Amazon Mechanical Turk in relation to other Turkers and requesters, and some self-reports as to how they might hypothetically engage. But it was only by actually designing and implementing the *research through evocative play* design of Turker Tales that I saw how participants revealed unique aspects of their identity that are normally veiled by the platform's design, critiqued power dynamics on the platform, and communicated with other participants via shared cultural and political understandings.

Furthermore, *research through evocative play* is unique from other empirical methods I have mentioned in that it places the researcher in an active role as a co-participant, and encourages researchers to take an inward gaze in the process. This also reveals further insights into power dynamics in the context, in which the researcher has necessarily become involved (if she was not already so prior to the study). For example, with Turker Tales, it was from the *research through evocative play* design that I was able to draw out participants' animosity and distrust torwards me as a researcher-requester, which led me to further consider the roles and responsibilities of researchers in crowd work platforms. Where other empirical methods strive to minimize the subjectivity of the researcher, such that studies that do *not* adequately

do so may be considered lacking or unsound, *research through evocative play* instead draws the biases of the researcher to the fore, asking researchers to place themselves under the lens of study as well as part of the research process.

In my own experience, I have found this emphasis on self-reflection to be a critical advantage of *research through evocative play* that goes hand-in-hand with the discussion of researcher well-being. In presenting my work as a researcher, I often feel pressured to aggrandize rather than reflect critically on the work and my role in it. I believe this pressure is detrimental to both the individual and to the research itself. By hiding their own bias and subjectivity, the researcher may then experience symptoms of imposter syndrome, viewing themself as a fraud by silencing critical reflection. Researchers may then be discouraged from digging deeper into a particular research path or direction, viewing the presence of their own subjectivity and biases as research flaws rather than as opportunities for exploration and knowledge gain. The research community at large may also suffer from this lack of self-reflection, wherein researchers are encouraged to view their peers' work (and evidence of bias therein) through a lens of judgment instead of curiosity.

Thus, the emphasis on self-reflection as well as the specific requirement of the adoption of a lusory attitude and engagement as a player for researchers sets *research through evocative play* apart from other empirical research methods. In *research through evocative play*, researchers are seen as visible, active participants. Research is framed directly as a form of play in which the researcher engages, an activity to be approached with curisoity, openness, and most importantly, playfulness. By giving attention to the position and humanity of the researcher, *research through evocative play* may thereby have unique advantages over other empirical methods in not just the outcomes of the research, but in the *experience* of conducting research.

7.4 Preparing Yourself for Conducting *Research through Evocative Play*: Recommended Reading and Skill-building

The following is a curated list of resources. Certainly, readers may draw inspiration and guidance from any of the sources I have cited throughout this document, but here, I will highlight what I see as the most essential readings to prepare yourself to begin conducting *research through evocative play*. Where possible, I opt for informal, short, primers in this list. In some cases, you may want to start with just the abstract (or even a summary of a book) before diving fully in, though this will depend largely on your personal preferences. I'll remind you that *research through evocative play* is

very much about learning through doing, so I'd encourage you to not let yourself get *too* bogged down with reading. I should also acknowledge my own bias here, which is that although I love reading certain texts, I also can get easily emotionally affected by what I read, and too much reading can send me into existential crisis mode, which I have found is generally *not* an effective means to prepare oneself for conducting research. I also have a tendency to never feel I am truly "ready" or knowledgeable enough about a subject, and reading can end up becoming a form of fearful procrastination to avoid actually "doing the thing."

If this sounds a bit like you, then you have my full permission— and in fact, encouragement— to limit your reading. You do not have to be an expert in a given domain or topic matter before beginning *research through evocative play*; at least to an extent, allow the process of *research through evocative play* to reveal gaps in your knowledge and expertise, and continue learning and improving as you conduct research. If you are making you way through the readings in this list, from elsewhere in this document, or from other sources, and the readings are inspiring you, teaching you, helping you generate new ideas, causing you to questions your assumptions, altering your perspectives, and so forth, then by all means: keep reading (and taking notes)! If, however, you hit a point where more reading is getting you down, and/or making you feel confused and uncertain as to whether you should even pursue your idea or area of research at all, or if you consider yourself a terribly slow reader [2] and feel overwhelmed by all the texts you feel you "need" to read in order to move forward with *research through evocative play*, then it's definitely time to stop reading, and start *doing*.

Alright, now that we've gotten the cautionary notes out of the way, here are the readings and skill-building activities from (1) human-computer interaction and related domains and (2) play theory and research that I most recommend to help equip you to conduct *research through evocative play*.

7.4.1 Readings and Skill-building in HCI and Related Domains

Human-computer interaction is a multidisciplinary field that integrates perspectives from a variety of other fields, including (but not limited to) cognitive science, psychology, other behavioral sciences, design, learning sciences, and computer

---

[2]You're probably not that slow a reader. In my experience, I find that most people (including myself) consider themselves "slow readers." Perhaps we're all just reading at the normal, human speed, and have unrealistic expectations about how fast we *should* be able to read. In any case, I say embrace being a slow reader; consider slowing yourself down even further! I find reading more rewarding the slower I take the process, allowing myself extra time to stare into space and reflect on what I'm reading.

science. As an HCI researcher, I often don't draw distinct boundaries separating these domains from HCI, more broadly, and thus I will present readings as broadly "HCI," though some readings I suggest here may actually hail more specifically from domains such as, say, social psychology.

Luma Institute. *Innovating for people: Handbook of human-centered design methods*. **LUMA Institute, LLC, 2012.** First, if you have no prior exposure to HCI research or practice, then I highly recommend the Luma's Institute's Innovating for People handbook. Intended for practitioners of Human-Centered Design, the book is helpful for HCI researchers and practitioners alike, and presents 36 methods well organized into sections of "Looking, Understanding, and Making." Each method is presently very clearly and briefly, making it easy to jump in. I especially like the Luma book and its sections on Concept Ideation and Modeling & Prototyping for the initial exploration stage of *research through evocative play*. Use the activities in the book to get the brain juices flowing, think about the context you are studying from different angles, and brainstorm ideas and possibilities for declared play therein. The section on "Rough & Ready Prototyping" is similar to articles and blog posts you can find online on paper prototyping; supplement with Internet resources as needed. Especially if you plan to work in digital or interactive domains, then you'll want to use paper prototyping at some point during the process (personally, I prefer to use paper prototyping during the brainstorming process).

Jakob Nielsen. *10 Heuristics for User Interface Design*. **1994.** The Nielsen Norman Group website[3] is a great resource for learning about or brushing up on standards of user experience research. It's geared towards user experience and design practitioners, but is also helpful for HCI researchers, and covers in condensed form some of the crucial concepts in user experience design. If you are designing for digital interaction, I would at the very least recommend becoming acquainted with Nielsen's heuristics, which are 10 general principles or rules of thumb for interaction design.

Even if you choose to violate certain heuristics with your declared play, it's important to be familiar with these general principles and aware of when you might be violating them. For example, a design that violates all of Nielsen's heuristics is likely to be totally unusable, and thereby will elicit very little information for you, as users will be unable to engage regardless of their perspectives on the concept of declared play in the context.

Conversely, keep in mind that designing your play superbly and in line with all of Nielsen's heuristics, making it as user-friendly as possible, could also negatively impact your *research through evocative play* project. In doing so, you are may be

---

[3]https://www.nngroup.com

actively persuading users to get on board with the declared play in ways that users may not even be consciously aware. This is most likely *not* the goal of your *research through evocative play* study.

Thus, I posit that a somewhat clunky design that partially violates certain heuristics (e.g., aesthetic design) may actually be quite useful for *research through evocative play*. Purposefully integrating some clunkiness or roughness into your design is not required for *research through evocative play*, but I tend to be of the mind that doing so increases the ambiguity of the declared play, and thus can be instrumental in drawing out participants' discomforts and pushing on boundaries in the context. Much of this will depend on your particular context and the nature of your declared play, but remember that it may not always be beneficial from a *research through evocative play* perspective to design for the most enjoyable, addictive, or otherwise persuasive form of play.

In general, before you officially deploy your *research through evocative play* design, you'll want to do at least *some* user testing or pilot testing; the Nielsen site also gives some nice overviews on user testing methods. In *research through evocative play*, such pilot testing needn't be extensive. In designing for *research through evocative play*, you are not attempting to create a polished, viable product that participants will *want* to engage in, as observing participant engagement rather than persuading users to engage playfully is the end goal. Similar to my argument regarding Nielsen's heuristics, you'll want to employ some (generally, informal) user testing to make sure your declared play is not totally unusable regardless of participants' unique perspectives and opinions, because then you'll miss out gaining insights into the nuances of power dynamics in the context.

Three recommended design readings. Of the design sources from HCI that I cited in the Background chapter (Chapter 3), there are three that I most highly recommend to give you more orientation on design work that inspired *research through evocative play*. While these readings might not be crucial to getting started right away with *research through evocative play*, they were quite critical to my own development of *research through evocative play*, and I thus include them here.

John Zimmerman and Jodi Forlizzi. "Research through design in HCI". in: *Ways of Knowing in HCI*. Springer, 2014, pp. 167–189.

William Gaver et al. "Cultural probes and the value of uncertainty". In: *Interactions-Funology* 11.5 (2004), pp. 53–56.

James Pierce et al. "Expanding and refining design and criticality in HCI". in: *Pro-*

ceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM. 2015, pp. 2083–2092.

Geoff F. Kaufman and Lisa K. Libby. "Changing beliefs and behavior through experience-taking." In: *Journal of Personality and Social Psychology* 103.2 (2012), pp. 1–19. Because as a *research through evocative play* researcher, you will need to come to terms with your own biases in the space, I highly recommend doing at least some reading in the space of bias. I have been very influenced by my advisor Geoff Kaufman's work in experience-taking, which touches upon concepts of bias and perspective. Remember, however, that unlike in Kaufman's study, with *research through evocative play*, the goal is not to directly *change* beliefs and behaviors through play declarations, but rather, to draw out, observe, and analyze those beliefs and behaviors. You might also want to check out a literature review on bias, such as:

Anthony G Greenwald and Linda Hamilton Krieger. "Implicit bias: Scientific foundations". In: *California Law Review* 94.4 (2006), pp. 945–967.

Qualitative and quantitative research methods: Skill-building. In order to conduct *research through evocative play*, you will need at least *some* familiarity with basic statistics or quantitative methods, as well as qualitative research. I recommend learning skills as you need them rather than trying to learn everything from the get-go. It's a good idea to read and keep abreast of analysis methods that exist, but you can build that knowledge gradually over time; don't try to learn everything at once, as (1) you'll just get overwhelmed and (2) you definitely *won't* ever use *all* the methods. In my own research, the research methods that have proven most useful to me are survey design and implementation, semi-structured interviews, and basic statistical analyses, so I will focus my recommendations on these areas. Note that some of the resources and methods— particularly those referring to specific software or web-based resources— may eventually become obsolete.

First, the Luma Institute's Innovating for People handbook I mentioned above [103] (focus on the "Understanding" section) is a good resource for introducing yourself to a variety of research methods that can be useful to HCI research and design. Now, to more specific resources.

To conduct **surveys**, I recommend reading relevant sections from:

> Elizabeth Goodman et al. *Observing the user experience: A practitioner's guide to user research*. Elsevier, 2012.

Check out Chapter 12, the sections on When to Conduct Surveys and How to Field a Survey. I personally like to use Qualtrics[4] software for creating surveys, which offers a free tier.

For **qualitative research**, I recommend:

> John W Creswell and Cheryl N Poth. *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications, 2017.

Note that for qualitative coding (covered in the book), I personally highly recommend using old-fashioned Excel (or your spreadsheet software of choice) to get the job done. I prefer more flexibility when coding and analyzing qualitative results, and much prefer spreadsheets to "official" qualitative analysis packages like NVivo or Atlas.ti. If you're not familiar with Pivot Tables, do a quick Internet search to learn the basics, as they can be especially helpful for facilitating analysis during qualitative coding.

For **transcribing any audio** you might collect as part of *research through evocative play*, I recommend using f4.[5]

For **statistics**, I recommend checking out Professor Koji Yatani's statistics webpage,[6] which is widely used among HCI PhD students to learn the basics of or brush up on specific statistical methods; also draw from other Internet resources as needed. If you're already familiar with Excel and not comfortable with statistical coding, know that you can do a lot in Excel, alone. If you've never learned any formulas in Excel, then you should definitely do some Internet searching to learn the basics. For example, I've found combining Excel's INDEX and MATCH formulas[7] to be invaluable for data cleaning and tidying in Excel. As you advance to more complex statistics, then a statistical package will become necessary. I recommend using the R programming language via RStudio,[8] which is free and open-source, and allows users to choose between scripting options and a graphical user interface (aka non-coding options). Once again, Internet searches are your friends; there are tons of resources online for conducting statistics in R.

---

[4]https://qualtrics.com/
[5]https://www.audiotranskription.de/english/f4
[6]http://yatani.jp/teaching/doku.php?id=hcistats:start
[7]https://exceljet.net/index-and-match
[8]https://www.rstudio.com/

7.4.2 Readings and Skill-building in Play

As I've mentioned elsewhere in this document, I am not a play researcher or designer. Though like most humans, I've played some games in my lifetime, I do not even consider myself much of a game-player. I did, however, have to do quite a bit of reading in play research and theory to catch up, and also benefited from exposure to games, play, and play research through the Human-Computer Interaction Institute and other related departments at Carnegie Mellon (e.g., the Entertainment Technology Center). Do, at the very least, read my chapter on background/related work (Chapter 3) to gain an overview of relevant play readings. For example, while I recommend reading Bernard Suits' Grasshopper, as well as a number of readings related to the magic circle (including Salen and Zimmerman's work), it's more important to understand the general concepts of the magic circle and the lusory attitude than to dive fully into the vast amount of literature on the magic circle, so you may be able to get by (at least to start) with the overview I provide. (That said, recall that I am not a play researcher or theorist, so I simultaneously encourage you to question my presentation of salient play concepts, and to return to original sources whenever possible).

Here, I want to highlight the three play readings that were most instrumental and inspiration to me while developing *research through evocative play*.

Miguel Sicart. *Play matters*. MIT Press, 2014.   This is a short book, and I highly suggest reading it in full. Sicart's perspectives on revealing context through play were by far the most inspirational to me in developing *research through evocative play*.

