The Effects of Ambiguity: A Feminist Study of Human Signifiers in Human-Computer Interaction

by

Cayley MacArthur

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Arts
in
English Language and Literature & Systems Design Engineering

Waterloo, Ontario, Canada, 2015

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

I would like to acknowledge the names of my co-authors who contributed to the study described in this thesis, these include:

- Adam Bradley,
- Dr. Mark Hancock, and
- Dr. Sheelagh Carpendale.

Abstract

A lack of diversity in STEM (science, technology, engineering and math) fields has been a popular topic of discussion and a persistent challenge in terms of recruitment, engagement, opportunity and equality spanning decades. It is not well-understood how new technologies created by the human-computer interaction (HCI) community affect identity construction in minorities such as women, people of colour, and persons with disabilities. The Feminist HCI movement calls on the community to engage with broader moral commitments such as empowerment, diversity, identity and equity, as well as epistemologically embracing multiplicity over universality. However, a gap remains between theory and practice.

I argue that goals of feminist rhetoric must be met in the community's discourse habits as an essential step to more inclusivity in HCI; these goals include dismantling certain oppressive language structures. Feminist theory suggests that the abstract, gender-neutral language used to talk about people in male-dominated systems actually elicits imagery perceived to be male, which would function in contrast to the intentions of writers who want to be gender-neutral when using these words. In this thesis, I present a study used to determine whether the human signifiers used in HCI publications have the same effect. Findings suggest that these HCI "people" words do generally have a tendency to be perceived as male. Insights from this study reveal some tendencies in how these words are thought of among a general audience. I recommend a stronger commitment to Feminist HCI in theory and practice, and greater awareness and sensitivity towards the connotations elicited by these falsely universal terms that are not representative of the diverse community within and outside of HCI.

Acknowledgements

The title of this thesis is a reference to Simone de Beauvoir's "The Ethics of Ambiguity" (1948). This research would not be possible were it not for a long line of inspiring researchers before me, nor without the support and input of many others whom I would like to thank here.

Thank you to my supervisors, Dr. Neil Randall and Dr. Mark Hancock. With your encouragement and support I have pursued new challenges and opportunities to become a better student and researcher. You go above and beyond for your students and it is very appreciated.

Thank you to Dr. Cathy Burns and Dr. Michael MacDonald for serving as readers of this thesis. I also would like to thank my collaborators on the study contained in this thesis, Dr. Hancock, Dr. Sheelagh Carpendale, and special thanks to Adam Bradley, who has been an excellent mentor, role model, constructive critic, and all-around life coach over the past few years.

Thank you to the members of the Touchlab, the Collaborative Systems Laboratory, the Social Usable Interactive Technologies group, and the wider Games Institute community: I am lucky to work among such a particularly fun, motivating, and bright group of people.

Thank you to the IMMERSe network and the English department at the University of Waterloo for providing funding during my degree.

Thank you to my family for keeping me afloat, and to my friends for being there even when I spent more quality time with my thesis. Most of all, thank you Mike for your unwavering positivity and contagious enthusiasm throughout this process.

This thesis is dedicated to Abby, who didn't get to see it through.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Throughout history, the language used by scientists and engineers has developed in lockstep with the evolution of their theoretical approaches. Within a Western cultural framework, the rationalist tradition of STEM disciplines is reflected in a communication style that supports their commitment to objectivity: the language used is typically abstract and disembodied, with the words used to describe *people* in these domains following the same pattern. As such, when scientific literature references a "subject", it ends up being conflated with a "universal". These subject terms act as *false universals*: they are meant to represent everyone, but signify a masculine epistemological subject in the mind of the reader. Significant advances have been made in science and engineering to be more considerate and inclusive of women and minorities, with the common practice in scientific writing being to use gender-neutral, ambiguous terms to refer to people. In the specific area of Human-Computer Interaction (HCI), at its top conference in 2014 the most frequently used words to describe people in the literature were *user*, *participant*, *person*, *designer*, and *researcher*. In this thesis, I explore the use of these "people words" in HCI literature to determine whether the ambiguity of these terms lend them to functioning as universal signifiers, or whether they carry gendered connotations. I also present the study method developed to conduct this investigation.

Core to research and practice in HCI are the concepts of *user*-centered and *human*-centered design. There is ongoing debate about which of these linguistic frames is most appropriate to meet the aims of the discipline (i.e. Gasson, 2003), demonstrating that there is already concern in the HCI community around the way people are thought of in the literature and whether it impacts their work. The community actively encourages diversity, and the development of Feminist HCI calls for more explicit engagement with the central commitments of feminism: "agency, fulfillment, identity and the self, equity, empowerment, diversity, and social justice" (Bardzell, 2010, p. 1301). Despite efforts to remain "gender neutral" in their writing, there is insufficient research exploring whether the language used in HCI literature introduces bias.

The binary language of universal rationality in "objective" disciplines may prevent meaningful engagement with the central commitments called on by Feminist HCI researchers. Questioning the formation of "subjects" in feminist criticism is crucial for these scholars to reveal social structures that exist but remain invisible in well-established disciplines, and by doing so they create opportunities for new types of involvement. A key goal of feminist rhetoric is to dismantle

structures that insist on universal truths in order to open them up to a plurality of meanings, which is why I will take this approach to explore whether the "people words" in HCI are acting as false universals. Using these kinds of words may lead to a distorted and one-sided perception of the subject during design and research, rather than the consideration of marginalized perspectives at the heart of feminist HCI.

1.1 Motivation

The language used in HCI literature affects how we research, design, and experience technology. Despite intentions of researchers to remain "gender neutral" in their communication, there is insufficient research exploring whether this language introduces (potentially unconscious) bias. The effects of universalizing terminology on products of the HCI community are detectable. Huff & Cooper (1987) show that software designed for boys, students, or girls results in one solution for "boys" and "students" and a different approach for "girls". When the abstract term is used, gender is no longer implicated and a masculine approach dominates. Testing men and women in the same environments on the same tasks may lead to an assumption that women perform worse than men, when in reality slight tweaks are needed in designing the task that significantly reduce differences in performance (Czerwinski, Tan, & Robertson, 2002; Tan, Czerwinski, & Robertson, 2003). Using static binary categories to distinguish between men and women in such research constrains the "imaginable domain of gender" (Butler, 1990, p. 9), making it more difficult to realize that the issue is not that women perform worse, but that the approach to design favours a masculine perspective in the first place.

Word use in scientific literature becomes problematic when it can be shown to "produce gendered subjects along a differential axis of domination or to produce subjects who are presumed to be masculine"—an appeal to such a system for release will be self-defeating (Butler, 1990, p. 2). In other words, the "subject" is conceived within a representational system where they are outside of the norm, and in such a system where abstract referents to people tend to be thought of as male, the other viewpoints may not be sufficiently or adequately represented. On foundations laid by Luce Irigaray and Simone de Beauvoir, Judith Butler argues that the relation between masculine and feminine cannot even be properly represented in a signifying economy in which the masculine "constitutes the closed circle of signifier and signified" (1990, p. 11). An ability to demonstrate that the signifiers used to describe people in HCI literature are actually presumed to be male by readers would raise

questions about the discursive structures in place, prevent those terms from continuing to be thought of as universal, and motivate linguistic change within the community to become more inclusive.

Gender and identity are constructed through everyday experience, and the experiences with technology created by the HCI community have a large opportunity to influence such aspects of daily life: feminist HCI emphasizes a strong commitment to people's experiences (Bardzell & Bardzell, 2011, p. 682). A key component of feminist epistemologies is an emphasis on the experiences of the marginalized; Bardzell & Bardzell (2011) raise this as both an ethical and epistemological imperative. However, if the language used to describe people in HCI literature has a tendency to frame them as masculine, then this would stand in contrast with the authors' intentions when they chose to use words believed to be neutral or universal.

Though feminist research methods are agnostic about quantitative versus qualitative methods, recent work specifically in the HCI domain highlights that historically the methods put forth in this genre favour qualitative methods, which suggests an opportunity for innovation in more quantitative approaches in keeping with a goal to engender diversity in data sources and interpretation (Bardzell & Bardzell, 2011, p. 681). In this spirit, I will address the problem of gendered or universalizing language in HCI by incorporating research methods based on the rhetorical analysis of language in other domains, and collecting data from a large pool of participants using Mechanical Turk, an established approach to data collection within the HCI community (Mason & Suri, 2012). Butler (1990) showed that language is where issues in "representational politics" come to a head, and in her analysis she questions the signifier-signified relationship in traditionally male-dominated disciplines, challenging the ways people in these fields think and talk about other people to be more inclusive. In this thesis I evaluate the signified-signifier relationship specifically in HCI, which has not been questioned in such a way before. To date, work in feminist HCI has focussed on shifting problem-solving methods to be more inclusive, but no work has been done to examine the structures and invisible constraints in which feminist HCI is bound to operate to determine whether this is, as Butler suggests, a self-defeating system.

1.2 Research Problem

Feminist scholars throughout history have problematized the universalizing nature of language used in historically masculine, objective fields such as the sciences. The implications for using totalizing language when referring to people in such systems is that the abstract, disembodied subject is thought

to be masculine and the needs of the "other" in this hegemonic structure are not considered. In practicality, this would lead to what might be thought of as vague "gender-neutral" language being used in research & design in HCI: words such as *user*. Due to existing language, social and political systems, these terms may not truly be gender-neutral and may lead to exclusion of those not within the normative masculine group. In order to be aware of the implications of language used in HCI literature, it was necessary to develop a method for investigating the perception of these words among a general audience for a baseline to determine if they are functioning as "universals" or if they carry more complex meanings.

1.3 Research Question and Contributions

The overall research question that is addressed in this thesis is:

Are the words used to describe people in HCI papers perceived as gender neutral?

In the process of addressing the above research question, three main contributions were made:

- **Determining the theoretical framework for evaluating the implications of gendered language in HCI.** I conducted a literature review to determine the intersections between feminist theory, language, and HCI. This review, presented in Chapter 2, also focuses on previous advances in feminist approaches to HCI, highlights the applicability of feminist HCI methods, and reveals areas requiring further research.
- **Designing a research study to investigate how the most common words that have been used to refer to people in recent HCI literature are received by non-experts.** Based on a review of past work, an interdisciplinary approach was used to collect imagery elicited by HCI's "people words" along with demographic questionnaire data. We developed a new study method to collect and evaluate language used in HCI research papers and we found that this was an effective way to conduct research studies of this type. A detailed description of the study, as well as a discussion of the rationale behind its design, is presented in Chapter 3.
- **Evaluating the results on the perception of gender in HCI language.** HCI's "people words" (user, participant, person, designer, and researcher) were presented as five separate study conditions in order to elicit drawings from non-experts indicating how they perceive these words. We found that these words were predominantly perceived as male by our

participants, with the exception of females being just as likely to draw “participant” and “person” as male or female. The findings and discussion based on study results are described in Chapter 4 and Chapter 5, respectively.

1.4 Thesis Organization

This thesis is organized into five chapters following the introduction:

- Chapter 2, Background – In this chapter, I review research literature related to the role of language in HCI, feminist issues in HCI, contributions of feminist research to HCI, and research methods that support the goal of understanding language within the HCI framework.
- Chapter 3, Study Method and Findings – I describe the design of the Mechanical Turk study used to collect drawings and questionnaire data about drawing contents as well as respondents.
- Chapter 4, Qualitative Analysis – In this chapter I present the qualitative and quantitative results of the study.
- Chapter 5, Discussion – Here, I situate the findings from the study in the larger context of gendered language and feminist HCI, presents considerations for language use going forward, and discusses the limitations of this research.
- Chapter 6, Conclusions – Finally, I discuss how the research objectives were met and presents recommendations for future work.

Chapter 2
Background

This chapter provides an overview of the literature and previous work around the relationship between language and human-computer interaction, and complemented by feminist theory where applicable. The background provided here is divided into four primary categories. First, in Section 2.1 I discuss existing feminist issues within the HCI community in terms of software design, hardware design, online, in the design of "gendered systems" (i.e. those with a voice), and in theory. I also highlight some previous efforts to address these issues, instances where feminist theory has been applied in addressing challenges in HCI, and reflect on where gaps remain. In Section 2.2 I raise a specific question of the "gender neutral" or "universal" language commonly used in HCI and discuss why this is of concern. In Section 2.3 I discuss past work in feminist theory that provides context for my own frame when investigating gendered language in HCI. Finally, in Section 2.4 I show how these threads come together to frame my methodological approach to understanding the perception and implications of the "people" words in HCI that are intended to be all-encompassing.

2.1 Existing Diversity Challenges in HCI

At the core of feminist philosophy is the idea that the views informing Western popular culture and philosophy including predominant concepts of "the self" are derived from the experience of the "predominantly white and heterosexual, mostly economically advantaged men who have wielded social, economic, and political power and who have dominated the arts, literature, the media, and scholarship" (Willett, Anderson, & Meyers, 2015). As a result, a feminist approach to various disciplines tends to involve taking into account individuals' unique standpoints and how social, environmental and other factors shape their perspectives. Feminist theorist Sandra Harding's neologism "strong objectivity" refers to a notion that the perspectives of marginalized groups can actually create more objective, or stronger accounts of the world, in opposition to the "weak" objectivity posited by supposed value-neutral research. As such, one can find traditionally "value-neutral" disciplines such as architecture looking to a feminist approach to breathe life into established traditions, calling to "reconstruct the unheard languages of the modern landscape as a means to reinvigorate contemporary design practice" (Meyer, 1997). Some degree of similar progress has been made in human-computer interaction, which has been productive in engaging practitioners but does leave more to be desired. In the following subsections I discuss this progress in more detail in terms

of software design challenges (Section 2.1.1), the design of hardware and other physical products (Section 2.1.2), research on the Internet (Section 2.1.3), in the design of interactive systems (Section 2.1.4), and in how the HCI community reads its own theories (Section 2.1.5). In Section 2.1.6 I will discuss efforts to address problems raised in the aforementioned areas.

2.1.1 "Agnostic" Software Design

One of the most literal ways software designers might account for gender differences is thinking about how people interact with a system and whether one can design to account for these differences. Burnett et al. (2011) focus on software tools that are intended to support end-users in problem-solving tasks, raising the point that such tools typically are designed to be "gender agnostic", meaning that they are made for a genderless user with universal characteristics. Although gender agnosticism in software design is well intended or even an unconscious design choice, Burnett et al. found that gender differences occurred in "feature usage, feature-related confidence, and tinkering (playful exploration) with features" (2011, p. 450). In their study, they found females had lower confidence but the same speed of learning new features as males. Due to lower confidence, they were slower to adopt new tools, which hampered their performance and turned their negative self-perceptions into a self-fulfilling prophecy. While confidence levels differed, these differences "clearly were not the sole explanation for the differences in feature usage and tinkering" (2011, p. 459). An important contribution of this work is highlighting the fact that "software that assumes that users will discover advanced features solely through tinkering and exploring will likely marginalize females" (2011, p. 459). In this case, it is shown that designing for a universal leads to assumptions that disadvantage those outside of the normative group.

Making changes to account for gender differences doesn't mean that one group would have an unfair advantage. In their paper *Women Go With The (Optical) Flow*, Tan et al. (2003) demonstrate that features such as visual cues can be added to reduce the performance gap between men and women in their navigation of 3D virtual environments, without negatively impacting male performance. These cues help to support a variety of navigation styles that accommodate numerous ways of interacting with an environment, regardless of one's gender.

The impact of designing "gender agnostically" is to risk marginalizing groups by failing to take into account that a one-size-fits-all solution will not work to everyone's advantage. The research shows that by spending more time to understand the broader population of potential users, simple accommodations can enhance performance by that marginalized group with no negative impact on

others. Before that can happen, it takes an awareness of designers' assumptions about those who will use their product, and a method is needed to help bring those assumptions or expectations to the surface.

2.1.2 Hardware & Product Design

There are numerous publications outlining gender differences in male and female performance in spatial tests and navigation tasks. Kimura (1999) summarizes that most spatial tests demonstrate some advantage for men, "who excel particularly at imaginal rotation and targeting", however women are "generally better at recalling the positions of objects in an array, and at remembering landmarks along a route" (1999, p. 64). Overall, men are said to navigate with different wayfinding or spatial reference techniques than women. These findings cross into spatial tests and navigation tasks in virtual environments, where gender differences are even more exaggerated (Astur, Ortiz, & Sutherland, 1998; Sandstrom, Kaufman, & A. Huettel, 1998; Wallet, Sauzeon, Rodrigues, & N'Kaoua, n.d.). Beyond the design of the actual virtual environment, researchers have shown there are other options for manipulations to improve the performance gap. Czerwinski et al. (2002) demonstrated that by providing a wider field of view in the design of the display, females were able to offload certain cognitive processes to their perceptual system. This improved their performance with no decrement in male performance.

Outside the frame of computer hardware, the design of other "things" such as furniture and consumer products has also functioned to perpetuate gender gaps. In her book *The Hoosier Cabinet and Kitchen History*, Nancy Hiller shows how such an innocuous invention as the Hoosier cabinet (a cabinet similar to a tall buffet or kitchen baking cabinet) had a large impact on how housewives were connected to the idea of domestic work. Rather than liberating women from household work, it ended up identifying them with it. Many other physical products were designed in ways that were influenced by gender identity of the designer and assumptions about the target audience. For example, Cockburn & Ormrod (1993) have written an entire book about the microwave and note that early models were targeted at women because they were thought to do all of the food preparation. Steven Gelber (2000) describes how suburban men can assert their role in the home through the use of power tools for DIY projects, drawing on "the masculine legitimacy of skilled labour" such that they can "stay at home without feeling emasculated or being subsumed into an undifferentiated entity with his wife". The result is that men are participating in domestic labour while still preserving the boundaries between traditional gender roles. In *The Significant Screwdriver*, Bardzell et al. sought to "transgress social

norms regarding the gendered division of labour in the domestic sphere in hopes of yielding insights or orientations toward improving the quality of domestic life" (2011, p. 371). Foregrounded in this project was the idea that interaction designs, although they are usually designed for a particular problem, can have both positive and negative effects when they reach the real world and become more widespread. The work of researchers and designers is increasingly embedded in everyday life, and the authors call for the HCI field to both understand and be "accountable to the ways that interaction designs affect stakeholders in the broadest sense" (2011, p. 371).