Ian Bogost. *Play anything: The pleasure of limits, the uses of boredom, and the secret of games*. Basic Books, 2016.   Although I did not formally cite this book in my background chapter, some of the concepts it presents greatly affected my research approach. I'll confess, however, that it was actually not the book itself, but rather, a podcast interview about the book[9] that most inspired my thinking. The concepts of applying play to situations that are not normally viewed as "fun" or "playful," and finding pleasure in the limits of daily experiences helped me to reconsider how I approach research. I do not always agree with all of Bogost's arguments in the book (e.g., we have much different views on "Ironoia"), but the spirit of of embracing constraints and re-framing situations as playful inspired me to consider research through a playful lens. I'd recommend first taking the shortcut and listening to a podcast about the book, and then reading the full book as you see fit.

---

[9]https://www.artofmanliness.com/articles/podcast-247-pleasure-limits-uses-boredom-antidote-excessive-irony/

Nicole Lazzaro. **"Why we play games: Four keys to more emotion without story".** **In: (2004).** Lazzaro's Four Keys to Fun (published under slightly different names and forms; the version I cite is most easily accessible from the Internet) is another must-read if you do not come from a games/play background. Lazzaro's work both forced me to think more deeply about what the term "fun" really means, as well as to consider my own goals in developing *research through evocative play*. It was through reading Lazzaro that I came to the realization that my goal and measure of success in *research through evocative play* was actually *not* fun, as is often the case in games and play, more generally.

7.4.3 Other Readings and Skill-building

Three other favorite readings that I have found myself returning to time and time again are:

Whitney Phillips. *This is why we can't have nice things: Mapping the relationship between online trolling and mainstream culture*. MIT Press, 2015.

Whitney Phillips and Ryan M Milner. *The ambivalent internet: Mischief, oddity, and antagonism online*. John Wiley & Sons, 2018.

Adrienne L Massanari. *Participatory culture, community, and play*. 2015.

These particular readings may not be immediately relevant to the power dynamics and contexts you choose to study using *research through evocative play*, but I urge you to at least read book summaries of these three (even the back covers will do, to give you an initial sense), and return to them if and when they become relevant. All three books cover elements of and concepts related to Internet culture that I've found to be relevant to a wide array of online spaces, behaviors, and communities.

When conducting *research through evocative play*, you'll also need to read up on the specific context you are studying (and power dynamics therein). I'll cover this in more detail in describing the "initial exploration stage" of conducting a *research through evocative play* project.

You'll also find that there will be additional variegated skills relevant to your project and form of declared play that you'll need to pick up. For example, when I created the monster vending machine discussed in the Preface (Chapter 1, I needed to learn basic rules about applying acrylic paint and protecting the paint from the

elements (as I carted the machine about campus during the initial stage of design). For YouMercials and Turker Tales, I had to learn more about creating Google Chrome extensions and backend options (Firebase/Firestore), and the platform designs of MTurk and YouTube. I had some prior knowledge of and experience with HTML/CSS/Javascript and basic database management, but was forced to stretch the limits of that knowledge and expand my skill base in the process of conducting *research through evocative play*.

On the one hand, take the path of least resistance; consider what is the easiest, most efficient way to implement your *research through evocative play* concept. For example, if you have no background or experience with programming, you certainly don't need to learn programming to implement a *research through evocative play* study design (even if you are studying a digital domain). A survey, paper-based or slideshow-based prototype might be just as effective.

On the other hand, if you are not a programmer but you *want* to learn programming, then I encourage you to use *research through evocative play* as an excuse to expand your skill set. Recalling that an intended advantage of *research through evocative play* includes supporting the well-being of the researcher and engaging the researcher in play, it's your job to enjoy the process of *research through evocative play*. This may not be absolutely necessary for a *research through evocative play* to be successful, at least in the short term, but I do believe it to be critical to your long-term resilience as a researcher and maintaining your commitment to projects. Remember to enjoy the process of *research through evocative play*. If for you, that means pursuing a slightly less efficient path of learning to achieve your goals, then I encourage you to enjoy the more circuitous route.

7.5 Designing and Implementing *Research through Evocative Play*

The following is a rough guideline for implementing a *research through evocative play* from start to finish. My main cautionary note before you proceed: Remember that *research through evocative play* is a learning-through-doing research approach. It's okay if you don't do everything perfectly. It's expected that you will realize after the fact myriad ways in which you could have done things better; this is true of virtually any research method, in fact. Do not let fear of "screwing up" keep you from moving on to the next stage of a *research through evocative play* project.

7.5.1 Initial Exploration Stage

Any *research through evocative play* should first begin with an initial exploration stage, which involves a combination of brainstorming, creative exercises, background reading, and reflection.

First, identify your context, and the power dynamics you are interested in exploring. Identify the stakeholders in your context. Activities such as the Luma Institute handbook's [103] stakeholder mapping exercise can be especially instrumental here. To the best of your own ability, articulate (and document) your own biases related to the context. Be prepared to continue reflecting on, re-evaluating, and re-assessing your biases as you continue down the *research through evocative play* path. If possible, allow space and time for both individual and group-based reflection. For example, in my initial exploration phase for YouMercials, I conducted the very individual-based daily measures project, but then later also worked in a more collaborative context through the "Roast Me, Part 2" course. The individual component helped me reflect on and articulate my own perspectives on and orientations towards online advertising, and the shared, collaborative component highlighted to me ways in which my perspectives may be particularly biased.

Next, do some background reading and research related to the context you are studying. What has already been written or done in the context? This background reading and research can range from informal to formal; you can draw on both academic and non-academic sources, and can include both general, informal observation of the context as well as more formal research methodologies. For example, with YouMercials, at one point I did some very informal observation and data collection on Twitter, seeking mentions of YouTube advertising to get a sense of how people discuss YouTube advertisements, and how that compares to my own assumptions. Background research may also involve conducting original research to inform your approach. With Turker Tales, for example, I conducted speed dating sessions— a method from research through design— to probe on Turkers' perceptions towards anonymity and privacy, power dynamics between requesters and workers and among workers, and notions of play during crowd work. Be sure to document your processes, emergent findings, and changing perceptions as you go through the intitial exploration phase.

7.5.2 Defining your Research Focus and Specific Tools of Analysis

Following the initial exploration stage, you should define your research focus and specific tools of analysis.

When I say "research focus" I am not necessarily referring to, say, formal hypotheses. A research focus can be articulated broadly, such as for YouMercials: "understanding power dynamics involving YouTube viewers in relation to YouTube advertisers, the YouTube corporation, and other viewers of YouTube advertisements," though you'll then want to break that down for yourself by listing key questions you hope to be able to answer through your study. In fact, I would urge users of *research through evocative play not* to rely too heavily on hypothesis-driven research, but instead, to directly acknowledge their own biases when articulating their research focus. In my case for YouMercials, that would mean acknowledging a bias against YouTube advertisers, and in favor of surfacing user empowerment and resistance towards YouTube advertisers. Again, be prepared to re-evaluate these biases throughout the *research through evocative play* process, and strive to enter the research space as open and flexible as possible while simultaneously ready to accept that you are never as open as you think you are.

That said, if you are trained as an experimental researcher, or experimental research is more heavily valued in your domain or department, then it is perfectly fine to conduct this acknowledgement of biases through the lens of hypotheses or very specific research questions. In fact, I myself took a partially experimental design approach to YouMercials. Just keep in mind that the findings of *research through evocative play* should ultimately be emergent, and that your orientation towards results should be outcome-ambiguous. Consider: will your results be interesting and informative regardless of whether or not the null hypotheses you lay out are negated? If the answer is "no," then you need to reconsider your approach.

Next, determine how you will measure the results of *research through evocative play*. It might seem a little premature to think about analyses methods when you haven't yet designed the declared play intervention; I'll circle back to this point in a moment. The methods you choose are those which can best answer the questions you want to address. As I've discussed, there are tradeoffs, but no clear "rights" or "wrongs" when it comes to choosing analysis methods. If you're starting to get confused about the terminology here (but wait, I thought *research through evocative play* is itself a method, so why the heck are we talking about choosing analysis methods?), I'll direct you back to parallels to research through design. *Research through evocative play* is defined by its use of declared play in a context to reveal power dynamics in that

context as studied through emergent reactions, behaviors and attitudes displayed in relation to that declared play. *Research through evocative play* is also characterized by its view of the researcher-designer as an integral player and active participant in the study, whose behaviors and attitudes should be studied in tandem with those of the other study participants. Analyzing the results of a *research through evocative play* study, however, can draw on a variety of research methods of analysis, including both qualitative and quantitative approaches.

In choosing analysis methods, I first want to emphasize starting with what you know. Yes, you'll probably find that you'll want to incorporate at least some basic statistics eventually when presenting your results, but if you've never done any statistics before and don't generally think in terms of quantitative analyses, then don't start with statistics! I did not discuss the use of playtesting as a method because although I've had some exposure to playtesting, it's not my area of expertise. But if *your* area of expertise is playtesting, and a bunch of ideas about how to study the space using playtesting methods pops into your head, then by all mean, use playtesting approaches to analyzing the data.

Second, I want to emphasize the importance of including at least *some* form of qualitative analyses— even if you tend to be more of a quantitative researcher— in analyzing the results of your *research through evocative play* study. Of course, I am biased here, as my own research (both inside and outside of *research through evocative play* while a HCI researcher) has been more qualitative-based. I'm an especially big fan of semi-structured interviews, and if I could return to the projects I presented in this document, I would include even more qualitative analysis (e.g., post-interviews with participants in Turker Tales and YouMercials). Because you are seeking to understand nuances of power dynamics, the richness of qualitative inquiry will be especially useful to *research through evocative play*. I want to avoid setting strict guidelines on the implementation of *research through evocative play* here, but I would venture to say that the incorporation of at least *some* qualitative analysis techniques are almost always necessary for *research through evocative play*.

Now, I'd like to circle back to the question of why you would choose your analysis methods at this point in the process and not *after* you've designed the play declarations and cues you'll be using. Certainly, once you actually design the *research through evocative play* declared play you will deploy, you might end up altering the specific analysis methods you choose to employ, and that is perfectly fine and expected. I argue that you should nonetheless articulate what tools of analysis you will use from the get-go because I've found in my own experience that designing the declared play first, without fully thinking through how implementing that declared play will answer your research questions or how you will determine those answers (e.g., what methods of analysis you will use), will render it quite dificult

to later convince both yourself and others that what you did was truly "research." The numerous "non-research" projects that I discussed in the Preface (Chapter 1) could have been re-framed as *research through evocative play* had I reflected more deeply on and articulated both my research questions and the analysis methods I (tentatively) planned to employ *before* designing and implementing the play cues.

7.5.3 Choosing your Play Cues

My own use of *research through evocative play* centered around the use of four specific play cues: (1) direct play declarations, (2) harsh humor, (3) perspective-taking (e.g., via short *who, why, what* scenarios), and (4) the creation of shared artifacts (in anonymous or semi-anonymous spaces). Thus, the advice I can give in relation to types of cues to use, and responses to expect from those cues, is largely limited to these four types of cues. However, you need not limit yourself to these four types of play cues in your own *research through evocative play* projects. I also encourage you to do some creative brainstorming first rather than choosing your play cues from a purely logical and strategic standpoint. Again, I recommend trying out different activites from the Innovating for People handbook [103] (especially the section on Concept Ideation).

Direct Play Declarations

One of the biggest surprises for me of conducting *research through evocative play* was how readily participants engaged in play. In retrospect, this is perhaps not altogether surprising. Humans are playful animals, after all [13]. With Turker Tales, for example, I was skeptical that participants would engage playfully. TUrkers are on the Amazon Mechanical Turk platform to earn money, and tasks are often low-paying, such that maximizing efficiency is often a central concern for crowd workers. With Turker Tales, I declared a play activity that doesn't require playful engagement to receive payment, and that could be more profitable if one whipped out non-playful responses. I thus was uncertain as to whether participants would actually engage playfully.

However, in Turker Tales and again in YouMercials, I found participants to be readily accepting of viewing an activity as play simply because I had declared it as play. Thus, researchers in *research through evocative play* should expect that by declaring play in a space, many participants will be inclined to play. In this way, when anticipating responses to direct declarations of play, plan to focus on the specific *manners* of play that emerge. E.g., does the play engagement come with a side of snark that criticizes the declared play itself, or some aspects of the context in some way? Is the

play relevant and in conversation with the given context, or does it go off-topic, and in what ways? Instances of *disengagement* from play should be seen as especially noteworthy.

In directly declaring play in both YouMercials and Turker Tales, I also found that participants either made assumptions about my role and affiliation in the design, or responded according to how I primed them to think of me as the researcher-designer. Thus, recalling that *research through evocative play* is as much a process of self-exploration and reflection as it is an outward-facing research approach, researchers using *research through evocative play* should expect the participants' responses to direct play declarations to be embedded with their views and opinions on the researcher's role in the process and placement in the power dynamics being studied.

Harsh Humor

My own observation of and study of harsh humor, both as part of a direct play declaration (RoastMe and YouMercials) and as an emergent behavior (Turker Tales), led me both to see the value of harsh humor in certain contexts (e.g., allowing for types of criticisms and freedom of expression that might otherwise be discouraged or silenced) as well as its dangers (specifically, the possibility of psychological injury to those on the receiving end of harsh humor). Integrating harsh humor as a play cue can lead to interesting insights about power dynamics, but because of its potential for harm, should be approached very cautiously.

One question that should be central in considering the use of harsh humor is: are there opportunities for reciprocity and/or retaliation (aka turnabout)? The context in which you are operating, as well as the way in which you declare play, will have a large influence on the answer to this question. For example, in RoastMe, participants had the option to engage as both a roaster *and* a roastee. Although this system was far from perfect in terms of mitigating harm in the community, it did at least allow for some degree of reciprocity. Many users operated exclusively as roasters, but there was at least an implied expectation that eventually, you would also submit yourself to a roasting, and experience the same level of harsh humor that you had levied at others. This helped participants on RoastMe see the harsh humor they engaged in as more equitable, though it didn't necessarily reduce harm to individual participants being roasted. It was also assumed in the community of declared play that those submitting themselves to be roasted had *consented* to the process of roasting, and were both expecting of and accepting of being the target of harsh humor. As I've already discussed, this does not guarantee the safety of using harsh humor, but as

participants attested, it did help both roasters and roastees view harsh humor as generally acceptable within the context.

In Turker Tales, the emergent harsh humor often criticized general concepts and actions, such as low paying HITs posted by requesters, and tended not to target individuals, as the scenarios participants created were in response to work tasks. In this way, I was not immediately faced with the ethical concerns of encouraging harsh humor in the crowd work context. However, in YouMercials, where a subset of participants were directly instructed to use harsh humor, the presence of actors in the commercials who had *not* consented or expected to be roasted opened up ethical concerns. Moreover, by priming participants to use roasting humor, YouMercials creators also ended up roasting potential users of products in ways that could be injurious to other viewers watching the advertisements that might identify with the group or individual being ridiculed. Thus, in considering the direct encouragement of harsh humor as a play cue, researchers should reflect on and articulate what groups or individuals could potentially be at risk for harm given the context and the other components of the play declaration. You should weigh the ethical pros and cons of encouraging humor in the declared play context.