Overall, these examples show how conscious or subconscious theories about "the nature of biological existence, about language, and about the nature of human action" can have a big impact on the shape of "what we build and how we use it" (Winograd, 1986, p. xii). The design of technologies can support or blur traditional gender roles, and researchers should think about the ways their own work does this and can be accountable for it. Research-through-design projects like *The Significant Screwdriver* can help the HCI community understand the assumptions behind more traditional ways of thinking that influence what is made and why it is made that way. By experimenting with the tradition of design and making its assumptions explicit, researchers and designers can "open [them]selves to alternatives and to the new design possibilities that flow from them" (1986, p. 5).

2.1.3 Gender on the Internet

Online communities are an important place where gendered identities can be acted out, explored, shed, or re-enforced. Early research in computer-mediated communication by Hiltz et al. (1986) found that increasing anonymity of individual members resulted in their engaging in more normative behaviour than they would otherwise. Savicki et al. (2006) write that as a consequence, "in situations where participants cannot be identified, the groups themselves may become 'gendered'" in that they would depend on more masculine or feminine modes of communication depending on how the group was composed. Researchers Hemphill & Otterbacher (2012) use the online community IMDb (the Internet Movie Database) to explore how two groups, men and women, use "linguistic patterns considered gendered" in the movie reviews that they contribute to the site, hypothesizing that due to a male majority, prolific female reviewers may shift their linguistic style to align with males. The authors report that male-authored reviews have a median utility score of 0.67 (meaning that 2/3 of readers found the reviews helpful) and received more votes, whereas female-authored reviews had a median of 0. Not only does this result in female-authored reviews receiving less attention and votes, IMDb filters search results on utility scores which presents readers with a predominantly male view.

If female reviewers want their voices to be heard, will they begin to communicate like someone else? Hemphill & Otterbacher demonstrate that the female reviewers' language does change over time and identifiable "female" markers decrease, but the perceived utility of their reviews by the online community does not. While there are not many definite conclusions that can be made, the study highlights an interesting relationship between language use, gender, and the perceived value of authors' contributions.

Another large online community suffering from a lack of minority representation is Wikipedia. A 2011 survey by the Wikimedia Foundation found that only 8.5% of Wikipedia editors are women (*Wikipedia Editors Study*, n.d.). Lam et al. (2011) find that even when controlling for variables that affect article length, articles of particular interest to women tend to be shorter and have a lower overall assessment of quality, suggesting that the gender gap has had a biased effect on content coverage. In their study, they also found a greater concentration of females in areas like People and Arts, whereas male editors were concentrated in Geography and Science. These findings come full-circle in a paper by Lim & Kwon (2010) wherein female Wikipedia readers had lower outcome expectations, perception of lower quality, lower confidence in evaluating information quality, and lower amounts of exploration or idle reading. Lam et al. postulate that the Wikipedia community may actually be resistant to female participation. They found a higher proportion of 'protected' articles, suggesting higher contention of female articles. Females are significantly more likely to have their edits reverted, and have a shorter activity lifespan. The results point to a concerning and unfortunate trend, where females are resistant to participate, and the resultant product does not represent a female point of view, therefore appearing less valid and of less interest to the female reader.

In recent years, events like Art+Feminism's Wikipedia Edit-A-Thon have been growing in popularity as Wikipedia's gender gap becomes better known. Art+Feminism is "a campaign to improve coverage of women and the arts on Wikipedia, and to encourage female editorship" ("About," n.d.). The Wikipedia Edit-A-Thon first occurred on International Women's Day in 2014, bringing together numerous volunteers all over the world to improve the coverage of female artists. Interestingly, Art+Feminism organizers state that Wikipedia has been very supportive of their initiative, even providing funding for childcare during the event, despite the site having banned "feminist" editors on accusations of not being "objective" during the GamerGate scandal when prominent women in the video game community were being harassed (Mandiberg, 2015).

The HCI community gets closer to introspection in studies like *Gender, Representation and Online Participation* (Vasilescu, Capiluppi, & Serebrenik, 2014) where the focus is on communities like StackOverflow (a programming Q&A community): "such dedicated online communities are an integral part of the working lives of software developers, exponents of Science, Technology, Engineering and Mathematics (STEM) professionals" (2014, p. 489). The authors suggest that communities utilizing competition and prizes to increase reputation may be detrimental to encouraging female participation, based on related research (Gneezy, Niederle, & Rustichini, 2003). The number of female contributors on StackOverflow was commensurate with the numbers of graduates from STEM fields at the time of the study, and like Wikipedia, they showed similar activity to men but departed from the community sooner. The types of incentives that encourage participation in a community can both influence female participation and foster either inclusive or exclusive atmospheres.

In a study of Facebook usage, Van House (2011) focuses on Judith Butler's theory of performativity to explain behaviour, the idea that "social reproduction and subject formation take place through … reiterations of existing forms" (2011, p. 423). By allowing performativity to shape understanding of social networking sites like Facebook, Van House argues that those designing and creating such interactions can be more accountable for the "sociotechnical assemblages" that these sites help to create and their role in the "configuration of identities" (2011, p. 423). Interestingly, there are a number of publications about "performative experience design" (e.g., Spence, Frohlich, & Andrews, 2013) that have nothing to do with this same concept of performativity, but with theatre and drama. This approach, too, is an important way of understanding how interactions and the shaping of identity might be linked. The overlapping terminology speaks to the pace at which the field rapidly elaborates its understanding following contributions from an increasing number of disciplines.

In all of these examples, HCI researchers have shown great interest in how or whether gender can be an important factor in the social fabric of online communities. Hemphill & Otterbacher (2012) show that in situations where women's reviews are less valued, they may adopt masculine language habits to fit in. Lam et al. (2011) highlight an imbalanced environment at Wikipedia, and Lim & Kwon (2010) find that female readers are less swayed by the resource. Vasilescu et al. (2014) show that environments can dissuade female participation not only through forms of discourse but by how they are constructed and how members are expected to contribute. This work acknowledges imbalances in STEM, but does not turn the mirror on itself to see how it structurally perpetuates

culturally constructed gender concepts. Additionally, these resources explore gendered behaviours in an old-fashioned, binary way. Alternatives to this lens are discussed in Section 2.3.

2.1.4 Gendered Systems

Hemphill & Otterbacher (2012) and Vasilescu et al. (2014) both demonstrate that whether an author is perceived as male or female can have an impact on how useful their contributions are perceived to be by others. On the other hand, researchers have investigated whether attributing a gender to a *system* can also affect people's perceptions. Perceptions of and reactions to gendered systems could be influenced by pop culture: think of Siri in iPhones, Computer in Star Trek, or the more menacing examples of HAL 9000 and the autopilot function in the spaceship in Wall-E (Griggs, 2011). Takayama & Nass (2008) recall an incident in the 1990s when BMW recalled its 5-series cars due to overwhelming complaints about taking GPS directions from a female voice. Human factors researchers Edworthy et al. (2003) show that knowledge about the sex of a speaker has no effect on judgements of perceived urgency, but this has also been challenged more recently by Arrabito (2009), who showed that male speech parameters can result in more accurate actions being taken in a cockpit. Regardless of what types of voices result in what outcomes, terms like "Bitching Betty" and "Nagging Nora" still exist, implying that a "female" system is culturally stereotyped as inferior (see Tupper, 2014; Ward, 2012).

Nass, Moon & Green (1997) found in a study with computer tutors that positive evaluations from male-voiced computers had a larger effect on subjects than female-voiced computers. They also found that female-voiced computers in dominant roles are perceived more negatively than male-voiced computers in the same role, and that male- or female-voice computers were perceived as more informative regarding topics stereotypically attributed to either gender. In a more recent study, Posard (2014) found that participants in an experiment estimated that a computer named *James* would cost significantly more than a computer named *Julie*, despite having rated them the same in terms of both accuracy and performance.

At a time when humans and computers interact on a regular basis, it is important to consider inequalities that exist, that could be created, or that could be exacerbated when gender is implicated in the design of systems that elicit gender-biased reactions. Studies such as Nass et al. (1997) and Posard (2014) demonstrate that gender is an influencing factor in how humans form judgements about technology through the process of interaction.

2.1.5 Theories of Competence and Motivation

Gender-related HCI research often discusses feelings of adequacy and competence among female users of technology. Following the Burnett (2011) study on tinkering and other behaviours in spreadsheet software, Beckwith et al. (2006) tried to determine whether encouraging tinkering could increase females' feelings of self-efficacy, since males had high confidence and tinkered freely. They found that males' tinkering habits actually could be counterproductive to their effectiveness in the task at hand, revealing some nuance in what "tinkering" can mean and how it factors in to or detracts from a workflow. Turkle & Papert (1991) use "bricolage" as an alternative working style that may serve females better – however, as Beckwith et al. (2006) point out, the French meaning of this word is essentially "DIY", the usage of which in France is equivalent to the masculine usage of the word "tinkering" in North America. DIY has emerged as a way to "cultivate the legitimacy of masculine skilled labour" in the home in an era where men no longer have the "aura of pre-industrial vocational masculinity" (Bardzell, Gross, et al., 2011, p. 373). Encouraging tinkering is essentially encouraging women to behave in a masculine way in order to enjoy the perceived benefits that men have, without really asking if this is what women need to succeed.