Perspective-Taking

Where perspective-taking emerged unprompted on RoastMe, in Turker Tales and YouMercials, I directly instructed participants to engage in perspective-taking activities. In investigating power dynamics, you will necessarily be studying relationships between individuals in a context, and as such, perspective-taking may be particularly useful either as a direct instruction for users to employ as part of the play declarations, or as a point of analysis for emergent behaviors. Of course, I am biased in that my advisor has expertise in perspective-taking and experience-taking that has influenced my own perspectives (see [90, 121] for further reading on perspective- and experience-taking).

In RoastMe, which again, I view as a precursor to *research through evocative play*, perspective-taking emerged naturally. Participants viewed the roaster-roastee relationship as reciprocal, given implicit expectations that roasters should also eventually post as roastees, and be willing to "handle the heat" in addition to dishing it out. In accordance with this implied rule of reciprocity, roasters often considered what it would be like to be in the shoes of the roastee. Thus, researchers of *research through evocative play* should keep in mind that play declarations that encourage reciprocity may naturally enable more perspective-taking, and they should be on the lookout for examples and manifestations of perspective-taking.

With Turker Tales and YouMercials, by contrast, I directly cued for perspective-taking. However, it would have been interesting to instead (or in addition) design declared play that was directly reciprocal in nature, and allowed perspective-taking strategies and mechanisms to emerge naturally. At the same time, note that the perspective-taking strategies employed in RoastMe were not self-evident; I only learned about participants' use of such strategies through the semi-structured interviews. Thus, if you plan to study emergent rather than instructed perspective-taking strategies, you will need to plan for follow-up inquiry, such as semi-structured interviews or a survey, to understand how and when such strategeies were enacted. You cannot expect to find evidence of perspective-taking strategies from observing and analyzing engagement in the declared play, alone.

If however, you directly cue for perspective-taking, you can observe how such perspective-cues are interpreted and responded to directly in your design. In cuing directly for perspective-taking, you'll need to think about audience; *who* can you engage in perspective-taking? For example, in Turker Tales, it might have been interesting and informative to have requesters take the perspective of Turkers, but both the platform design (through which it is easier to engage Turkers in tasks) and my own choice to focus on understanding the perspectives of Turkers, not requesters, meant that I designed perspective-taking for Turkers. As such, I knew that my study would be unable to provide insights into requesters' perspectives towards Turkers. You should be able to list and acknowledge such limitations prior to implementing the study.

When cuing for perspective-taking, you'll subsequently need to decide *whose* perspective you plan to instruct participants to adopt or consider. For Turker Tales, I used two layers of perspective-taking, asking participants to first put themselves in the shoes of another Turker, and then to instruct other Turkers to put themselves in the shoes of someone else. By so doing, I was able to investigate Turkers' relationships to one another, as well as to study what kinds of perspectives Turkers would choose to create for others, and how that might reveal aspects of themselves and their perspectives in relation to the context. Although I omitted requesters as "players" when designing the cues for Turker Tales, attitudes towards requesters still emerged and thus factored into my analysis. Participants expressed opinions of requesters and the platform at large both in conversation with other peers (as they were prompted), as well as in conversation with me, the researcher. For example, those in the pilot that critiqued low-paying receipt transcription HITs often intimated that I myself was part of the requester problem by asking them to complete the perspective-taking exercises in the context of receipt transcription. Thus, you should know that in instructing for specific types of perspective-taking, it is likely that participants will incorporate and reflect on power dynamics related to other stakeholders in the space through the perspective-taking exercise (that is, even though they are not

directly taking those stakeholders' perspectives or designing an activity for other stakeholders).

Note that I might have instead chosen to ask Turker Tales participants to take the perspective of the requesters. In YouMercials, I built off the methods from Turker Tales to include components that did require taking on the perspective of the YouTube advertisers, especially for those in the *help advertisers* condition. Here, I want to highlight that any time you are asking participants to take the perspective of not just another individual in their same peer group (e.g., other YouTube advertisement viewers), but of individuals with differential power in the space (e.g., advertisers), you are essentially asking them to empathize with those individuals. In so doing, you may elicit responses from participants that critically reflect on the stakeholders or directly resist the perspective-taking activity. However, as I saw hints of in YouMercials, you may also engage participants in an empathy-building activity, and accordingly, you should weigh the pros, cons, and ethical implications of the potential for encouraging empathy for other stakeholders in the space in light of the power dynamics operating in the context.

Regardless of whether you choose to employ perspective-taking as an instructed cue, or whether you plan to observe and analyze any emerging perspective-taking strategies, be sure to document and articulate the implications and possible responses of participants, and to reflect on both the ethics and the potential insights that will be gained before implementing your design. Directly laying out for yourself the specifics of how (e.g., emergent through reciprocity versus directly instructed) for whom (e.g., what stakeholders in the context) and whose perspective (thinking especially about status and power differentials, e.g. Turkers versus requesters) you will ask participants to take will help you clearly implement your design and study the results moving forward.

Social Component: Creation of Shared Artifacts

In my own research, I focused on digital domains where anonymity or semi-anonymity governed what (limited) communication options were available to participants via the platform, and where peer communication was either limited (e.g., in RoastMe, there was an implicit rule that roastees should not "talk back") or essentially non-existent within the context (e.g., crowd worker and user-viewer participants lack means of platform-supported peer communication on Amazon Mechanical Turk and YouTube advertising, respectively). This impelled me to explore play design options that would allow for sharing among individuals that cannot normally share or communicate freely with one another through the platform.

Given that (a) play is already often a social experience (see Lazzaro's discussion of the social experience key to fun [135]), and that (b) you want to study power dynamics between people using *research through evocative play,* I highly recommend including play cues that allow for social components in order to draw out and observe those interpersonal relationships. If you decide to do so via the creation of shared artifacts, as I did, note that the power dynamics implied by certain kinds of artifact creation may not always be self-evident. For example, with Turker Tales, I was able to observe directly aspects of the self that Turkers revealed, but could not always as easily discern Turkers' views of *other* Turkers from those artifacts. In retrospect, my analysis of Turker Tales could have benefited from follow-up interview work (my go-to would be semi-structured interviews, but other methods like contextual inquiry while participants are engaging in the application, or a post-survey, like I used in YouMercials, could also be appropriate).

In both Turker Tales and YouMercials, I used ratings to gain feedback on participants' reactions to one another's ratings, but did not allow for direct interaction around created artifacts. I chose this approach in large part to protect participants from potentially negative reactions to the artifacts they had created. However, this also created a sort of black box nature to the artifact creation. Participants created artifacts, but then never knew how others viewed their artifacts, nor could viewers communicate directly with the creators. In this way, I placed myself in a power position, controlling and limiting communications, just as the platforms themselves had controlled and limited communications among participants.

In RoastMe, interactions tended to be more reciprocal in that most roastees had also participated as a roaster at some point (though the inverse was not true for the large number of exclusive roasters in the community). Interactions around posts were also more direct. For example, when someone posts a photo of themselves to be roasted, they receive direct feedback from roasters. Likewise, roasters' comments are upvoted or downvoted by other RoastMe members; some comments that are considered either particularly funny or particularly lame will receive comments back from other roasters either applauding, riffing on, degrading, or lampooning the roast. In this way, the social interactions in RoastMe were more reciprocal and interactive than sharing engendered by the artifact creation play cues I used in Turker Tales and YouMercials. Still, with RoastMe it was the interviews rather than the interactions themselves that revealed the nuances of power behind those interactions (for example, the use of perspective-taking by roasters, or the benefits of "thick skin" gained by roastees).

Thus, researchers should be aware of how the manner of sharing artifacts (or other forms of social interaction) may affect the nature of those interactions. For example, should artifacts be shared publicly or privately, anonymously or identifiably,

reciprocally or as one-off creations "into the void?" The same statement could be applied to otherwise enabling or encouraging social interactions through the declared play. Regardless of the specific nature of either the artifact creation or other form of social exchange you choose to engender through your *research through evocative play* design, I would advise you to employ methods such as interviews or surveys in order to gain greater insights into the power dynamics revealed by those interactions. It is likely that the interactions alone will not tell you the full story.

Other Considerations

Incentivized play. In designing and implementing *research through evocative play*, especially if you are working in a formal research environment in academia or industry, you'll need to decide on whether and how to pay participants. Personally, I'm of the mindset of paying if you have any funds available. If your career or schooling might benefit from this study, it seems only ethical to also remunerate the participants that are critical for your study's existence. Moreover, if you are assuming that the declared play you design will cause ambivalent or even negative reactions in participants, then you'll need some form of external incentive to persuade participants to consider the declared play at all. If instead you position a declared play intervention as fully voluntary and uncompensated, only those interested in engaging in play in the context will participate, which will skew your results.

However, keep in mind that incentivizing play will almost certainly affect how participants engage in the declared play. For example, in Turker Tales and YouMercials, even though I didn't specify that participants *had* to engage playfully in order to receive compensation, participants may have assumed so, anyway. In RoastMe, participants already belonged to the context, and thereby their engagement in the declared play was not affected by external incentives. But because I *did* pay participants for the semi-structured interviews, I can conjecture that participants' responses might also have been affected by the pay. They might have aimed to tell me what I wanted to hear since I was paying them for their time.

Yet the same could be true of an uncompensated *research through evocative play* project. If participants view me as the researcher, then they may still be subject to demand characteristics and the Hawthorne effect [151] regardless of whether I pay them or not; participants may still be incentivized to change their behaviors or responses in order to help me in my research. I therefore think it's important to acknowledge that the power of the researcher is always going to be an active component of any *research through evocative play* project (and really, any research project). Although I did not always do this in my own work, I recommend asking

follow-up questions post-engagement regarding participants' views towards the researcher and/or participants' perceptions of the research intentions or goals.

Other ways to address the payment issue are to use differential incentives and/or deception. With Turker Tales, I used a differential incentive in that participants received the highest payment for simply downloading the extension (regardless of whether they engaged or not), and then provided only very small payments for additional tasks (e.g., artifact creation and ratings). My intent was to reduce feelings of pressure to engage in the play as part of the study, while still providing some payment in order to abide by both Amazon Mechanical Turk rules and expectations, as well as ethical research guidelines for implementing studies on Amazon Mechanical Turk.

Deception is another option, provided you are able to receive approval for a deceptive study design from your university's or company's Institutional Reivew Board. For example, you could consider informing participants that they won't be compensated for their participation, and then at the end of of the study, pay all participants. Similarly, you might tell participants that a small subset will be paid at random, and then ultimately pay everyone who participants. You might even choose to deceive participants about your own role in the process, perhaps posing as a non-researcher in some way (and only revealing your true identity after the study has come to a close).

Effects of modeled behavior.   Another item to consider is the visibility of behaviors in the declared play intervention, and how participants might influence each others' behaviors. In RoastMe, for example, more experienced members of the community model proper roasting behaviors; newcomers to the subreddit can review and learn from historical examples of acceptable roasting behavior on RoastMe. Longitudinal *research through evocative play* study designs, which I did *not* explore in my own research, might thereby become more similar to online communities over time in that newcomers to the declared play can see concrete examples of expected engagement in the space.

I would assume that most *research through evocative play* studies, however, will be more short-term in nature, similar to Turker Tales and YouMercials. Participants will thereby need very explicit rules of action, and may not be very attuned to the rules when first entering the space. As the researcher-designer, you should also consider that even in short-term scenarios, participants may be affected by the behaviors of other participants. For example, in YouMercials, I found that participants in the *take back power* condition, especially, did not always behave in adherence with their assigned rules of play. I had included steps to reinforce participants' study

conditions (by repeating the instructions multiple times) and had also checked that they had paid attention and absorbed the information (by asking them to summarize the instructions in their own words). However, these actions might still not have been enough to adequately establish the rules of declared play for participants. Participants may also have been affected by the YouMercials they viewed, created by participants who had been assigned different rules, and changed their behavior accordingly. In such designs, it will be especially important to engage in follow-up analysis (such as a post-survey) to understand whether deviations from the rules of declared play stem from personal decisions to resist or subvert, or are the results of modeled behavior effects.

Another option is to model expected or possible behaviors for participants in order to mimic or simulate more longitudinal situations. In Turker Tales, for example, I seeded the in-the-wild deployment with stories collected from the pilot. However, you'll want to consider the possible impacts of behavior modeling in your specific design. In most situations, you'll want to avoid biasing participants to engage in specific ways, as this might alter your findings. Often, using behaviors collected from formal or informal pilot testing is a reasonable middle ground in that participants will not necessarily follow the rules perfectly, and thus open up a range of possible engagement options for participants. This is a strategy I also used in the YouMercials controlled experiment, wherein participants were asked to rate YouMercials that had been created by pilot participants.

Other forms of play cues.   As a reminder, I focused on the four play cues of direct declarations of play, harsh humor, perspective-taking, and the creation of shared artifacts, but you need not limit yourself to these cues in your own *research through evocative play* study designs. In my designs, I also used primarily text-based forms of shared artifacts, only branching out a bit with YouMercials to include the creation of audio artifacts. Play artifacts or modes of interaction could instead include drawing or visual elements, for example. Other play mechanisms could include self-control (à la Simon Says), or cooperative play, sharing or creation. I cannot yet speak with any authority about the types of responses to expect from declared play using mechanisms I did not attempt in my own research, but your own research could help blaze the way.

7.5.4 Pilot Testing and Deploying your Design

You're now at the point where you've done a lot of preliminary exploration, reading, reflection, and planning; it's time to dive in. You'll first want to conduct at least some type of pilot testing, (formal or informal), if you haven't already done so during the

initial exploration phase. Refer back to my reading recommendations and tips about skill-building for HCI earlier in this chapter.

Next, you're ready for full deployment. See my notes on "Other Readings and Skill-building," above, and remember to choose methods of deployment that (1) get the job done and (2) get the job done in a way that's fun for you.

Whichever deployment method you choose, just be sure to acknowledge (and document) the pros and cons. For example, for YouMercials, I could have used a *research through evocative play* approach wherein I had people create YouMercials in-person (perhaps in collaboration with other participants). This would have also pushed on boundaries of social anonymity. Such a design is not necessarily better or worse, but would likely have yielded different insights. I might have gained richer insights into participants' creative ideas for altering YouTube advertisements, but I would have been doing so in a context that is quite distinct from the original YouTube advertising context, and thus potentially less insightful in terms of revealing participants' attitudes towards YouTube advertising. Assuming that in this scenario, I (the researcher) would be present in the room, an in-person design would also alter researcher-participant dynamics, perhaps encouraging participants to respond in ways that they felt I *wanted* them to behave. In making your choice of deployment method then, consider and document how your decision might affect participant responses.

Also note that you can and should seek collaborators with specific skill sets if you feel a certain deployment approach would be especially useful to your research. For example, for Turker Tales, it would have been much harder (though not impossible) to even consider implementing an in-person deployment approach, given the distributed (and less widely utilized) nature of the platform. Doing so would have also made it quite difficult to simultaneously consider the implications of actual play on Turker Tales, which I conjectured would probably need to be distributed and online. So in this sense, I viewed a coded, platform-integrated deployment approach as quite necessary for my *research through evocative play* study design.

Now if you don't already have those coding skills, and have the energy, time and desire to work towards building them, I highly encourage you to do so, in keeping with my previous advice to make the process fun for you (even if that means it ends up being somewhat more circuitous). When learning new skills, always allow for *way* more time than you think you need, be proud of small accomplishments, and don't get too caught up in the "I'm not a ..." (coder, play designer, qualitative researcher, etc.) mindset. At the same time, when you get stuck (you most likely will), and/or if you do not have the time/energy/desire to throw yourself into learning a new set of skills, don't be afraid to seek the support and collaboration of

others, especially if "doing it yourself" is holding you back from making progress. I'll admit that as someone who is very stubborn, and very easily frustrated with my perceived lack of ability and progress when learning new skills, this is not advice that I always follow well myself. At the very least, talk with others. In a human-computer interaction department, I often found it encouraging to talk to those who *do* have computer science backgrounds about their own struggles (past and present) with programming projects, as it would remind me that even experts in a domain hit roadblocks, encounter challenges, and feel they are not progressing as quickly as they should be.

7.5.5 Analyzing, Reflecting on, and Synthesizing Results

After deploying your *research through evocative play* design, remember that you already laid out your plan of analysis; you should return to that plan. But at this point, you'll want to re-assess the analysis methods you had chosen. Remember that you should be incorporating at least some qualitative analyses of the emergent responses to the declared play intervention you designed and implemented. Beyond that, the world (of qualitative and quantitative data analysis methods) is your oyster. You will likely find in analyzing emergent findings that there are additional areas you want to explore, and this may require some additional reading and skill-building. Refer back to my advice on quantitative and qualitative research methods for resources on appropriate analyses as needed. Moreover, similar to my advice on the design and deployment, don't be afraid to reach out to others for input or help on analyses methods with which you're not familiar. Recall also that for most analyses methods— especially quantitative analysis methods— simple search engine queries will produce a plentiful array of free and useful explanatory resources, examples, and usage guides.

7.6 Final Note: Forks Encouraged

For those of you with some familiarity with GitHub (an online space for storing, sharing, and collaborating on coding projects), you know that "forking" a project means copying someone else's project to use it as a starting point for your own ideas, modifying, redirecting and adding to the original project as you see fit. In presenting these guidelines for conducting *research through evocative play* in your own work, I encourage "forks." The rules of *research through evocative play* are not set in stone, and you may come up with novel and creative ways to improve upon the methodology, or use the concept of *research through evocative play* as a jumping-off point for other methodologies or research approaches. I am excited to see how others choose to adapt and extend *research through evocative play*.

If you have questions or comments about the use of *research through evocative play*, or want to share your own experiences with using or modifying *research through evocative play*, please feel free to reach out to me. You can best reach me via email, amkasunic@gmail.com.

In the following final chapter, I will conclude the presentation and discussion of my work with *research through evocative play* by summarizing its contributions as well as its limitations, and discussing interesting ways in which I hope to see *research through evocative play* extended in future research.

Conclusion

8

I have introduced *research through evocative play* as an empirical research methodology in which the researcher-observer draws a circle of play in a context, declaring play, and instructing participants to enter into that context. In so doing, the researcher-observer positions themself as an active player and participant, while the circle declared is uncertain and ambiguous in that not all participants will view the circle as magic, or enter in with a lusory (playful) attitude. By introducing more ambiguity in a space, and asking participants to engage in the circle of declared play without first confirming that participants *want* to engage in play and view the declared circle as magical, *research through evocative play* invites participants to reflect on and respond to the power complexities embedded in the context. The play engendered by *research through evocative play* is thereby "evocative" in that it evokes reflection and engagement that reveals and deepens our understanding of participants, attitudes, and power dynamics in a context. *Research through evocative play* can be especially useful for eliciting criticism of a context's power dynamics, as well as ambivalent or conflicted attitudes held by context inhabitants.

Research through evocative play is also unique in the way it places a lens of analysis on the researcher, and engages the researcher in play. In viewing the researcher-observer-"declared play designer" as holding an active position of power, *research through evocative play* encourages critical reflection on the research process and the researcher's own biases and perspectives. Coupled with the engagement of the researcher in play, *research through evocative play* is also a method that concerns itself with the visibility and well-being of the researcher.

8.1 Summarizing RoastMe, Turker Tales, and YouMercials

I began my discussion of my work in *research through evocative play* with the study of r/RoastMe. With RoastMe, I entered into the declared play setting as a researcher-observer rather than a researcher-designer, and studied how an existing declared play circle operated, with special attention to where uncertainties, harms, challenges and rewards lay. I found that, on the one hand, the form of declared play meets desires unfulfilled in other online spaces by encouraging humorously-intended, harsh

criticism of online photographic self-presentations. For many RoastMe participants, the declarations and rules of play the subreddit puts forth *do* allow them to take on lusory attitudes. By casting photographic self-presentation and the use of harsh language as play, RoastMe also led both roaster and roastee participants to reflect on contexts of self-presentation more broadly. For example, roastee participants relished the opportunity to take themselves less seriously and acknowledge and have fun with their perceived flaws rather than trying to hide them.

My study of RoastMe was a precursor to the *research through evocative play* approach I developed and implemented with Turker Tales and YouMercials in that the subreddit was a pre-existing space of play that participants at least initially and ostensibly entered into voluntarily. That is, in RoastMe, a circle of play was first circumscribed by the subreddit's rules, and then behaviors of photographic self-presentation and harsh humor were transported to the space, thereby taking behaviors that aren't always seen as playful, and naming them as play. The circle of play on RoastMe is also part of a larger culture of irreverent play that permeates Reddit [148]. In this way, we might presume a more equal balance of power. For example, roastees, by consciously choosing to receive criticism rather than acting as unwitting targets, could be seen as gaining power through the experience, e.g. by thickening their own skin and not taking themselves (or their perceived flaws) seriously. The presence of reciprocity in the subreddit (the expectation that participants should consider engaging as both roasters and roastees) also mitigates the power differentials between roasters and roastees, reducing the presence of traditional target-perpetrator dynamics in interactions such as cyberbullying. However, as my interviews of RoastMe participants showed, the power dynamics in RoastMe are more complex than this surface level understanding would suggest.

As emerged from my qualitative investigation, the power dynamics on RoastMe were actually not so clear-cut. For example, many of the interview participants were hesitant to view the play circle of RoastMe as fully magical; a number of interview participants expressed uncertainty, concern, and guilt about their participation. They expressed doubts about unequal power dynamics and lack of protection for vulnerable individuals. The quantitative data analysis I conducted showed there was validity to their concern, with female-appearing roastees, as well roastees experiencing mental health struggles (as evidenced by their posts on other mental-health related subreddits) being disproportionately targeted in RoastMe. Participants worried about redditors who may participate without fully understanding or consenting to the rules of RoastMe play, and questioned whether their own engagement in RoastMe was thereby potentially unethical. In RoastMe, declaring harsh humor and online self-presentation as play may evoke playful attitudes, but it also opens up questions of power and consent. RoastMe participants presumably entered the subreddit's circle with more of a lusory attitude than we would expect from a *research through*

evocative play declared play scenario, in that participants for the most part willingly joined the playful subreddit and consented to its rules of declared play, rather than having play declared and overlaid atop a non-play context. Nonetheless, the very presence of a play declaration still drew ambiguities of power to the fore. The revelatory power of declared play I observed in RoastMe led me to further explore ways to use declared play in research.

My work with RoastMe inspired and influenced the *research through evocative play* approach I next took to Turker Tales. On the one hand, Turker Tales demonstrated possibilities for supporting Turkers in novel ways through play, as many participants appeared to adopt playful attitudes in engaging with Turker Tales, and the system also allowed workers to share aspects of themselves and their experiences in an online platform where communication between peers is normally suppressed. The work suggests alternative designs for MTurk that could support play, especially paid play untied from work efficiency outcomes, and calls into question the roles and responsibilities of both researchers and requesters in terms of how they support and interact with crowd workers.

At the same time, with Turker Tales, and drawing on my prior research with RoastMe, I began to formally introduce *research through evocative play* as a design and play research method and concept. In contrast to RoastMe, in which the play context was created from the start by community originators, and new members voluntarily entered into a community of play, with Turker Tales and later YouMercials, I introduced top-down declared play as the researcher. Whereas in my study of RoastMe, I was an observer (or "researcher-observer") of an existing community, starting with Turker Tales, I shifted my role to one of researcher-observer-designer. In so doing, I too, joined the context and became an active player, myself. In this way, any critiques, critical reflection, perspective-taking, and power considerations that emerge from *research through evocative play* studies necessarily include me in the mix, and place a lens of analysis on me, the researcher, as much as it does on the other, non-researcher participants.

Critical responses and reflections on power imbalances in Amazon Mechanical Turk emerged from Turker Tales, especially when participants viewed the researcher-designer as aligned with unethical MTurk requester practices such as low pay. I therefore posited that from a *research through evocative play* perspective, encouraging views of paternalistic power, in which the researcher-designer is viewed as aligned with oppressive forces within a context, could be especially beneficial to revealing participants' negative views of a context. Such a design could be especially illustrative in a community like MTurk, where such communication might otherwise be prohibited or suppressed. However, Turker Tales also calls into question the implications of declaring play in a context where power dynamics are unequal.

Playing with a capitalist system, after all, could be a way of contributing to that capitalist system even when the form of play is attempting to criticize the system.

Finally, with YouMercials, I incorporated aspects of both RoastMe and Turker Tales in my *research through evocative play* study design. Like RoastMe, YouTube is a space of play and entertainment with aspects of participatory culture. Like Amazon Mechanical Turk, user participation in YouTube (and especially the more specific context of YouTube advertising) contributes to a capitalist system, and power imbalances between user-producers (recall again the concept of "produsage" [23]) are ambiguous. Play in all three study designs was declared directly by the designer or research ("This is play, now"), included elements of perspective-taking, and centered around communication and content-sharing with anonymous or pseudonymous peers. Inspired by the declared play on RoastMe, in the YouMercials study, I assigned a subset of participants to the *roast* condition, prompting them to use a roasting, or critical/ridiculing style of humor, in creating YouMercials. As in Turker Tales, participants also engaged in perspective-taking directly by creating similar identity imagination exercises for other participants to view and engage in, and like Turker Tales, YouMercials featured the creation and sharing of artifacts with other participants.

With YouMercials, I found that aligning the researcher-designer with (potentially) oppressive powers in a *research through evocative play* study design may not always be effective in eliciting criticism of or resistance to a context. I observed a lack of harsh critique of YouTube and YouTube advertising among those that, using a paternalistic paradigm inspired by the results of Turker Tales, had been explicitly primed to consider the potential benefits of YouMercials to power-holding stakeholders (advertisers). What's more, there was an the absence of such criticism even when participants had been explicitly primed to "take power back" from YouTube and its advertisers through YouMercials. For the most part, participants adhered to YouMercials' rules of play, and when I did see evidence of deviation from those rules, it was often in the opposite direction than I had anticipated, with *take back power* participants displaying more support for brands, products and advertising stakeholders than expected.

YouMercials thereby also placed an analytic lens on me as the researcher, and the assumptions and biases against YouTube advertising through which I approached the context. I conceptualized YouMercials as a way to elicit critique of YouTube advertising power dynamics, using strategies such as direct empowerment messages (*take back power*), paternalistic priming (*help advertisers*), and harsh humor (*roast*). However, participants' engagement in the controlled experiment and Google Chrome extension revealed a more nuanced reality, with users viewing advertisers as stakeholders in the context bringing value to the table (e.g., allowing for "free" viewing of

YouTube and support of content creators, and alerting users to potentially relevant products and services) rather than purely oppressive, power-wielding forces. Similar to Turker Tales, my work with YouMercials once again led me to further question the ethics of play in the context of *research through evocative play* and beyond.

8.2 *Research through Evocative Play*: What it Is and Is Not

I now turn to laying out a brief definition of *research through evocative play* based on my own work in this approach. I consider this definition nascent, however, because it is my hope that other researchers extend, complicate, shift, and even re-define the concept and application of *research through evocative play* in future work.

Research through Evocative Play Is:

1. A research method designed to produce empirical contributions.

2. A research and design method that uses play declarations to reveal and understand nuances of a context and the complexities of power dynamics therein.

3. A research method heavily influenced by both design research (especially research through design, critical design, and ludic design) as well as play theory and research.

4. A research method well-suited to tackle "wicked problems," where there is not a clear computational solution. Particularly, *research through evocative play* is intended for studying power dynamics in a context.

5. A research method that encourages critical reflection on the context being studied, concepts of play considered in that context, and the study design process, itself.

6. A reflective practice that acknowledges and draws into the analysis the role of the researcher-designer in the study design context.

7. A methodology that, in requiring the researcher to approach their work playfully, concerns itself with the well-being of the researcher.

Research through Evocative Play Is *Not*:

1. The promotion of a specific play design or system.

2. Advocacy for the use of play or games in a given context.

3. A research method that will result in clear and unambiguous conclusions about a context.

4. A prescriptive methodology with a fixed set of rules or approaches.

8.3 Effectiveness of Specific Cues for Play and Engagement

As discussed, the goal of *research through evocative play* is not necessarily to engage participants in play— a "success" in *research through evocative play* is often *not* measured by enjoyment of the play or assumption of lusory attitudes, as the implications of playing in and with a given context under *research through evocative play* are recognized as complex, ambiguous, and problematic by nature. Nonetheless, in the course of my research, I learned lessons about how specific forms of play may be effective (or ineffective) in engaging participants in play. These insights can be of use to researchers and designers of play, more broadly, as well as for scholars using *research through evocative play* that seek to either encourage or discourage engagement through specific forms of play declarations in a given context. Below, I summarize lessons learned for each of the four play cues I focused on during my development of the *research through evocative play* approach.