Rather than any one behaviour (such as tinkering) being correlated with perceptions of self-efficacy, Huffman et al. (2013) specifically linked the concept of gender to feelings of confidence in using technology. Their research corroborated previous evidence that males report higher ratings of technology self-efficacy than females. More significantly, the authors found that "gender roles, specifically masculinity, are more predictive of higher ratings than the designation of biological sex" (2013, p. 1783). The findings accept the concept of gender as a social construct rather than being founded on biological sex: their study shows that "although biological sex can be used to show a difference between male's and female's technology self-efficacy, this singularity is mostly explained by gender roles and not by biology" (2013, p. 1783). Even when students arrive in universities with plenty of previous experience with computers, it is masculinity that continues to be a "stronger predictive factor of technology self-efficacy" than any previous preparation or support (2013, p. 1783). The authors propose that because men have "more socially acceptable" interactions with technology, this results in forming masculine gender roles that include the various "skills, motives and beliefs" necessary to be successful in technological tasks (2013, p. 1783).

Self-determination theory (SDT) is the concept that autonomy, competence, and relatedness are three human needs that need to be satisfied for enhanced motivation and well-being (Ryan &

Deci, 2000). This concept is used at length in HCI research and Przybylski et al. (2010) propose a motivational model for video game engagement based on the idea that engaging games satisfy basic psychological needs of competence, autonomy, and relatedness. Psychologist Albert Bandura (1977) describes self-efficacy as one's belief in one's ability to succeed in specific situations, which can alternately be described as a belief in one's own competence: with more confidence due to higher perception of competence comes belief in a higher likelihood of success.

Self-determination theory does not question whether the same definitions of autonomy, competence, and relatedness are universally motivating for all genders. The conceptual frameworks that guide the organization of official knowledge in our culture "tend to reflect the experience and interests of men, particularly racially and economically privileged men" (Sprague & Hayes, 2000, p. 674). From the standpoint of the privileged, our understanding is often organized such that it is "(1) objectifying, (2) decontextualized, and (3) logically dichotomous" (2000, p. 674). While *competence* might lend itself well to SDT without much tailoring for different groups, *autonomy* suffers from decontextualization which creates a model for intrinsic motivation that is not inclusive.

Kant once said: "Autonomy of the will is the property the will has of being a law unto itself (independently of every property belonging to the objects of volition)" (1785, p. 108). The post-enlightenment rational self is "atomistic" and "independent" in that it is abstracted from the social context in which it is embedded, it has fallen to objectification; this is in opposition to a feminist concept of the self that is both socially constructed and relational (Heikes, 2010, p. 65). The autonomous free agent implicitly corresponds to a "masculine ideal", an idea of being self-made and that this independence is valuable unto itself. If a masculine concept of autonomy values independence at the expense of social relationships and care, a feminist concept of the self accepts dependency and revalues it, rather than allowing it to devalue the self. Some feminists dismiss autonomy as "an androcentric relic of modernity", whereas others assert women's need for self-determination (Willett et al., 2015). Autonomy is important in the feminist discourse, in fact, some have recontextualized it as a distinctly feminine value (Meyers, 2002). Lugones & Spelman (1983) equate autonomy with having a voice, which is "integral to leading a life rather than being lead through it" and that "being silenced in one's own account of one's life is a kind of amputation that signals oppression" (1983, p. 19). The feminist reconceptualization of autonomy has been called *relational autonomy*. It is relational in that it denies the idea that autonomy requires self-sufficiency:

the autonomous self is not atomistic, it is socially and historically embedded, and shaped by factors such as race and class (Stoljar, 2014).

This concept of relational autonomy is distinct from *relatedness* in self-determination theory in that the latter describes how "social environments can facilitate or forestall intrinsic motivation by supporting versus thwarting people's innate psychological needs" (Ryan & Deci, 2000, p. 71). The former references an individual's own socially-defined situation, and the latter refers to connectedness in the context of intrinsic motivation (i.e., the activity that the person is/isn't motivated to do). Ultimately, the point to consider here is that the feelings of autonomy that may be traditionally accepted as essential for an activity to be intrinsically motivating are not one-size-fits-all, due to the idea that historically masculine ways of knowing have failed to account for different ways that people might feel autonomous and engaged when using technology. To some, dependency may feel utterly discouraging, and for others it is a positive, motivating feeling.

2.1.6 Attempts to Reconcile Gaps in Gender HCI Work

There are numerous opportunities to address some of the issues raised in Sections 2.1.1 – 2.1.5. In this section I will discuss some of the ways in which researchers in HCI or other areas of STEM have acknowledged and worked to change emerging negative trends in the areas of: software, hardware, the internet, design of systems, and encouraging involvement and learning among women or other marginalized groups.

Burnett et al. (2011) brought light to the fact that software designed with 'gender agnosticism' in mind risks glossing over differences during design decisions and putting one group at a performance disadvantage compared to another. Brunner, Bennett & Honey (qtd. in Cassell & Jenkins, 1998) surveyed a number of adults about their technological fantasies, and they interviewed children to see what futuristic technological inventions they could imagine. They found that both women and girls were more likely to be concerned with "how new technologies can fit into the social and environmental surroundings", whereas men and boys are "much more likely to be preoccupied with doing things faster, more powerfully, and more efficiently regardless of social and environmental consequences" (1998, p. 77). In short, women were more likely to see technology as a tool that can assist in the current environment, whereas men may see it more as an interesting artefact in its own right. Interestingly, these perceptions by both boys and girls seem to be co-opted by designers as well when designing for either audience. Huff & Cooper (1987) asked 43 educators with programming experience to design software for "boys", "girls", or "students". They found that the

programs made for girls were “learning tools” whereas programs for boys were “games”, and when they were just designing for students they were “games” as well. The implication of this study is the finding that the differences in what was produced were a function of designers’ *expectations* about their target audience, and this resulted in stereotyped software (1987, p. 519). The software itself is not biased and the designer may not be explicitly biased either, but what they produce acts as a vehicle for the expectations or stereotypes that may be held and that may implicitly guide design decisions. Designers should be aware of their own social and cultural context and how this may influence their approach to design problems. Maggie Johnson, Director of Education and University Relations at Google (2015), suggests tackling this problem right at the source, by getting underrepresented groups to be the designers. She writes that in 2013, only 14% of computer science graduates were women and 20% were underrepresented minorities. This lack of representation needs to be addressed because “[t]he workforce that creates technology should be representative of the people who use it, or there will be an inherent bias in design and interfaces” (2015).

There are tangible impacts on women resulting from the design of hardware failing to effectively design for a universal by failing to consider that women may be included in the group of potential users. Margolis (2002) shows that some early voice-recognition systems were calibrated to typical male voices which left women literally unheard, such as a video conferencing system that focused automatically on the speaker: if women could not be heard they could not be seen, and were excluded from participating. An alarming example from the automotive industry recalls a group of male engineers who fit the first generation of car airbags to average adult male bodies, which resulted in avoidable deaths of women and children (2002, p. 3). The HCI community should not only be aware of the potential audience for their work, but also of the gendered discourses that could implicitly influence the things they design. In the software world this arose in the problem of educational software being designed differently for both boys and girls. In the world of product design, the worst reaction to “considering” gender differences is the unfortunate tendency to “shrink it and pink it” (Tischler, 2009). In the design of new technologies, gender is often either a factor that is not explicitly considered, or it is considered in the sense of idealized models for products made “for men” or “for women” (Dray, Peer, et al., 2013). The *Perspectives on Gender and Product Design* workshop held at CHI in 2014 addressed this issue within a specifically HCI context. It served as a checkpoint to review how the discussion is framed, and the current status of gender-related knowledge and practice in the field. The workshop’s organizers also explored whether gender-

sensitive or gender-neutral products can exist or if these are a desirable goal, and reflected on gender-sensitive design processes and practices (Dray et al., 2014).

Hemphill & Otterbacher's (2012) extensive analysis of the language patterns on IMDb revealed differences between genders and that some female reviewers may change their language over time. The theoretical basis for this study utilizes a binary opposition between genders with "feminine" characteristics on the one side and "masculine" on the other. This dichotomy is culturally constructed and is perceived differently depending on cultures, time, and context. In their study, the characteristics of women's communication was that they posted "fewer and shorter messages, are more polite, and write in a manner that aligns themselves with others" whereas men "tend to dominate conversations … and are more likely to post aggressive messages" (2012, p. 2). They also write that "while women seek and offer social support …, men seek to achieve and maintain status by showcasing their knowledge to others" (2012, p. 2). The Malagasy of Madagascar "attribute indirect, ornate, and respectful speech that avoids confrontation to men; women are held to be overly direct and incapable of repressing their excitability and anger" (Cassell, 2002, pp. 1–2). The discussion around what constitutes "feminine" or "masculine" behaviour drifts dangerously close to essentialism and in inherently global situations such as the discourse on IMDb, researchers must be sensitive to the potential for misreading language patterns that come from other cultures. A workshop at the Computer Supported Co-operative Work (CSCW) conference in 2012 called *Feminism and Social Media Research* intended to "identify ways to improve social media research by leveraging feminist approaches" (2014, p. 319). The authors cite the *Handbook of Feminist Research* (Hesse-Biber, 2012) with some starting points: including more diverse researchers, constant opportunities for epistemological criticism by different inquirers, doing research from the perspective of those marginalized stakeholders potentially impacted by the research, and a plurality of theoretical frameworks and models to understand scientific phenomena (Hemphill et al., 2014, p. 320).

Lam et al. (2011)'s exploration of the gender imbalance on Wikipedia noted that the Wikimedia foundation set a goal to increase the number of female editors to 25% by 2015, up from 13% in 2009 (2011, p. 1). Vasilescu et al. (2014) showed us that certain structures designed to incentivize participation and contributions may actually dissuade women from engaging, and Wikipedia is no exception. The HCI community should be sensitive to the successes and failures of previous efforts such as these to motivate participation, and to work on understanding what discourse models lead to more inclusivity. Efforts to increase numbers will not succeed if the structure of the

community itself never changes. Margolis et al. (2002) outline a strategy that, in five years, increased the proportion of women entering Carnegie Mellon's computer science program from 7% to 42% (2002, p. 6). The results are encouraging and show that structural changes are needed to correct the trajectory of low female involvement: the changes made were both formal and informal, and ranged from "the design of the curriculum to the education of teaching assistants to the way the university thinks about admissions" (2002, p. 6).

Huffman et al. (2013) showed how the construct of masculinity, rather than biological sex, was a more powerful predictor of a person's technology self-efficacy. The authors suggest that universities adjust their teaching methods and intervention strategies to accommodate for gender differences, and that this will level the playing field for future opportunities for advancement (2013, p. 1784). The disillusionment and divide that contributes to the marginalization and self-exclusion of women and minorities in STEM begins far earlier than in university, and this is likely too late to make a significant impact. Nicholson et al. (1998) found that in a first grade classroom, girls working in a mixed-gender group on a computer activity were more likely to be laughed at, criticized, or have their competence questioned than when they worked alone or in all-girl groups. Male students also frequently interrupted the girls, but the reverse was not true. More than a level playing field is needed: the conditions need to change for people to get there.

Another important note is to remain aware that this is more than a problem for women or femininity alone. As Margolis & Fisher (2002) point out, "the shortage of people of color in the computing profession is even more dire than the shortage of women" (2002, p. 10). Approximately 5% of graduates from computer science bachelor programs are disabled, but this population shrinks with further specialization with only 0.8% of graduates from computer science PhD programs from 2000-2004 having disabilities (Burgstahler & Ladner, 2006). Furthermore, statistics regarding disability are often unclear, can be underreported, and disability can happen at any point in a person's lifetime. There is a need for HCI to move beyond second-wave feminism and to adopt more diverse ways of thinking about its "users", new technologies, and HCI researchers themselves, so that work such as Stepulevage (2001) or Landström (2007) are no longer an exception to the norm.

In summary, it is clear that the HCI community is to some degree aware of some of the issues it faces in terms of designing inclusive software, or preventing the perpetuation of stereotypes in hardware or product design. In more abstract areas that do not have a direct application in the design process, primarily *approaches to conducting research about gender* and *investigating ways to reduce*

barriers to entry, solutions are lacking. Researchers must be aware that studying gendered behaviour based on perceived differences contributes to essentialist models of user behaviour that are not universal nor generalizable, since what could be perceived as "masculine" or "feminine" is culturally dependent. Additionally, attempts to empower the marginalized should not happen after the marginalization has happened – it needs to be preventative. As I will discuss in Section 2.4, schemas of what a "scientist" and other heavily stereotyped words represent are sedimented in childhood. The ways that people in STEM disciplines talk about the people affected by their work has a role to play in how others grow to identify with or feel empowered by that persona. By investigating whether the language we use in HCI contributes to these stereotypical images causing disillusionment, I aim to open up new opportunities to prevent or address the problem before it grows.

2.1.7 Reflection and Next Steps

Through Section 2.1 I have demonstrated how existing work relevant to feminism in the context of the HCI domain has cleared the way for more pluralistic scholarship. Early work on the implications of gender in the design of new technologies was discussed in Sections 2.1.1, 2.1.2 and 2.1.4. In Section 2.1.3, I described how the design of online communities can have social implications that extend beyond the computer and affect the participation and enjoyment of female users. Feminist theory reveals nuance in the way HCI researchers and practitioners approach understanding motivation in HCI work, as discussed in Section 2.1.5. Steps towards addressing some of the issues highlighted in the first half of this chapter were overviewed in Section 2.1.6. While this coverage is quite diverse, there is ample room for future research, which is needed to avoid perpetuating some of the methodological and ideological issues implicated in Section 2.1.6. In the following sections I describe the specific problem facing HCI that I will study.

2.2 Is the "User" in HCI Literature a Neutral or Gendered Term?

The HCI community is always expanding and collaborating, looking for new ways to investigate, analyze and think about human-computer interactions. A large part of understanding of these interactions relies on concepts of what is known, or what they think they know, about the human component and how it is situated in the world. Identity and the self lay the foundations for the conceptual models that contribute to people's semiotic understanding of the world. The prevailing strategy for representing knowledge, however, arose out of the rationalistic tradition, and has been "to categorize the world into domains of knowledge … and then to enumerate facts about the domain and

relationships between them" (Suchman, 1987, p. 43). This process divorces the knowledge from individuals and their particular circumstances. The HCI community frequently makes use of the word "user" in papers, reports and conversations when describing people who use or interact with technology, but previous research has not been able to shed sufficient light on who the user might be, how this term is perceived, or what some implications of the term are.

"User" is considered a fairly benign term, however, "[a]s the use of a new technology changes human practices, our ways of speaking about that technology change our language and our understanding. This new way of speaking in turn creates changes in the world we construct" (Winograd, 1986, p. 6). In HCI, "interaction" implies a presence of an autonomous human agent. Yet the word "user" implies "dependency, addiction, and, ultimately, objectification where one is conscripted into the logic of whatever is being used, be it a technology, drug, etc." (Coleman, 2011, p. 38). With the development of gestural input techniques (such as *tap*, *pinch*, *swipe*, *flick*), for example, non-verbal communication is naturalized through so-called "intuitive" and "natural" inputs. The identity of the human actor is rescinded: they must perform the role of the user to be acknowledged by the system, and its influence over the person's behaviour thus extends from within the interface to how they move in space lest their gestural cues go unnoticed.

Naturalizing certain actions has not resulted in complete silencing of the user as an individual: "[i]t is not media technologies that reposition us, but rather how we engage them in a symbiotic relationship" (Coleman, 2011, p. 38). Psychologist Albert Bandura explains that human agency and sense of self is an interactive phenomenon, "a product of a reciprocal interplay of intrapersonal, behavioral, and environmental determinants" (2006, p. 165). "Affordance" describes the set of possibilities for action that an environment presents to its users (Morrison, 2013, p. 117). By foregrounding certain affordances and creating constraints to encourage particular behaviours, the designer manipulates the available means of persuasion and engages in design as a rhetorical practice. When a designer constrains which actions may be performed, they limit the acts the social agent may use to "constitute social reality, through language, gesture, and all manner of symbolic social sign" (Butler, 1988, p. 1). It follows from Butler's theory of performative acts and gender constitution that the concept of the "user" is an implicitly exclusive term, acting as a "false universal" term in the same culture where "the false universal of 'man' has for the most part been presupposed as coextensive with humanness itself" (1988, p. 4). The prescription of acceptable, preferred or target behaviours and their naturalization within scientific literature has a normative effect in which those outside the ideal

embodiment of "user" may be subject to alienation and exclusivity, for "embodiment is not thinkable without relation to a norm, or a set of norms" (Butler, 2004, p. 28). The new subdiscipline of Feminist HCI is concerned with the design and evaluation of interactive systems that are "imbued with sensitivity to the central commitments of feminism—agency, fulfillment, identity and the self, equity, empowerment, diversity, and social justice" (Bardzell, 2010, p. 1301). In addition, it assists with revealing unspoken values within HCI's dominant research and design paradigms, underpinning the development of new approaches and methods.

In Section 2.1 I described how existing research and design paradigms failed to account for a number of issues leading to marginalization of women and other minority groups in areas directly pertaining to or associated with HCI. Feminist theory and other philosophical discussions have raised the idea that "user" might present potentially problematic connotations meriting further investigation. In Section 2.3 I will review literature around issues leading up to the "user" debate, to assist with navigating the specific problem in this research by exploring whether using such neutral language has been seen as advisable in past scenarios and why or why not. In Section 2.4 I return to the original research problem to lay the groundwork for investigating the question of gendered language in HCI.

2.3 Perspectives on Equality and Intentionally Neutral Language

The research in this thesis is specifically concerned with the language used to describe people in HCI literature and whether these words are perceived as gender-neutral. While researchers may intentionally use neutral language in the interest of encouraging equality, whether this is the best option is not generally called into question. This very issue has been a topic of discussion in feminist rhetoric and opinions have shifted throughout the three main waves of the feminist movement. Without an overview of the feminist theory relevant to this study, it would be difficult to contextualize my discussion of the use of ambiguous words like "user" in HCI research papers. In this section, I will also show how the different approaches to the debate over universalizing language may inform current researchers' choices in their writing and approach to gender-related problems.