- **Direct play declarations.** Simply declaring a space as play— essentially, stating "This is play, now"— was surprisingly effective in persuading participants to accept activities as play. This did not necessarily mean that participants subsequently viewed the circle as fully magical, or adopted wholly lusory attitudes. But in the same way I may dislike a certain game and question its value as a form of play while still viewing that game as a form of play, participants did not question the entire activity as un-playful. Now, other cues for play also co-existed and no doubt interacted with these direct play declarations, but it appears that direct play declarations, in conjunction with other design cues for play, can be an effective piece of the puzzle to prompt engagement in playfulness within contexts or sets of behaviors that are not normally framed as playful.

- **Harsh humor.** Harsh or "roasting" humor *can* be effective for engaging participants, but given the specificity and particular flavor of this form of humor,

might be difficult for some participants to latch onto. In YouMercials, for example, a number of participants correctly summarized the spirit behind roasting, but had difficulty actually implementing this form of humor, and expressed self-consciousness and discomfort regarding the use of roasting humor in the post-survey. Depending on the aims and approach of the researcher or designer, using roasting humor to draw out these discomforts could be useful, but roasting humor should not be seen as an "easy" way to draw participants in or incite ludic curiosity. Further, the potential harm that harsh humor can generate should give researchers and designers pause; the direct prompting of harsh humor should be very carefully considered, and only included if it is deemed that the benefits of its inclusion outweigh the potential for harm.

- **Perspective-taking through short *who, why, what* scenarios**. In RoastMe, perspective-taking was not an explicit rule of play, but roaster participants often placed themselves in the shoes of the roastees to consider what kind of humorous remarks they might want to hear, or in what ways the roastee is like themselves. Inspired by the emergent use of perspective-taking in RoastMe, I directly incorporated design cues for perspective-taking in both Turker Tales and YouMercials. For both, I used a version of this *who, why, what* approach, and found it to be a relatively short, simple, and unintimidating way for participants to engage in the system, while still providing rich content for analysis that revealed their relationships to the context and peers within the context. For example, Turker Tales participants created a range of creative scenarios despite the more obvious incentive for them to "spam" the system to maximize earnings, and YouMercials participants created far more "imagine yourself as" YouMercials than they did audio-dubbed YouMercials, with participants measuring higher in enjoyment for creating and viewing "imagine yourself as" YouMercials than audio-dubbed YouMercials. However, there may be context-dependent limitations to the success of such a play form. For example, in YouMercials, many of the "imagine yourself as" YouMercials that took the rules of "play" seriously by earnestly supporting the products or brands simultaneously did so with a lack of creativity, not attempting to surprise, enchant or otherwise engage readers of the text in interesting, novel or creative ways.

- **Creation of shared artifacts in anonymous or semi-anonymous spaces.** With RoastMe, roasts (comments) allow for social interactions in a pseudonymous online space, often taking on collaborative elements as roasters build or riff off each others' roasts. The process of roasting is reciprocal in that there are implicit norms that you should be able to both "dish it out and take it," that is, engage in the subreddit as both a roaster and a roastee. Drawing from the public nature of comment sharing in RoastMe under a shared task structure

(the directive to roast participants), I incorporated the creation and sharing of (primarily text-based) artifacts in my *research through evocative play* study designs of Turker Tales and YouMercials as an analogous means to enable social interactions in the two contexts. In Turker Tales and YouMercials, the creation and sharing of artifacts allowed for maintaining anonymity (especially important to the Amazon Mechanical Turk participants), while still sharing connections with and inciting curiosity about other participants. All three examples of declared play— the pre-existing context of RoastMe, as well as the *research through evocative play* designs of Turker Tales and YouMercials— thereby fit in with Lazzaro's [135] "social experience key" of play enjoyment.

8.4 Limitations and Dangers of *Research through Evocative Play*

In my conclusion of Chapter 6, I spoke of *research through evocative play* as relatively safe and protected, but not without its dangers. In all of the studies I presented, we saw examples in which both the ramifications of playing in a context, as well as the direct implications of play declarations and enacted play and engagement within the study itself, might be harmful, unethical, or marginalizing. In RoastMe, where the declared play *is,* in fact, in the wild (not confined to a study design), both the qualitative and quantitative analyses I conducted suggested that in keeping with participants' concerns, members of the RoastMe community *are* in fact harmed in the process of participating in RoastMe. Declaring behaviors that might normally be considered harmful or undesirable as play does *not* fully remove that original level of harm.

Likewise, taking aspects of (for example) the RoastMe paradigm and confining them to a research study may not necessarily have as widespread or long term an effect, but could be injurious, all the same. In Turker Tales and YouMercials, I spoke of the ethical ambiguity of playing in and with a context with oppressive elements— that danger remains even if the intent is to critique or rail against those elements. *Research through evocative play* can be a way to observe, critically consider, and reflect upon such dynamics, but offers a shield in that the declared play is not designed to fully operate in the wild over a longer time period, or to be viewed as a direct model for play within a context. Nonetheless, *research through evocative play* is by no means commensurate to "safe play," and can result in undesired consequences. For example, in my initial exploration and design stage of YouMercials, I considered how engaging in YouMercials might draw criticsms of power structures to the fore, but instead found that some participants actually viewed advertisers more positively after the

study, or self-reported paying *more* heed to advertisements after participating in the study.

Similarly, by instructing a subset of YouMercials participants to employ "roasting" humor when creating their YouMercials, I inadvertently brought harm to individuals (such as actors) and other groups of individuals (such as potential users of a product that share certain characteristics) that the participants chose to judge and roast. Where RoastMe at least requires that roastee participants hold the "RoastMe" sign indicating their consent and understanding of the rules of play, the design of YouMercials essentially allowed for and even encouraged the use of harsh humor towards people who had *not* consented to be ridiculed.

Moreover, nothing is stopping other individuals, groups, or institutions/corporations outside the research study from being inspired by and ultimately co-opting a *research through evocative play* concept to meet ends that conflict with the values tested and explored through the original design. For example, ideas of play from Turker Tales could be used to engage crowd workers in play and increase general work satisfaction while maintaining low wages and power imbalances that characterize Amazon Mechanical Turk.

Research is often motivated by "good intentions," but this does not necessarily mean the resulting research and its implications are ethically sound. By the same token, critically reflecting on potential negative impacts of one's work, and acknowledging ways in which the work could be improved, or the work raises ethical issues, should not discredit the value of a research contribution. Many research fields, including human-computer interaction, subscribe to this line of thought, at least in theory. For example, a "limitations" section much like this one that acknowledges imperfections or areas for concern, is standard and expected in academic papers. In practice, however, researchers may feel pressure to downplay their work's ethical issues or possibilities for improvement.

In laying out *research through evocative play*, I thus very purposefully position critical reflection on the role of the researcher and the research process as a core component of the methodology. *Research through evocative play* researchers should enter into their work as consciously as possible, exploring the problem space and beginning to come to terms with their own biases in the initial exploration phase. For example, researchers should consider the negative externalities their work could engender, and protect against such externalities as much as possible within the scope of their research. However, doing so does not ensure that the final *research through evocative play* design you deploy will be free of ethical or other issues; there is no such thing as flawless research. The presence of biases and ethical concerns should not be seen as a *research through evocative play* failure, but as an expected component of

the research process that the researcher should be ready to analyze, reflect on, and discuss.

8.5 Future Work for *Research through Evocative Play*

There are a number of ways in which my own work in *research through evocative play* could be improved upon and extended. A research methodology should not be created in isolation, but should instead incorporate diverse ideas and perspectives. For this reason, I view *research through evocative play* as a budding research approach, and very much hope that other researchers choose to use, adapt, and extend the methodology.

As I discussed in Chapter 7, *research through evocative play* does not prescribe the use of a specific analytic approach in assessing the results of a *research through evocative play* study design, though I did discuss that in general, qualitative approaches should at least play *some* role in the plan of analysis. Following, I would encourage future research to place even greater emphasis on the importance of qualitative inquiry as part of *research through evocative play* than I did in my own work. In designing and implementing a *research through evocative play* prototype or system, greater attention should be paid to extensive qualitative interviews after the implementation of a system or prototype featuring play declarations. For example with RoastMe, I was able to gain insights into participants' relationship with the declared play of RoastMe through in-depth, semi-structured interviews.

The use of deception within *research through evocative play* is another area ripe for exploration. In YouMercials, I used priming to encourage certain possible inter-pretations of the system's (and the researcher-designer's) intents, without formally introducing deception into the study design. The use of deception, which has been successfully used in other play-related research, e.g. [119, 67, 118], could be especially useful in further investigating the potentials of evoking criticism and reflection on a context through paternalistic design in *research through evocative play*, whereby participants may purposefully resist against the rules of declared play if they believe that the researcher-designers intentions and affiliations conflict with their own values.

Future work could also further study the effects of manipulating the level of anonymity and presence within contexts that assume anonymity in creating shared artifacts. In my work with both Turker Tales and YouMercials, I was hesitant to break anonymity given both crowd workers' professed values for anonymity (as revealed

by my speed-dating sessions), and research study conventions that seek to protect the anonymity of participants. However, pushing the boundaries of anonymity in anonymous contexts through *research through evocative play*, especially in cases where the anonymous nature of communication in a context may hamper or otherwise limit communications or reinforce unequal power dynamics, could prove especially compelling and insightful as a research approach.

Lastly, I hope that in time and in collaboration with distributed colleagues from diverse perspectives, *research through evocative play* can not only explore critiques of contexts within those contexts, thereby limiting themselves to operating within the confines of existing power structures, but can also dream up and then create new hybrid spaces of research in which to explore the ways in which play declarations can reveal context and power dynamics therein in nuanced, problematic and conflicting ways. Moreover, I currently view *research through evocative play* as best applied to short-term projects, but the approach could also be iteratively adapted and applied to longitudinal research.

In my research, I introduce *research through evocative play* as a methods contribution with relevance to human-computer interaction research, design research, and play theory and research. As a side product of developing this methodology, my work also contributes directly to play theory and design by providing lessons learned about the effectiveness of different forms of play cues in engaging participants.

The full potentials of *research through evocative play* as a methodology have yet to be realized. I urge fellow researchers and practitioners in human-computer interaction and play, scholars of power dynamics in other disciplines, and other curious humans seeking to better understand the complexities of power dynamics in a context to draw from and build upon the *research through evocative play* methodology I have laid out here. I look forward to collectively discovering what the future of *research through evocative play* holds.

Bibliography

[1] Gregory A Aarons and Angelina C Sawitzky. "Organizational climate partially mediates the effect of culture on work attitudes and staff turnover in mental health services". In: *Administration and Policy in Mental Health and Mental Health Services Research* 33.3 (2006), p. 289 (cit. on p. 63).

[2] Joey Adams. *Here's to the Friars: The Heart of Show Business*. Crown Publishers, 1976 (cit. on p. 40).

[3] Turki Alelyani, Ke Mao, and Ye Yang. "Context-Centric Pricing: Early Pricing Models for Software Crowdsourcing Tasks". In: *Proceedings of the 13th International Conference on Predictive Models and Data Analytics in Software Engineering*. PROMISE. Toronto, Canada: ACM, 2017, pp. 63–72 (cit. on p. 63).

[4] Joseph V Anderson. "Creativity and play: A systematic approach to managing innovation". In: *Business Horizons* 37.2 (1994), pp. 80–86 (cit. on p. 149).

[5] Mark Andrejevic. "Exploiting YouTube: Contradictions of user-generated labor". In: *The YouTube Reader* 413 (2009) (cit. on p. 36).

[6] Anonymous. *Dirtbag requesters on Amazon Mechanical Turk*. 2013. URL: `http:// scumbagrequester.blogspot.com/2013/08/what-is-jon-brelig-and-oscar- smith.html` (cit. on p. 75).

[7] Jane Arthurs, Sophia Drakopoulou, and Alessandro Gandini. *Researching YouTube*. 2018 (cit. on p. 36).

[8] Mitja D Back, Juliane M Stopfer, Simine Vazire, et al. "Facebook profiles reflect actual personality, not self-idealization". In: *Psychological Science* 21.3 (2010), pp. 372–374 (cit. on p. 33).

[9] Shaowen Bardzell, Jeffrey Bardzell, Jodi Forlizzi, John Zimmerman, and John Antanitis. "Critical design and critical theory: the challenge of designing for provocation". In: *Proceedings of the Designing Interactive Systems Conference*. ACM. 2012, pp. 288–297 (cit. on pp. 22, 26).

[10] John A Bargh, Katelyn YA McKenna, and Grainne M Fitzsimons. "Can you see the real me? Activation and expression of the "true self" on the Internet". In: *Journal of social issues* 58.1 (2002), pp. 33–48 (cit. on p. 33).

[11] P Barreira and V Bolotnyy. *Graduate Student Mental Health: Lessons from American Economics Departments*. `https://scholar.harvard.edu/bolotnyy/publications/ graduate-student-mental-health-lessons-american-economics-departments` (cit. on p. 148).

[12] Roland Barthes. *Camera lucida: Reflections on photography*. Macmillan, 1981 (cit. on p. 34).

[13] Patrick Bateson. "The role of play in the evolution of great apes and humans". In: *The nature of play: Great apes and humans* 1324 (2005) (cit. on p. 167).

[14] Benjamin B Bederson and Alexander J Quinn. "Web workers unite! addressing challenges of online laborers". In: *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2011, pp. 97–106 (cit. on p. 35).

[15] Janine Berg. "Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers". In: *Comp. Lab. L. & Pol'y J.* 37 (2015), p. 543 (cit. on p. 34).

[16] Gregory S Berns, Kristina Blaine, Michael J Prietula, and Brandon E Pye. "Short-and long-term effects of a novel on connectivity in the brain". In: *Brain connectivity* 3.6 (2013), pp. 590–600 (cit. on p. 28).

[17] Christian Blauvelt. *The Comedy Central Roast of Donald Trump*. 2011 (cit. on p. 40).

[18] David M Blei, Andrew Y Ng, and Michael I Jordan. "Latent dirichlet allocation". In: *Journal of machine Learning research* 3.Jan (2003), pp. 993–1022 (cit. on p. 71).

[19] Laurie Block and Johanna Leseho. ""Listen and I tell you something": Storytelling and social action in the healing of the oppressed". In: *British Journal of Guidance & Counselling* 33.2 (2005), pp. 175–184 (cit. on p. 65).

[20] Ian Bogost. *Play anything: The pleasure of limits, the uses of boredom, and the secret of games*. Basic Books, 2016 (cit. on pp. 150, 161).