Feminist critiques of science span decades, with prominent contributions coinciding with feminism's first wave, the 19th- and 20th-century movements focused on voting rights and other legal challenges. In 1938, Virginia Woolf wrote "[s]cience, it would seem, is not sexless; she is a man, a father, and infected too" (1966, p. 127). First-wave women's rights movements sought to garner the same legal rights enjoyed by men, but did not extend the same rights to all women and did not

dismantle the larger power structures that maintained unfavourable social and power relationships between these groups. In their book on gender and computer games, Cassell and Jenkins recall a 1997 episode of *Saturday Night Live* with a pseudocommercial advertising "Chess for Girls," asking why such gendering of chess can seem so absurd yet similar efforts to bring computer games to girls has generally strong support (1998, p. 3). Attempts to invite women into technology risk being informed by this outdated approach, extending things that already exist for men to women by relabeling them or marketing them in a way that seems more "feminine" under existing cultural schema. This method also presumes a meaning of "gender" characterized by sexual difference rather than systems of relations. When movements aim for a redistribution of power with the existing power structure intact, rather than challenging the "forms and nature of political life," they resubject themselves to a "phallocratic order" and may end up constituting a more "concealed exploitation of women" (Irigaray, 1987, p. 81). This positions women as the "other side" of the gender that holds a monopoly on value. The practical application here is to ask whether something is "good" for women in its own right or whether it simply awards some privilege elevating their status closer to that of men.

Second-wave feminism worked within existing power structures while demanding access for women: it was based on "an empiricist view of science as (gender) neutral" and denied any fundamental sex differences between genders (Wajcman, 1991, p. 8). Rather than designing for gender differences, the tendency may be to design for a universal group. Washing over these differences falls into the same "sexual indifference that underlies the truth of any science, the logic of every discourse" (Irigaray, 1987, p. 69). Thus it is not sufficient to assume that leveling the playing field would create the conditions for success. A later, more radical approach to science sought to understand what cultural and social factors were preventing equality. It also sought to incorporate women's values and experiences, however these "values" that are attributed to women are the result of the historically ongoing subordination of women, a belief that they have a kind of unchanging nature or inherent set of values. There cannot be a new approach to science based on the standpoint of all women's experience, as this experience is fractured across race, class, sexuality, and culture.

HCI researchers have in their own work assumed that in conditions of equality, all genders will have positive experiences. In their analysis of IMDb, Hemphill & Otterbacher (2012) hypothesized that: first, prolific female reviewers would write more like males over time and therefore achieve more prestige; and second, prolific female reviewers would correspondingly see improved utility ratings over time (2012, p. 3). By acting more like males, and writing "in a manner

that appeals to the IMDb audience", women were expected to come out as equals. In reality, they found that while women did adapt their language use to sound more like male reviewers over time, their scores never caught up to their male counterparts (2012, p. 7). At the end of the day, female contributors are changing themselves to be valued within a system that most highly prizes male contribution. Czerwinski et al. (2002) noticed that systems designed for navigation through virtual worlds saw significantly better performance by males than females. These systems were designed for equality rather than difference, and the authors asked "what if females can navigate as effectively as males … but simply have not been provided with the proper display parameters that best support their navigation strategies" (2002, p. 195)? The improvements made to accommodate all potential users of their system did not result in harm to the performance of the males.

The first and second waves of feminism introduced to HCI a philosophy of difference and then of equality. The third responded to what some women perceived as failures of the second wave: its essentialist tendencies, its lack of intersectionality (focusing on white, privileged women), and perception of gender as dichotomous. For example, Kotamraju (2011) pays attention to calls for feminist approaches to research in HCI, and asks *which* kind *of feminism*? In their paper, the author focuses on three behaviours in usability practice that can be read as gendered, and shows that within a liberal feminist framework these behaviours seem anti-feminist, but in many other frameworks (relational/care-giving, sex-positive, multicultural, post-colonial and Third Wave) they appear feminist. Due to this intersectionality, this third wave has introduced some more complex concepts that could be applied in research and design. As discussed in Section 2.1.3, *performativity* has been adopted in HCI with a number of nuanced interpretations, Van House (2011) reflects on performativity in the context of social networking sites. The concept of performativity as defined by Butler (1990) remains central to current feminist theorizing about identity: that "social reproduction and subject formation take place through (largely unquestioned, but not necessarily faithful) reiterations of existing forms" (Van House, 2011, p. 423). Suchman (2007) asks readers to think about interactions with technology in the same way that Butler asks them to think about interactions in day-to-day life. She writes that the way in which Butler shows how gendered bodies materialize over time through reiterating norms is "suggestive for a view of technology construction as a process of materialization through a reiteration of forms" (2007, p. 9). Feminist theory rethinks what constitutes the subject and the object, the active and passive, in an interaction. It thus challenges the subject/object framing in and evaluation of HCI experimentation, which often takes place through the lens of the rationalist and traditionally masculine scientific method. Although traditional science has

emphasized "objectivity and detachment as virtues in a researcher or a scientist," feminism helps to address the perspective an *individual* brings to the table and how this may be embedded in their work (Muller, 2011, p. 448).

There remains contention over whether language as a structure can be inherently oppressive or whether it is only such in its applications (see Butler, 1990, p. 26), but common threads remain: first, that language is an important site for feminist scholarship to break away from existing oppressive structures; and second, that universality inhibits progress in "tearing down the phallogocentric tower which insists on singular meaning and universal truth" in order to replace it with a "plurality of meaning and a concept of truth which incorporates multiplicity" (Carnegie, 1998, p. 5). This is one reason why "identification" is not a key term that works as well for feminist rhetoric, because it is seen as a rhetorical tool used for totalization. Writing about such terms in feminist rhetoric, Carnegie (1998) argues that the accepted key terms of rhetoric (ethos, logos, pathos, analogy, persuasion, identification, and so on) maintain existing power relations and have, until recently, been established by those who "benefit from the structures of a patriarchal society" (1998, p. 2). It is from this perspective that Carnegie argues that rhetorical identification is totalizing, as it erases rather than embraces differences. On this note, she proposes "experience" as a more appropriate key term for feminist rhetoric (1998, p. 10). Understanding people's experiences is an important component of HCI research. User-centered design asks practitioners to "draw on the user's knowledge of the world" (Abras, Maloney-Krichmar, & Preece, 2004, p. 2). Abras et al. (2004) name the major advantage of user-centered design as "a deeper understanding of the psychological, organizational, social and ergonomic factors that affect the use of computer technology" (2004, p. 10).

As an attempt to bridge the gap between these concerns and HCI research, as well as an attempt to frame future research attempts, Bardzell & Bardzell (2011) examined existing methods of feminist social science and proposed an outline for an actual methodology for what they call Feminist HCI. In this new discipline, there are four precursory assumptions to account for: first, that there "is no feminism but rather feminisms"; second, that feminism is not the "investigation of gender"; third, that feminist theory has no preference of qualitative approaches over quantitative; and fourth, that there is no concept of a "conflict" between feminism and HCI, in that HCI is *not* inherently anti-feminist nor sexist, and feminism "does not operate in a privileged space over and above HCI" (2011, p. 677). It does assume a relatively high level of understanding of some of complexities of feminist

theory to which the HCI community may not be sensitive, which could limit its benefit. It also assumes agreement about what "feminism" even entails, and does not make explicit the schools of feminist thought to which it subscribes. Feminist discourse runs the risk of universalizing itself, as it has often relied on the term "woman" to account for some shared cultural experience which "in its universal status, provides a false ontological promise of eventual political solidarity" (Butler, 1988, p. 4). Recognizing the effects of such terms is as important inside feminism as it is outside. In Western culture, Butler argues that the term "man" has for the most part been "presupposed as coextensive with humanness itself" (1988, p. 4). Such terms, coined "false universals", fail to acknowledge the presence, influence, and oppression of those absorbed into its self-definition: they use circular logic that "reduces differences to sameness and universalizes and totalizes by substituting part for the whole" (Carnegie, 1998, p. 7). Through identification, bodies become "sexually undifferentiated" and cannot be interrogated outside of a model where they are unwoven and shared with "man" (1998, p. 7). These false universals are the kinds of issues that HCI researchers may encounter when adopting feminist HCI practices without being provided the context around feminism's own discursive issues.

There is a trend in the design of human-computer interactions and experiences to make them *intuitive* and *natural*. Controlled experimental studies show what kinds of gestures and interactions are, scientifically speaking, statistically significantly, the most *intuitive*. These models are being constructed within the objective and abstracted scientific method. Feminism looks to the phenomenological theory of "acts" to understand how "social agents constitute reality, through language, gesture, and all manner of symbolic social sign" (Butler, 1988, p. 1). These acts constitute identity; they are the mundane ways in which sex-related constructs are "produced, re-produced, and maintained within the field of, bodies" (1988, p. 5). In a 1991 article anticipating ubiquitous computing, Weiser wrote that the most profound technologies are "those that disappear … they weave themselves into the fabric of everyday life until they are indistinguishable from it (1991, p. 1). Building a model of perceived "natural" behaviours within the lens of traditional science creates a new totalizing structure that disappears and delimits the acts constituting a successful interaction with some technology. Repetition and reuse of these socially-established meanings is the "mundane and ritualized form of their legitimation" (Butler, 1988, p. 6). With this new method of identification, bodies are once again undifferentiated and consubstantial under a falsely universal set of actions constituting social reality and identity. Thus it is essential for work in HCI going forward to be aware of the construction of this new set of norms and the implications on an individual's sense of self when they are asked to align their body with a system reiterating a gendered identity due to interaction

models constructed from a masculine perspective. Indeed, reworking the norms by which bodies are experienced is crucial not only for gender or disability politics, but also to the intersex and transgendered movements as they "contest forcibly imposed ideals of what bodies ought to be like" (Butler, 2004, p. 28).