[21] Allum Bokhari. *#RedditRevolt Rages as Redditors Condemn 'Chairman Pao's' Crackdown*. 2015 (cit. on p. 56).

[22] Erin Brady, Meredith Ringel Morris, and Jeffrey P Bigham. "Gauging Receptiveness to Social Microvolunteering". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 1055–1064 (cit. on p. 35).

[23] Axel Bruns. "Produsage: Towards a broader framework for user-led content creation". In: *Creativity and Cognition: Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*. ACM. 2007 (cit. on pp. 36, 182).

[24] Richard Buchanan. "Wicked problems in design thinking". In: *Design issues* 8.2 (1992), pp. 5–21 (cit. on pp. 21, 146).

[25] Madeleine Bunting. *Willing slaves: How the overwork culture is ruling our lives*. Harper-Collins UK, 2011 (cit. on p. 63).

[26] Jean Burgess. "'All your chocolate rain are belong to us?'Viral Video, YouTube and the dynamics of participatory culture". In: *Video vortex reader: Responses to YouTube*. Institute of Network Cultures, 2008, pp. 101–109 (cit. on p. 35).

[27] Ronald J Burke. "Workaholism in organizations: the role of organizational values". In: *Personnel review* 30.6 (2001), pp. 637–645 (cit. on p. 63).

[28] Aleksejs Busarovs et al. "Ethical Aspects of Crowdsourcing, or is it a Modern Form of Exploitation". In: *International Journal of Economics & Business Administration (IJEBA)* 1.1 (2013), pp. 3–14 (cit. on p. 34).

[29] Gordon Calleja. "Erasing the magic circle". In: *The philosophy of computer games*. Springer, 2012, pp. 77–91 (cit. on p. 24).

[30] Chris Callison-Burch. "Crowd-workers: Aggregating information across turkers to help them find higher paying work". In: *Second AAAI Conference on Human Computation and Crowdsourcing*. 2014 (cit. on pp. 35, 63).

[31] A Carlick and Francis C Biley. "Thoughts on the therapeutic use of narrative in the promotion of coping in cancer care". In: *European Journal of Cancer Care* 13.4 (2004), pp. 308–317 (cit. on p. 65).

[32] Fabio Celli, Elia Bruni, and Bruno Lepri. "Automatic personality and interaction style recognition from Facebook profile pictures". In: *Proceedings of the 22nd ACM international conference on Multimedia*. ACM. 2014, pp. 1101–1104 (cit. on p. 33).

[33] Eshwar Chandrasekharan, Umashanthi Pavalanathan, Anirudh Srinivasan, et al. "You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech". In: *Proceedings of the ACM on Human-Computer Interaction* 1.CSCW (2017), p. 31 (cit. on pp. 32, 56).

[34] Jane Harper Chelf, Amy MB Deshler, Shauna Hillman, and Ramon Durazo-Arvizu. "Storytelling: A strategy for living and coping with cancer". In: *Cancer Nursing* 23.1 (2000), pp. 1–5 (cit. on p. 65).

[35] Chun-Wei Chiang, Anna Kasunic, and Saiph Savage. "Crowd Coach: Peer Coaching for Crowd Workers' Skill Growth". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (2018), p. 37 (cit. on pp. 35, 63).

[36] Judeth Oden Choi, Jodi Forlizzi, Michael Christel, et al. "Playtesting with a Purpose". In: *Proceedings of the 2016 annual symposium on computer-human interaction in play*. ACM. 2016, pp. 254–265 (cit. on p. 145).

[37] Giovanni Luca Ciampaglia and Dario Taraborelli. "MoodBar: Increasing new user retention in Wikipedia through lightweight socialization". In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM. 2015, pp. 734–742 (cit. on p. 35).

[38] Pauline Rose Clance and Suzanne Ament Imes. "The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention." In: *Psychotherapy: Theory, Research & Practice* 15.3 (1978), p. 241 (cit. on p. 149).

[39] Pauline Rose Clance and Maureen Ann OToole. "The imposter phenomenon: An internal barrier to empowerment and achievement". In: *Women & Therapy* 6.3 (1987), pp. 51–64 (cit. on p. 149).

[40] Stanford Crowd Research Collective. "Daemo: a self-governed crowdsourcing marketplace". In: *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM. 2015, pp. 101–102 (cit. on p. 35).

[41] *Commenting to Learn: Evidence of Language and Intercultural Learning in Comments on YouTube Videos*. https://scholarspace.manoa.hawaii.edu/bitstream/10125/44435/19_03_benson.pdf. (Accessed on 04/25/2019) (cit. on p. 35).

[42] Patsy Cooper. *When Stories Come to School: Telling, Writing, and Performing Stories in the Early Childhood Classroom*. ERIC, 1993 (cit. on p. 64).

[43] Marinka Copier. *Beyond the magic circle: A network perspective on role-play in online games*. Utrecht University, 2007 (cit. on p. 24).

[44] Marinka Copier. "Challenging the magic circle". In: *Digital Material* (2009), p. 159 (cit. on p. 24).

[45] John W Creswell and Cheryl N Poth. *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications, 2017 (cit. on p. 160).

[46] Arthur J Cropley. "Creativity and mental health in everyday life". In: *Creativity Research Journal* 3.3 (1990), pp. 167–178 (cit. on p. 149).

[47] Peng Dai, Jeffrey M Rzeszotarski, Praveen Paritosh, and Ed H Chi. "And now for something completely different: Improving crowdsourcing workflows with micro-diversions". In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM. 2015, pp. 628–638 (cit. on pp. 35, 63).

[48] Nicholas Sheep Dalton. "Neurodiversity & HCI". In: *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2013, pp. 2295–2304 (cit. on p. 13).

[49] Kerstin Dautenhahn and Chrystopher Nehaniv. "Artificial life and natural stories". In: *International Symposium on Artificial Life and Robotics*. Citeseer. 1998, pp. 435–439 (cit. on p. 64).

[50] Scott Davidoff, Min Kyung Lee, Anind K Dey, and John Zimmerman. "Rapidly exploring application design through speed dating". In: *International Conference on Ubiquitous Computing*. Springer. 2007, pp. 429–446 (cit. on p. 64).

[51] Nicola Dell and Neha Kumar. "The Ins and Outs of HCI for Development". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: ACM, 2016, pp. 2220–2232 (cit. on p. 13).

[52] Xuefei Nancy Deng and KD Joshi. "Is Crowdsourcing a Source of Worker Empowerment or Exploitation? Understanding Crowd Workers' Perceptions of Crowdsourcing Career". In: (2013) (cit. on p. 34).

[53] Jennifer DeWinter, Carly A Kocurek, and Randall Nichols. "Taylorism 2.0: Gamification, scientific management and the capitalist appropriation of play". In: *Journal of Gaming & Virtual Worlds* 6.2 (2014), pp. 109–127 (cit. on pp. 27, 79).

[54] Maja Djikic, Keith Oatley, Sara Zoeterman, and Jordan B Peterson. "On being moved by art: How reading fiction transforms the self". In: *Creativity Research Journal* 21.1 (2009), pp. 24–29 (cit. on p. 64).

[55] Mira Dontcheva, Robert R Morris, Joel R Brandt, and Elizabeth M Gerber. "Combining crowdsourcing and learning to improve engagement and performance". In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2014, pp. 3379–3388 (cit. on p. 35).

[56] Steven Dow, Anand Kulkarni, Scott Klemmer, and Björn Hartmann. "Shepherding the crowd yields better work". In: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM. 2012, pp. 1013–1022 (cit. on pp. 35, 63).

[57] Anthony Dunne. "Hertzian tales: Electronic products, aesthetic experience, and critical design". In: (2008) (cit. on p. 22).

[58] Carsten Eickhoff, Christopher G Harris, Arjen P de Vries, and Padmini Srinivasan. "Quality through flow and immersion: gamifying crowdsourced relevance assessments". In: *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*. ACM. 2012, pp. 871–880 (cit. on p. 35).

[59] Nicole Ellison, Rebecca Heino, and Jennifer Gibbs. "Managing impressions online: Self-presentation processes in the online dating environment". In: *Journal of Computer-Mediated Communication* 11.2 (2006), pp. 415–441 (cit. on p. 33).

[60] Teresa M Evans, Lindsay Bira, Jazmin Beltran Gastelum, L Todd Weiss, and Nathan L Vanderford. "Evidence for a mental health crisis in graduate education". In: *Nature biotechnology* 36.3 (2018), p. 282 (cit. on p. 148).

[61] Alexandra Eveleigh, Charlene Jennett, Stuart Lynn, and Anna L. Cox. ""I Want to Be a Captain! I Want to Be a Captain!": Gamification in the Old Weather Citizen Science Project". In: *Proceedings of the First International Conference on Gameful Design, Research, and Applications*. Gamification '13. Toronto, Ontario, Canada: ACM, 2013, pp. 79–82 (cit. on p. 35).

[62] ExpressExpense. *ExpressExpense's Massive Receipt Database*. 2016 (cit. on p. 68).

[63] Rosta Farzan, Joan M. DiMicco, David R. Millen, et al. "Results from Deploying a Participation Incentive Mechanism Within the Enterprise". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. Florence, Italy: ACM, 2008, pp. 563–572 (cit. on p. 35).

[64] Anne-Laure Fayard and John Weeks. "Photocopiers and water-coolers: The affordances of informal interaction". In: *Organization studies* 28.5 (2007), pp. 605–634 (cit. on p. 61).

[65] Casey Fiesler and Nicholas Proferes. ""Participant" Perceptions of Twitter Research Ethics". In: *Social Media+ Society* 4.1 (2018), p. 2056305118763366 (cit. on p. 13).

[66] Mary Flanagan and Peter Carini. "How games can help us access and understand cultural artifacts". In: *American Archivist* 75.2 (2012), pp. 514–537 (cit. on p. 79).

[67] Mary Flanagan and Geoff Kaufman. "Shifting Implicit Biases with Games Using Psychology". In: *Diversifying Barbie and Mortal Kombat* (2016), p. 219 (cit. on p. 188).

[68] Paul Florin and Abraham Wandersman. "An introduction to citizen participation, voluntary organizations, and community development: Insights for empowerment through research". In: *American Journal of Community Psychology* 18.1 (1990), pp. 41–54 (cit. on p. 35).

[69] National Science Foundation. *Doctorate recipients from US universities: 2014*. 2015 (cit. on p. 148).

[70] Sarah Fox, Jill Dimond, Lilly Irani, et al. "Social Justice and Design: Power and oppression in collaborative systems". In: *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM. 2017, pp. 117–122 (cit. on p. 13).

[71] Christian Fuchs. "Class and Exploitation on the Internet". In: *Digital labor. The Internet as playground and factory* (2013), pp. 211–224 (cit. on p. 36).

[72] Christian Fuchs. *Internet and society: Social theory in the information age*. Routledge, 2007 (cit. on p. 36).

[73] Mathias Fuchs. "Gamification as twenty-first-century ideology". In: *Journal of Gaming & Virtual Worlds* 6.2 (2014), pp. 143–157 (cit. on p. 36).

[74] Mathias Fuchs. "Subversive Gamification". In: *Playful Disruption of Digital Media*. Springer, 2018, pp. 181–191 (cit. on p. 36).

[75] Shira Gabriel and Ariana F Young. "Becoming a vampire without being bitten: The narrative collective-assimilation hypothesis". In: *Psychological Science* 22.8 (2011), pp. 990–994 (cit. on p. 28).

[76] Ujwal Gadiraju, Alessandro Checco, Neha Gupta, and Gianluca Demartini. "Modus operandi of crowd workers: The invisible role of microtask work environments". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.3 (2017), p. 49 (cit. on pp. 34, 63).

[77] William Gaver. "What should we expect from research through design?" In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM. 2012, pp. 937–946 (cit. on p. 21).

[78] William Gaver, Andy Boucher, Sarah Pennington, and Brendan Walker. "Cultural probes and the value of uncertainty". In: *Interactions-Funology* 11.5 (2004), pp. 53–56 (cit. on pp. 22, 23, 26, 158).

[79] William W Gaver, John Bowers, Andrew Boucher, et al. "The drift table: designing for ludic engagement". In: *CHI'04 extended abstracts on Human factors in computing systems*. ACM. 2004, pp. 885–900 (cit. on p. 22).

[80] Carolin Gerlitz and Anne Helmond. "The like economy: Social buttons and the data-intensive web". In: *New media & society* 15.8 (2013), pp. 1348–1365 (cit. on p. 36).

[81] Tony Gillam. "Creativity and mental health care". In: *Mental health practice* 16.9 (2013) (cit. on p. 149).

[82] Elena L. Glassman, Aaron Lin, Carrie J. Cai, and Robert C. Miller. "Learnersourcing Personalized Hints". In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. San Francisco, California, USA: ACM, 2016, pp. 1626–1636 (cit. on p. 35).

[83] Erving Goffman. "The presentation of self in everyday life". In: *American Journal of Sociology* 55 (1949), pp. 6–7 (cit. on p. 33).

[84] Elizabeth Goodman, Mike Kuniavsky, and Andrea Moed. *Observing the user experience: A practitioner's guide to user research*. Elsevier, 2012 (cit. on p. 159).

[85] Google. *BigQuery - Analytics Data Warehouse | BigQuery | Google Cloud*. 2018 (cit. on p. 71).

[86] Google. *Cloud Firestore | Firebase*. 2018 (cit. on p. 68).

[87] Alvin W Gouldner. "Alienation from Hegel to Marx". In: *The Two Marxisms*. Springer, 1980, pp. 177–198 (cit. on p. 34).

[88] Isabela Granic, Adam Lobel, and Rutger CME Engels. "The benefits of playing video games." In: *American psychologist* 69.1 (2014), p. 66 (cit. on p. 149).

[89] Mary L. Gray, Siddharth Suri, Syed Shoaib Ali, and Deepti Kulkarni. "The Crowd is a Collaborative Network". In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. San Francisco, California, USA: ACM, 2016, pp. 134–147 (cit. on pp. 35, 63).

[90] Melanie C Green, Timothy C Brock, and Geoff F Kaufman. "Understanding media enjoyment: The role of transportation into narrative worlds". In: *Communication Theory* 14.4 (2004), pp. 311–327 (cit. on p. 169).

[91] Anthony G Greenwald and Linda Hamilton Krieger. "Implicit bias: Scientific foundations". In: *California Law Review* 94.4 (2006), pp. 945–967 (cit. on p. 159).

[92] Ido Guy, Anat Hashavit, and Yaniv Corem. "Games for Crowds: A crowdsourcing game platform for the enterprise". In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM. 2015, pp. 1860–1871 (cit. on p. 79).