2.4 Solving a Feminist Language Problem in HCI

In the preceding sections I have shown the ways in which feminist thinking is essential to inclusive HCI practice, and how feminist research methods can both assist with development of new technologies and uncover problematic implications of existing ones. Judy Wajcman critiques much of feminist epistemology, saying that "however philosophically sophisticated … [it] is misdirected" (1991, p. 12). She argues that the more philosophically oriented work fails to do much in terms of dealing with social practices, instead getting stuck on abstract ideas and "questions of recruitment to science" (1991, p. 12). Language is a tool of inquiry often used successfully in other domains such that we can learn from its applications in feminist research elsewhere and apply these ideas to HCI for richer analysis.

Feminist theorists strive to reconceive both the gendered subject and re-conceptualize the object by looking for new grounds to understand the body outside of the mind/body split, and to understand our relations to the material other than the subject/object distinction (Suchman, 2007, p. 9). Such boundaries as those between human and machine are constructed and are important cognitively to be able to understand situations and experiences, but these boundaries are not value-free because they are positioned by an agent rather than naturally occurring (2007, p. 11). As a result, though no one is necessarily responsible for the world as it stands, everyone is accountable for it is reiterated and reinforced through the everyday practices that we do have a role in shaping (Barad qtd. in Suchman, 2007, p. 11). These practices encompass words spoken and actions taken, and often take the form of unconscious habits. Being more accountable is not necessarily pinpointing the source of a problem but rather being aware if there are issues at all: it is making an effort to understand the social and cultural structures that underpin a discipline, and working to find what kind of implications those may have; this responsibility is met through "ongoing practical, critical, and generative acts of engagement" (2007, p. 11). The goal of this work is to engage with the language of HCI as one component of its formative structures, and to become aware of any distributions that may or may not be engendered through particular practices, particularly any non-inclusive practices that may be encouraged by using ambiguous language.

For the research presented in this thesis, drawings were analyzed as a result of linguistic prompts drawn from HCI vocabulary. The collection and analysis of drawings is a popular and well-documented methodology in psychology (for example, see Brafman, 2012), a field with which HCI is closely aligned. Türkcan (2013) shows how drawings can be analyzed semiotically to diversify the techniques used to give meaning to images produced by participants. Such images show us "not the thing actually seen but its representations in human consciousness", which is informed through history, culture, and codes (2013, p. 601). By looking at drawings, researchers can find the processes informed by particular circumstances that coalesce to "render meaning shared and action accountably rational" instead of looking for a structure that is invariant across situations which is a more traditionally rationalistic approach (Suchman, 1987, p. 67). Past research demonstrating gender differences in drawings among children has been interpreted from an adult perspective and has also taken the images at face value, without understanding art-making processes, agency or artistic intentions, but Wright (2014), shows how drawings can be solicited with the goal of understanding "the relationships that the artists have with the object of their representations" as well as any "gendered productions" or the discourses that would account for these drawings (2014, p. 393).

Drawing also assists with conceptualizing the relational as opposed to static nature of the words used in HCI. Language can contain more stereotypes than the things people draw, so examining images can help to uncover meanings that would not necessarily be conveyed through words (Sontag, 2001). Chambers' (1983) influential *Draw-A-Scientist Test* (DAST) was a pioneering effort specifically meant to uncover stereotypes. Administered to children at different ages and over a period of 11 years, the researchers were able to identify trends in how children perceived the "scientist" as well as the ages at which certain symbolic imagery begins to appear. This research builds on Chambers by using a similar drawing method to investigate whether common "people" words used in HCI literature are gendered.

There is no better time to turn the lens on HCI discourse than the present. While there is an obvious commitment to a feminist research agenda by authors like Bardzell (2010), Cassell (2002) and Dray (2014), research showing a trend map of discussion topics over the past 5 years of CHI papers demonstrates that discussions to do with women were more often linked to their role in public or community life in the context of HCI projects rather than being linked to gender-related advocacy, and these discussions reached a peak over three years ago before starting to decline (Padilla, Methven, Corne, & Chantler, 2014).

The words and ways used to refer to others can have a larger impact than what might be realized or intended. Hurtig & Pichevin (in Cassell & Jenkins, 1998) showed that when study participants were asked to categorize people in a photo of "successful executives" they chose to sort them by gender, but when others were asked to categorize the same group but referred to them as a "group of friends", the gender categories did not come up. Their conclusion was that gender only becomes a variable when gender is being contended, in other words, when socially constructed categories are evoked that have to do with "what we expect of men and women" – the cultural construct being the binary opposition between masculine and feminine (1998, p. 6). Since the goal of feminist rhetoric is to break the social relationship where women are subservient to men (Carnegie, 1998, p. 5), it is important to understand whether gender is arising as a way of distinguishing between figures such as "users" or "designers" or other words used to categorize "people" in the field.

There are real effects of a failure to acknowledge patterns like the culturally constructed gender dichotomy influencing the way HCI researchers think or talk about people. Huff & Cooper (1987) were able to demonstrate that software designers created different products depending on whether they were asked to design for "boys", "girls", or "students". The programs for girls were classifiable as "learning tools" whereas those created for "boys" or "students" were like "games" – these differences occurred due to the designers' expectations about their potential users and resulted in gender stereotyped software (1987, p. 519). The authors conclude that it is therefore the designers' expectations and stereotypes at the root of gender-biased software.

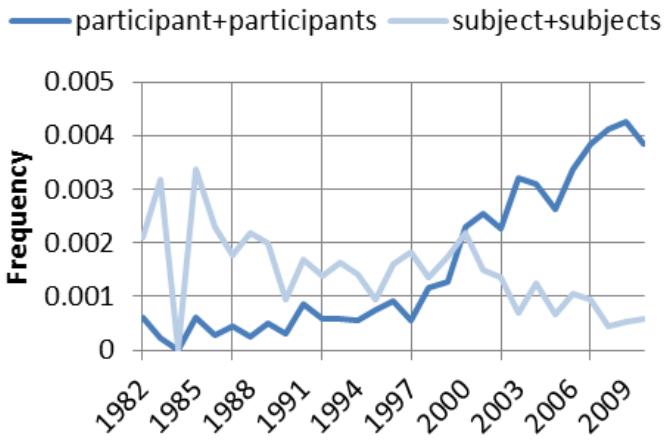

Figure 2-1: Word Frequency of "Participant" and "Subject" in CHI Papers from 1982-2011

If there was indeed a phenomenon in which the language of HCI had gendered implications, the community would be wary of how they use this language given Huff & Cooper's troubling findings. The psychology community had similar concerns about the effects of its vocabulary and this instigated a large shift in the way it spoke of its research participants. In 1995, Chalmers (1999) suggested that medical researchers follow the British Psychological Society's example by using "participants" rather than "subjects" after noting that "psychologists owe a debt to those who agree to take part in their studies" and that they "therefore deserve to be treated with the highest standards of consideration and respect". Boynton (1998) advised that using "subjects" actually conflicted with research policy, as it failed to regard people as active or consenting participants in research. What Chalmers had initially realized was demeaning terminology was then framed as a systematic and ethical imperative for change, and language policies changed accordingly. Psychology plays a large role in HCI, it can be seen through the n-gram analysis shown in Figure 2-1 of CHI papers published since the first SIGCHI conference in 1982 that in or around 1998 the use of the word "subject" was supplanted by "participant." In accordance with this, "participants" will be used to refer to any research participants in this thesis. What I suggest is that there may be other words in the lexicon, beginning with the word "user", that can act similarly negatively as did "subject". The previous work in feminist theory clearly lays an ethical foundation to motivate linguistic change, thus the methods for unearthing stereotypes in the language of HCI outlined here should be used to identify gendered phenomena in HCI vocabulary.

2.5 Chapter Summary

In this chapter I discussed areas related to HCI in which researchers and designers attempted to be deliberately agnostic about gender (Section 2.1.1) and when they purposely designed with gender in mind (Section 2.1.2). I reviewed past work about gendered behaviour and language in online scenarios (Section 2.1.3), and briefly reviewed some implications of actual systems designed with gendered attributes (Section 2.1.4). I explored how some theories within HCI as a discipline might be re-interpreted through an alternative, feminist lens to understand why some designs may not appeal equally to all people (Section 2.1.5). There have been attempts to acknowledge some of these issues which I discuss in Section 2.1.6. I called the ambiguous "people words" in HCI literature such as "user" into question in Section 2.2, and discussed the position on this neutral language from a feminist perspective in Section 2.3. Finally, in Section 2.4 I outlined the specific approach to how I will investigate the question of gendered language in HCI further.