[93] David Hankerson, Andrea R Marshall, Jennifer Booker, et al. "Does technology have race?" In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM. 2016, pp. 473–486 (cit. on p. 13).

[94] Michael Hanley and Michael Becker. "Cell phone usage and advertising acceptance among college students: A four-year analysis." In: *International Journal of Mobile Marketing* 3.1 (2008) (cit. on p. 132).

[95] Karin Hansson, Tanja Aitamurto, and Thomas Ludwig. "From alienation to relation: Examining the modes of production in crowdsourcing". In: *Proceedings of 15th European Conference on Computer-Supported Cooperative Work-Exploratory Papers*. European Society for Socially Embedded Technologies (EUSSET). 2017 (cit. on p. 34).

[96] Kotaro Hara, Abigail Adams, Kristy Milland, et al. "A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 449 (cit. on pp. 34, 63).

[97] Christine Hauser. *Reddit Bans Nazi Groups and Others in Crackdown on Violent Content*. 2017 (cit. on p. 56).

[98] *History of YouTube - Wikipedia*. https://en.wikipedia.org/wiki/History_of_YouTube. (Accessed on 04/25/2019) (cit. on p. 84).

[99] Thomas J. Holt, Kristie R. Blevins, and Natasha Burkert. "Considering the Pedophile Subculture Online". In: *Sexual Abuse: A Journal of Research and Treatment* 22.1 (2010), 3–24 (cit. on p. 32).

[100] Bell Hooks. "The oppositional gaze: Black female spectators". In: *The feminism and visual cultural reader* (2003), pp. 94–105 (cit. on p. 90).

[101] Skye Hudson. *Free Speech vs. Harassment: Why Did Reddit Ban Five Subreddits?* 2015 (cit. on p. 56).

[102] Holly M Hutchins. "Outing the imposter: A study exploring imposter phenomenon among higher education faculty". In: *New Horizons in Adult Education and Human Resource Development* 27.2 (2015), pp. 3–12 (cit. on p. 149).

[103] Luma Institute. *Innovating for people: Handbook of human-centered design methods*. LUMA Institute, LLC, 2012 (cit. on pp. 157, 159, 164, 167).

[104] Lilly Irani. "The cultural work of microwork". In: *New Media & Society* 17.5 (2015), pp. 720–739 (cit. on pp. 34, 35, 78).

[105] Lilly C Irani and M Silberman. "Stories we tell about labor: Turkopticon and the trouble with design". In: *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM. 2016, pp. 4573–4586 (cit. on pp. 34, 35, 63).

[106] Lilly C. Irani and M. Six Silberman. "Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, 2013, pp. 611–620 (cit. on pp. 35, 65, 78).

[107] Henry Jenkins and Mark Deuze. *Convergence culture*. 2008 (cit. on p. 35).

[108] Huizinga Johan. *Homo ludens: A study of the play element in culture*. 1950 (cit. on pp. 9, 23, 27).

[109] Carl Gustav Jung. *Man and his symbols*. Laurel, 1964 (cit. on p. 64).

[110] Jesper Juul. "The magic circle and the puzzle piece". In: (2008) (cit. on p. 24).

[111] Bob Kanefsky, Nadine G Barlow, and Virginia C Gulick. "Can distributed volunteers accomplish massive data analysis tasks". In: *Lunar and Planetary Science* 1 (2001) (cit. on p. 35).

[112] Katherine A Karl, Joy V Peluchette, and Lynn Harland. "Is fun for everyone? Personality differences in healthcare providers' attitudes toward fun". In: *Journal of health and human services administration* (2007), pp. 409–447 (cit. on p. 63).

[113] Anna Kasunic and Geoff Kaufman. "" At Least the Pizzas You Make Are Hot": Norms, Values, and Abrasive Humor on the Subreddit r/RoastMe." In: *ICWSM*. 2018, pp. 161–170 (cit. on p. 39).

[114] Anna Kasunic and Geoff Kaufman. "Be Me or Be Mii?: A Study of Self-Presentation and Interaction in the Miitomo Mobile Application". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM. 2017, pp. 1973–1977 (cit. on p. 4).

[115] Anna Kasunic and Geoff Kaufman. "How We Create and Embody the Other: Implications for Diversity in Character-Centric Games and Media". In: *Meaningful Play*. 2018 (cit. on p. 13).

[116] Anna Kasunic, Chun Wei Chiang, Geoff Kauffman, and Saiph Savage. "Crowd Work on a CV? Understanding How AMT Fits into Turkers' Career Goals and Professional Profiles". In: *ACM Collective Intelligence Conference 2018*. 2018 (cit. on pp. 34, 63).

[117] Anna Kasunic, Chun-Wei Chiang, Geoff Kaufman, and Saiph Savage. "Turker Tales: Integrating Tangential Play into Crowd Work". In: *Proceedings of the 2019 on Designing Interactive Systems Conference*. ACM. 2019, pp. 21–34 (cit. on p. 61).

[118] Geoff Kaufman and Mary Flanagan. "A psychologically "embedded" approach to designing games for prosocial causes". In: *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 9.3 (2015) (cit. on p. 188).

[119] Geoff Kaufman, Mary Flanagan, and Max Seidman. "Creating Stealth Game Interventions for Attitude and Behavior Change". In: *Transactions of the Digital Games Research Association* 2.3 (2016), pp. 173–193 (cit. on p. 188).

[120] Geoff F. Kaufman and Lisa K. Libby. "Changing beliefs and behavior through experience-taking." In: *Journal of Personality and Social Psychology* 103.2 (2012), pp. 1–19 (cit. on pp. 28, 159).

[121] Geoff F Kaufman and Lisa K Libby. "Changing beliefs and behavior through experience-taking." In: *Journal of personality and social psychology* 103.1 (2012), p. 1 (cit. on p. 169).

[122] A Ianus Keller, Aldo Hoeben, and Aadjan Van der Helm. "Cabinet: merging designers' digital and physical collections of visual materials". In: *Personal and Ubiquitous Computing* 10.2-3 (2006), pp. 183–186 (cit. on p. 152).

[123] Os Keyes. "The misgendering machines: Trans/HCI implications of automatic gender recognition". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (2018), p. 88 (cit. on p. 13).

[124] David Comer Kidd and Emanuele Castano. "Reading literary fiction improves theory of mind". In: *Science* 342.6156 (2013), pp. 377–380 (cit. on p. 28).

[125] Youjeong Kim and S Shyam Sundar. "Visualizing ideal self vs. actual self through avatars: Impact on preventive health outcomes". In: *Computers in Human Behavior* 28.4 (2012), pp. 1356–1364 (cit. on p. 34).

[126] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, et al. "The Future of Crowd Work". In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. CSCW '13. San Antonio, Texas, USA: ACM, 2013, pp. 1301–1318 (cit. on p. 34).

[127] John Kolligian Jr and Robert J Sternberg. "Perceived Fraudulence in Young Adults: Is There an'Imposter Syndrome'?" In: *Journal of personality assessment* 56.2 (1991), pp. 308–326 (cit. on p. 149).

[128] Robert E. Kraut, Paul Resnick, and Sara Kiesler. *Building successful online communities: evidence-based social design*. The MIT Press, 2016 (cit. on pp. 30, 32).

[129] Ethan Kross and Ozlem Ayduk. "Making meaning out of negative experiences by self-distancing". In: *Current directions in psychological science* 20.3 (2011), pp. 187–191 (cit. on p. 149).

[130] Lisa Kurt, William Kurt, and Ann Medaille. "The power of play: Fostering creativity and innovation in libraries." In: *Journal of Library Innovation* 1.1 (2010) (cit. on p. 149).

[131] Sybille Lammes. "Spatial regimes of the digital playground: cultural functions of spatial identification in post-colonial computergames". In: *Conference proceedings Mediaterr@–Gaming realities: A challenge for digital culture, ed. Manthos Santorineos*. 2006, pp. 236–43 (cit. on p. 24).

[132] Rochelle LaPlante, M Six Silberman, and Industriegewerkschaft Metall. "Building trust in crowd worker forums: Worker ownership, governance, and work outcomes". In: *Proceedings of WebSci16. ACM* (2016) (cit. on pp. 35, 63).

[133] Edith Law, Ming Yin, Joslin Goh, et al. "Curiosity killed the cat, but makes crowdwork better". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM. 2016, pp. 4098–4110 (cit. on pp. 35, 63).

[134] Maurizio Lazzarato. "Immaterial labor". In: *Radical thought in Italy: A potential politics* 1996 (1996), pp. 133–47 (cit. on p. 36).

[135] Nicole Lazzaro. "Why we play games: Four keys to more emotion without story". In: (2004) (cit. on pp. 25, 30, 149, 162, 172, 186).

[136] John Le, Andy Edmonds, Vaughn Hester, and Lukas Biewald. "Ensuring quality in crowdsourced search relevance evaluation". In: *SIGIR workshop on crowdsourcing for search eval*. 2010 (cit. on p. 35).

[137] Katia Levecque, Frederik Anseel, Alain De Beuckelaer, Johan Van der Heyden, and Lydia Gisle. "Work organization and mental health problems in PhD students". In: *Research Policy* 46.4 (2017), pp. 868–879 (cit. on p. 148).

[138] Chih-Ping Li and En-Chi Chiu. "Construct validity of the Empowerment Scale in patients with schizophrenia". In: *Neuropsychiatry* 7.5 (2017), pp. 501–508 (cit. on p. 134).

[139] Alan Lundgard, Yiwei Yang, Maya L Foster, and Walter S Lasecki. "Bolt: Instantaneous Crowdsourcing via Just-in-Time Training". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 467 (cit. on pp. 35, 63).

[140] Anders Lundström. "Designing energy-sensitive interactions: conceptualising energy from the perspective of electric cars". PhD thesis. KTH Royal Institute of Technology, 2016 (cit. on p. 152).

[141] Catherine M Rawlins. "The ivory tower of academia and how mental health is often neglected". In: *Future Science OA* 5.4 (2019), FSO392 (cit. on p. 148).

[142] Charalampos Mainemelis and Sarah Ronson. "Ideas are born in fields of play: Towards a theory of play and creativity in organizational settings". In: *Research in Organizational Behavior* 27 (2006), pp. 81–131 (cit. on p. 149).

[143] *Majority of Americans now use Facebook, YouTube | Pew Research Center*. `https://www.pewinternet.org/2018/03/01/social-media-use-in-2018/pi_2018-03-01_social-media_0-01/`. (Accessed on 04/25/2019) (cit. on p. 84).

[144] Thomas M Malaby. "Beyond play: A new approach to games". In: *Games and culture* 2.2 (2007), pp. 95–113 (cit. on p. 24).

[145] David Martin, Benjamin V Hanrahan, Jacki O'Neill, and Neha Gupta. "Being a turker". In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM. 2014, pp. 224–235 (cit. on p. 34).

[146] Alice E Marwick. *Status update: Celebrity, publicity, and branding in the social media age*. Yale University Press, 2013 (cit. on p. 85).

[147] Adrienne L. Massanari. *Participatory Culture, Community, and Play*. Peter Lang, 2015 (cit. on pp. 31, 57).

[148] Adrienne L Massanari. *Participatory culture, community, and play*. 2015 (cit. on pp. 40, 57, 83, 162, 180).

[149] Rollo May. "Everyday creativity: Our hidden potential". In: *Everyday Creativity and New Views of Human Nature. American Psychological Association, Washington* (2007) (cit. on p. 149).

[150] Elijah Mayfield and Carolyn Penstein Rosé. "LightSIDE: Open source machine learning for text". In: *Handbook of Automated Essay Evaluation*. Routledge, 2013, pp. 146–157 (cit. on p. 71).

[151] Elton Mayo. "The Hawthorne Experiment. Western Electric Company". In: *2016). Classics of organization theory* (1933), pp. 134–141 (cit. on pp. 23, 137, 173).

[152] Paul McDonald. "Digital discords in the online media economy: Advertising versus content versus copyright". In: *The YouTube Reader*. National Library of Sweden, 2009, pp. 387–405 (cit. on p. 36).

[153] Tiffany McDowell. "Fun at work: Scale development, confirmatory factor analysis, and links to organizational outcomes." In: (2004) (cit. on p. 63).

[154] A Peter McGraw and Caleb Warren. "Benign violations: Making immoral behavior funny". In: *Psychological Science* 21.8 (2010). Ed. by Salvatore Attardo, pp. 1141–1149 (cit. on p. 29).

[155] Markus Montola. "On the edge of the magic circle: Understanding pervasive games and role-playing". In: (2012) (cit. on p. 24).

[156] Markus Montola, Jaakko Stenros, and Annika Waern. *Pervasive games: theory and design*. CRC Press, 2009 (cit. on p. 24).

[157] Antonio Negri, Michael Hardt, and David Camfield. "Multitude: war and democracy in the age of empire". In: *Labour* 56 (2005), p. 359 (cit. on p. 36).

[158] Collette Nicolle and Julio Abascal. *Inclusive design guidelines for HCI*. CRC Press, 2014 (cit. on p. 13).

[159] Jakob Nielsen. *10 Heuristics for User Interface Design*. 1994 (cit. on p. 157).

[160] Kristine L Nowak. "Choosing Buddy Icons that look like me or represent my personality: Using Buddy Icons for social presence". In: *Computers in Human Behavior* 29.4 (2013), pp. 1456–1464 (cit. on p. 33).

[161] William Odom, Richard Banks, Abigail Durrant, David Kirk, and James Pierce. "Slow technology: critical reflection and future directions". In: *Proceedings of the Designing Interactive Systems Conference*. ACM. 2012, pp. 816–817 (cit. on p. 22).

[162] Bradley M Okdie, Rosanna E Guadagno, Frank J Bernieri, Andrew L Geers, and Amber R Mclarney-Vesotski. "Getting to know you: Face-to-face versus online interactions". In: *Computers in human behavior* 27.1 (2011), pp. 153–159 (cit. on p. 33).

[163] David Oleson, Alexander Sorokin, Greg P Laughlin, et al. "Programmatic Gold: Targeted and Scalable Quality Assurance in Crowdsourcing." In: *Human computation* 11.11 (2011) (cit. on p. 35).

[164] Judith S Olson and Wendy A Kellogg. *Ways of Knowing in HCI*. Vol. 2. Springer, 2014 (cit. on p. 150).

[165] Ryan Olson, Jessica Verley, Lindsey Santos, and Coresta Salas. "What we teach students about the Hawthorne studies: A review of content within a sample of introductory IO and OB textbooks". In: *The Industrial-Organizational Psychologist* 41.3 (2004), pp. 23–39 (cit. on p. 23).

[166] Daniel Pargman and Peter Jakobsson. "Do you believe in magic? Computer games in everyday life". In: *European Journal of Cultural Studies* 11.2 (2008), pp. 225–244 (cit. on p. 24).

[167] Anna Parkman. "The imposter phenomenon in higher education: Incidence and impact". In: *Journal of Higher Education Theory and Practice* 16.1 (2016), p. 51 (cit. on p. 149).

[168] Max Pashkevich, Sundar Dorai-Raj, Melanie Kellar, and Dan Zigmond. "Empowering online advertisements by empowering viewers with the right to choose: the relative effectiveness of skippable video advertisements on YouTube". In: *Journal of Advertising Research* 52.4 (2012), pp. 451–457 (cit. on p. 86).

[169] Frank Pasquale. *The black box society*. Harvard University Press, 2015 (cit. on p. 36).

[170] Jessica Annette Pater, Yacin Nadji, Elizabeth D. Mynatt, and Amy S. Bruckman. "Just Awful Enough". In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI 14* (2014) (cit. on pp. 32, 46).

[171] James W Pennebaker. "Writing about emotional experiences as a therapeutic process". In: *Psychological science* 8.3 (1997), pp. 162–166 (cit. on p. 65).

[172] Whitney Phillips. *This is why we can't have nice things: Mapping the relationship between online trolling and mainstream culture*. MIT Press, 2015 (cit. on pp. 28, 29, 48, 162).

[173] Whitney Phillips and Ryan M Milner. *The ambivalent internet: Mischief, oddity, and antagonism online*. John Wiley & Sons, 2018 (cit. on pp. 28, 162).

[174] James Pierce. "Working by Not Quite Working: Resistance as a Technique for Alternative and Oppositional Designs". In: (2015) (cit. on p. 22).

[175] James Pierce, Phoebe Sengers, Tad Hirsch, et al. "Expanding and refining design and criticality in HCI". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 2083–2092 (cit. on pp. 22, 158).

[176] Bohdan Pikas and Gabi Sorrentino. "The effectiveness of online advertising: Consumer's perceptions of ads on Facebook, Twitter and YouTube". In: *Journal of Applied Business and Economics* 16.4 (2014), pp. 70–81 (cit. on p. 86).

[177] Barbara Plester and Ann Hutchison. "Fun times: the relationship between fun and workplace engagement". In: *Employee Relations* 38.3 (2016), pp. 332–350 (cit. on p. 63).

[178] Hector Postigo. "The socio-technical architecture of digital labor: Converting play into YouTube money". In: *New media & society* 18.2 (2016), pp. 332–349 (cit. on p. 36).

[179] Daniel Pressnitzer, Jackson Graves, Claire Chambers, Vincent De Gardelle, and Paul Egré. "Auditory Perception: Laurel and Yanny Together at Last". In: *Current Biology* 28.13 (2018), R739–R741 (cit. on p. 69).

[180] Stephen D Reicher, Russell Spears, and Tom Postmes. "A social identity model of deindividuation phenomena". In: *European review of social psychology* 6.1 (1995), pp. 161–198 (cit. on p. 32).

[181] Steven Reiss. "Multifaceted nature of intrinsic motivation: The theory of 16 basic desires". In: *Review of general psychology* 8.3 (2004), pp. 179–193 (cit. on p. 30).

[182] Horst WJ Rittel and Melvin M Webber. "Wicked problems". In: *Man-made Futures* 26.1 (1974), pp. 272–280 (cit. on pp. 21, 146).

[183] George Ritzer and Nathan Jurgenson. "Production, consumption, prosumption: The nature of capitalism in the age of the digital 'prosumer'". In: *Journal of consumer culture* 10.1 (2010), pp. 13–36 (cit. on p. 36).

[184] Joel Ross, Lilly Irani, M Silberman, Andrew Zaldivar, and Bill Tomlinson. "Who are the crowdworkers?: Shifting demographics in mechanical turk". In: *CHI'10 extended abstracts on Human factors in computing systems*. ACM. 2010, pp. 2863–2872 (cit. on p. 34).

[185] Hiro Saito, Gemma Irving, Jan Packer, David Solnet, et al. "The staff break room as an oasis: Emotional labour, restorative environments and employee wellbeing in the hospitality industry". In: *CAUTHE 2015: Rising Tides and Sea Changes: Adaptation and Innovation in Tourism and Hospitality* (2015), p. 694 (cit. on p. 63).

[186] Susumu Saito, Chun-Wei Chiang, Saiph Savage, et al. "TurkScanner: Predicting the Hourly Wage of Microtasks". In: *arXiv preprint arXiv:1903.07032* (2019) (cit. on p. 35).

[187] Niloufar Salehi, Lilly C. Irani, Michael S. Bernstein, et al. "We Are Dynamo: Overcoming Stalling and Friction in Collective Action for Crowd Workers". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, 2015, pp. 1621–1630 (cit. on pp. 35, 63).

[188] Katie Salen and Eric Zimmerman. "The game design reader". In: *USA: MIT Sida.(2007). Retrieved May* 15 (2006), p. 2007 (cit. on p. 24).

[189] Katie Salen, Katie Salen Tekinbaş, and Eric Zimmerman. *Rules of play: Game design fundamentals*. MIT press, 2004 (cit. on pp. 9, 23, 24).

[190] Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. "Intersectional HCI: Engaging Identity Through Gender, Race, and Class". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: ACM, 2017, pp. 5412–5427 (cit. on p. 13).

[191] Ari Schlesinger, Kenton P. O'Hara, and Alex S. Taylor. "Let's Talk About Race: Identity, Chatbots, and AI". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, 315:1–315:14 (cit. on p. 13).

[192] Trebor Scholz. "Introduction: Why does digital labor matter now?" In: *Digital Labor*. Routledge, 2012, pp. 9–18 (cit. on pp. 27, 36, 83).

[193] Peter Schultes, Verena Dorner, and Franz Lehner. "Leave a Comment! An In-Depth Analysis of User Comments on YouTube." In: *Wirtschaftsinformatik* 42 (2013), pp. 659–673 (cit. on p. 35).

[194] Jan H Schumann, Florian von Wangenheim, and Nicole Groene. "Targeted online advertising: Using reciprocity appeals to increase acceptance among users of free web services". In: *Journal of Marketing* 78.1 (2014), pp. 59–75 (cit. on p. 85).

[195] Avi Segal, Ya'akov (Kobi) Gal, Robert J. Simpson, et al. "Improving Productivity in Citizen Science Through Controlled Intervention". In: *Proceedings of the 24th International Conference on World Wide Web*. WWW '15 Companion. Florence, Italy: International World Wide Web Conferences Steering Committee, 2015, pp. 331–337 (cit. on p. 35).

[196] Gwendolyn Seidman. "Self-presentation and belonging on Facebook: How personality influences social media use and motivations". In: *Personality and Individual Differences* 54.3 (2013), pp. 402–407 (cit. on p. 33).

[197] Theresa M Senft. "Microcelebrity and the branded self". In: *A companion to new media dynamics* (2013), pp. 346–354 (cit. on p. 85).

[198] Nathan Shedroff. *Design is the problem: the future of design must be sustainable*. Rosenfeld Media, 2009 (cit. on p. 90).

[199] Miguel Sicart. "Darkly Playing Others: The Aesthetics of Abusive Game Design". In: (2015) (cit. on p. 26).

[200] Miguel Sicart. *Play matters*. MIT Press, 2014 (cit. on pp. 9, 26, 153, 161).

[201] Strike Social. *2018 YouTube Data Report*. http://go.strikesocial.com/rs/592-XWM-066/images/2018_YouTube_Data_Report.pdf. (Accessed on 04/25/2019). 2018 (cit. on p. 84).

[202] Olivia Solon. *Google's bad week: YouTube loses millions as advertising row reaches US | Technology | The Guardian*. https://www.theguardian.com/technology/2017/mar/25/google-youtube-advertising-extremist-content-att-verizon. (Accessed on 04/25/2019). 2017 (cit. on p. 36).

[203] Paul W Speer and Joseph Hughey. "Community organizing: An ecological route to empowerment and power". In: *American journal of community psychology* 23.5 (1995), pp. 729–748 (cit. on p. 35).

[204] Susan Leigh Star and Anselm Strauss. "Layers of silence, arenas of voice: The ecology of visible and invisible work". In: *Computer supported cooperative work (CSCW)* 8.1-2 (1999), pp. 9–30 (cit. on p. 34).

[205] Jaakko Stenros. "In defence of a magic circle: the social, mental and cultural boundaries of play". In: *Transactions of the Digital Games Research Association* 1.2 (2014) (cit. on pp. 24, 25).

[206] Matt Stopera. *I Followed My Stolen iPhone Across The World, Became A Celebrity In China, And Found A Friend For Life*. 2015 (cit. on p. 77).

[207] Bernard Suits. *The Grasshopper-: Games, Life and Utopia*. Broadview Press, 2014 (cit. on pp. 10, 24, 27).

[208] Brian Sutton-Smith. *The ambiguity of play*. Harvard University Press, 2009 (cit. on p. 27).

[209] Aida Isabel Tavares. "Telework and health effects review". In: *International Journal of Healthcare* 3.2 (2017), p. 30 (cit. on pp. 34, 63).

[210] Tina L Taylor. *Play between worlds: Exploring online game culture*. Mit Press, 2009 (cit. on p. 24).

[211] Alexandra To, Joselyn McDonald, Jarrek Holmes, Geoff Kaufman, and Jessica Hammer. "Character Diversity in Digital and Non-Digital Games". In: *Transactions of the Digital Games Research Association* 4.1 (2018) (cit. on p. 13).

[212] Alexandra To, Jessica Hammer, and Geoff Kaufman. "Promoting Digital Wellbeing by Empowering Users from Racial Minority Groups". In: *Proceedings of ACM CHI conference, May 2019*. ACM. 2019 (cit. on p. 13).

[213] Meera Vijay and Nitin Vazirani. "Emerging Paradigm-Fun in Workplace to Alleviate Stress." In: *SIES Journal of Management* 7.2 (2011) (cit. on p. 61).

[214] Nina Wallerstein and Edward Bernstein. "Introduction to community empowerment, participatory education, and health". In: *Health education quarterly* 21.2 (1994), pp. 141–148 (cit. on p. 35).

[215] Xinyi Wang, Haiyi Zhu, Yangyun Li, Yu Cui, and Joseph Konstan. "A Community Rather Than A Union: Understanding Self-Organization Phenomenon on MTurk and How It Impacts Turkers and Requesters". In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '17. Denver, Colorado, USA: ACM, 2017, pp. 2210–2216 (cit. on pp. 35, 63).

[216] Caleb Warren and A Peter McGraw. "Benign Violation Theory". In: *Encyclopedia of Humor Studies*. SAGE Publications, Inc, 2014 (cit. on pp. 29, 114).

[217] Stephan AG Wensveen, Johan Partomo Djajadiningrat, and CJ Overbeeke. "Interaction frogger: a design framework to couple action and function through feedback and feedforward". In: *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM. 2004, pp. 177–184 (cit. on p. 152).

[218] Samuel West and ECC Shiu. "Play as a facilitator of organizational creativity". In: *Creativity research: An inter-disciplinary and multi-disciplinary research handbook* (2014), pp. 191–206 (cit. on p. 149).

[219] Monica T Whitty. "Revealing the 'real'me, searching for the 'actual'you: Presentations of self on an internet dating site". In: *Computers in Human Behavior* 24.4 (2008), pp. 1707–1723 (cit. on p. 34).

[220] *Who uses Pinterest, Snapchat, YouTube and WhatsApp | Pew Research Center*. `https://www.pewinternet.org/chart/who-uses-pinterest-snapchat-youtube-and-whatsapp/`. (Accessed on 04/25/2019) (cit. on p. 84).

[221] Douglas Wilson and Miguel Sicart. "Now it's personal: on abusive game design". In: *Proceedings of the International Academic Conference on the Future of Game Design and Technology*. ACM. 2010, pp. 40–47 (cit. on p. 26).

[222] Jacob O Wobbrock and Julie A Kientz. "Research contributions in human-computer interaction". In: *interactions* 23.3 (2016), pp. 38–44 (cit. on pp. 142, 143, 150).

[223] Alia Wong. *Mental Health Disorders Disproportionately Affect Graduate School Students - The Atlantic*. `https://www.theatlantic.com/education/archive/2018/11/anxiety-depression-mental-health-graduate-school/576769/`. (Accessed on 06/13/2019). 2018 (cit. on p. 148).

[224] Allison Woodruff, Sarah E Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. "A qualitative exploration of perceptions of algorithmic fairness". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 656 (cit. on p. 13).

[225] Xiaohui Yan, Jiafeng Guo, Yanyan Lan, and Xueqi Cheng. "A biterm topic model for short texts". In: *Proceedings of the 22nd international conference on World Wide Web*. ACM. 2013, pp. 1445–1456 (cit. on p. 71).

[226] *Yanny or Laurel*. 2018 (cit. on p. 69).

[227] *YouTube ad boycott: AT&T, Disney drop out over pedophile concerns*. `https://www.usatoday.com/story/tech/news/2019/02/22/at-t-disney-epic-games-pull-youtube-ads-child-exploitation-concerns-pedophiles/2948825002/`. (Accessed on 04/25/2019) (cit. on p. 36).

[228] *YouTube as a participatory culture - Chau - 2010 - New Directions for Youth Development - Wiley Online Library*. `https://onlinelibrary.wiley.com/doi/abs/10.1002/yd.376`. (Accessed on 04/25/2019) (cit. on p. 35).

[229] José P Zagal, Staffan Björk, and Chris Lewis. "Dark patterns in the design of games". In: *Foundations of Digital Games 2013*. 2013 (cit. on p. 26).

[230] Yu Zhang and Mihaela Van der Schaar. "Reputation-based incentive protocols in crowdsourcing applications". In: *INFOCOM, 2012 Proceedings IEEE*. IEEE. 2012, pp. 2140–2148 (cit. on p. 35).

[231] John Zimmerman and Jodi Forlizzi. "Research through design in HCI". In: *Ways of Knowing in HCI*. Springer, 2014, pp. 167–189 (cit. on pp. 21, 158).

[232] John Zimmerman, Erik Stolterman, and Jodi Forlizzi. "An analysis and critique of Research through Design: towards a formalization of a research approach". In: *proceedings of the 8th ACM conference on designing interactive systems*. ACM. 2010, pp. 310–319 (cit. on p. 21).

[233] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. "Research through design as a method for interaction design research in HCI". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2007, pp. 493–502 (cit. on p. 21).

List of Figures

List of Tables