Chapter 3
Study Method and Findings

In Chapter 2, I identified that traditional approaches in HCI were not necessarily effective or appropriate to address feminist issues in the field, as well as how analysis of language and drawings could be used to broaden the scope of research and contribute as new ways to critically engage with the growing community. In this chapter, I describe the design of a Mechanical Turk study used to gather images associated with the five most common words to describe people used in HCI papers (user, participant, person, designer, researcher). Results of this study are described in the latter half of this chapter, with the results further contextualized through qualitative analysis in Chapter 4. Broader discussion of these results and their implications is found in Chapter 5.

Feminist rhetoric seeks to break social relationships wherein women are marginalized through sedimentation of socially constructed categories (i.e., their gender). Gender can be an influencing factor in software design decisions leading to different outcomes depending on the designer's expectations: the way we think or talk about end users influences what is designed for them (i.e. Huff & Cooper, 1987). When asked to talk about "successful executives" shown in a photo, people categorized them by gender, but when told to talk about them as a "group of friends", gender was no longer a category (Cassell & Jenkins, 1998). Do the words that we use to talk about people in HCI, specifically "user", "participant", "person", "designer" and "researcher", elicit the same tendency to stereotype by gender (as in the executives) or are they generic (as in a group of friends)? Words that are meant to be neutral but actually represent a masculine ideal are false universals, and this chapter describes the study method used to evaluate whether these HCI "people words" act as false universals.

Burnett et al. (2011) show how certain "gender agnostic" tools make assumptions that all users will exhibit masculine behaviours, such as tinkering to learn new software features, can inhibit the speed at which female users adapt to the same tool. Tan et al. (2003) show how female performance in 3D virtual environments can be improved at no detriment to male performance by adding simple visual cues. In both cases, designing "gender agnostically" resulted in designing for a false universal, and it marginalized women. The words the HCI community uses most often to talk about people appear to be agnostic as well, but until now there has been no way to determine whether this is the case. Chambers (1983) used the *Draw-A-Scientist Test* with the control "draw a person" to

investigate how stereotypes emerge in association to the language we use. Since then, similar versions of Chambers' test have been used to demonstrate gendered, racial, even socioeconomic stereotypes associated with the "scientist", but there have been no extensions to understand how other "people" words are perceived in other fields. Wajcman (1991) argues that the language and symbolism of technology, like science, is also masculine. A key goal of this research is to gain an understanding of how or whether language used to talk about people follows patterns identified by feminist theorists whereby gender-neutral or "universal" terms lead to an assumption that the subject is actually male among a general population. With better knowledge of this effect, the HCI community can take steps to use more inclusive language in its own research and practice. For example, if the language structures being used by the community do elicit imagery that is more frequently male, then the "people" words being used may not be conducive to designing gender inclusive, as opposed to falsely gender agnostic, tools. These types of agnostic tools, such as the examples discussed from Burnett et al. (2011) and Tan et al. (2003), create a disadvantage for women prompting lower performance and lower feelings of self-efficacy. Determining whether the "people" words of HCI are perceived as gender-neutral could motivate a shift to more inclusive language practice, similar to the shift seen in the psychology community from "subject" to "participant" (Boynton, 1998).

This chapter is organized as follows: first, I discuss the motivation for this study and the approach. Next, an overview of the participants is given along with justification of the sample used. Following this is a discussion of the materials/apparatus created for this experiment, then a description of the research design. An overview of the procedure is given with a description of the data analysis methods. Finally, I present the primary results of the study.

3.1 Study Motivation

Given the issues around gendered language discussed above, research is needed to understand how the general public perceives the HCI "people" words (user, participant, person, designer, researcher), and whether there are patterns in the way they perceive these words . By designing a study to investigate whether the words used in HCI research papers to describe people are perceived as gendered (whether these words are thought of as predominantly male or female), in future research we can see if terms like "user" contribute to gender-biased products with more targeted groups such as software engineers and interaction designers, as seen in Huff & Cooper's (1987) example of designing for "boys", "girls", or "students", where "boys" and "students" received the same result but "girls" received something totally different.

The images that people draw can point to higher-level themes showing how the imagery originates from a cultural or temporal context, as opposed to directly translating these images into specific motivations (Cohn, 1987). Collecting drawings helps us to understand mental processes; they show us representations of perspectives imbued with history and culture (Türkcan, 2013). Despite the diversity of respondents, Mechanical Turk workers have been shown to indicate consistent visual themes in their drawings, pointing to a globalizing effect of imagery (McMaster, 2012). Given this pattern, Mechanical Turk is an ideal place to begin this new research into themes and perceptions underlying the HCI "people" words. For this study, participants were asked to draw a representation of one of five words used to talk about people in HCI papers: user, participant, person, designer, and researcher. Through a questionnaire, information such as the gender, age, and race of the drawing was collected, in the interest of determining whether these representations are thought of as being similar and relatable to the respondents, or whether they are more abstract concepts. A combination of qualitative and quantitative analysis techniques were used to analyze the study data. The full study materials may be found in Appendix A.[1]

3.2 Participants

Participants were recruited from the Mechanical Turk marketplace. An advantage of Mechanical Turk is that the workers "tend to be from a very diverse background, spanning a wide range of age, ethnicity, socio-economic status, language, and country of origin" (Mason & Suri, 2012, p. 3). To better understand the representations that people associate with the "people" words of HCI and whether this relates to their own demographic information, it was necessary to first explore possible trends with a general audience before hypotheses could be made about how domain experts might perceive these words.

The participants completed Human Intelligence Tasks (HITs), which is how tasks are delivered to workers on Mechanical Turk; this is defined further in Section 3.3.1. We posted 5 unique HITs, one per word (user, participant, person, designer, and researcher) with 150 workers per HIT, for a total of 750 instances; however, some participants completed multiple HITs leaving only 433

[1] This study was published in the proceedings of Graphics Interface 2015 with co-authors Adam Bradley, Mark Hancock, and Sheelagh Carpendale (Bradley, MacArthur, Hancock, & Carpendale, 2015). My role in this work involved defining the research question, bringing in our theoretical and study motivations, designing the survey questions, and discussion of implications. Adam Bradley ran the analysis to find the "people" words, compiled the data set prior to analysis, and we wrote the final paper collaboratively. The analysis for gender, age, and race was done collaboratively with the co-authors with my involvement specifically in the interpretation and contextualization; Mark Hancock ran the statistics for these three areas. All additional analysis is my own.

unique respondents (58 participants completed all five HITs, 9 completed four, 12 completed three, and 34 completed two). We received 766 responses, which is greater than the number of requested HITs, perhaps due to participants restarting the study or revisiting the URL after problems such as system failure. Thus, in the data cleaning process, we used only complete responses, consisting of an image and corresponding questionnaire data. This gave us 756 unique responses. Therefore, the analysis is based on the unit of responses, rather than participants, affecting sample size, degrees of freedom, etc.

Although this research was an investigation primarily into gender of participants and their drawings to determine whether these drawings reflected the participants, other demographic information was collected during the study such as ethnicity and age of the participant and drawing. Additionally, participants could voluntarily indicate other information such as their first language and self-rated English proficiency. This became an interesting point for discussion in Chapter 5. As to English proficiency, 328 participants indicated English as their primary language. Another 316 indicated another language as primary, but 226 (72%) of these respondents indicated being fluent in English, 73 (23%) indicated functional knowledge, and 17 (5%) indicated limited or no knowledge of how to communicate in English, thus we can be reasonably confident in participants' comprehension of the questionnaire which was delivered in English only.

This research was reviewed and received clearance from the University of Waterloo Office of Research Ethics. This approval is provided for reference in Appendix B.

3.3 Materials/Apparatus

3.3.1 Mechanical Turk

The study was conducted using Amazon Mechanical Turk, which is an online, crowdsourced "labor market where requesters post jobs and workers choose which jobs to do for pay" (Mason & Suri, 2012, p. 1). These jobs, sometimes called "microtasks", take relatively short periods of time and payment is proportionate. These microtasks are formally called *Human Intelligence Tasks*, or HITs. When workers decide to select your HIT, they are paid on completion. As of 2011, Mechanical Turk claimed "hundreds of thousands of workers and roughly ten thousand employers", which makes it the source of a large set of persistently available people willing to participate in tasks including research studies (2012, p. 2). It is a beneficial tool for researchers not only for its subject pool access and diversity, but also because it is low-cost and the payment mechanism is built-in: Paolacci et al. (2010)

replicated studies from decision-making literature and received results comparable to those conducted in a laboratory with undergraduates, at a cost of approximately $1.71 per hour per subject. Mechanical Turk also allows researchers to quickly move from theory to experimenting; when recruiting and fixing errors in methodology would usually cause significant downtime, this method allows for rapid iteration of the theory and experimental execution (Mason & Suri, 2012, p. 3). As this study involved a new methodology it was an asset to test and refine the procedure rapidly and at low-cost during a pilot to ensure better comprehension among the participants and a higher number of usable responses in the actual study.

3.3.2 Word Selection

Forming the basis of our study were the five HCI "people" words that we chose to investigate. Due to the importance of these words to the study and a desire for them to be relevant to the HCI community, they were chosen in a systematic way. The list of five words was derived by gathering all of the papers published in the CHI 2014 proceedings, which is "the world's premiere conference in human factors & computing systems"[2] and generally regarded as the most impactful[3] in human-computer interaction. A frequency analysis of individual words was performed, and from this a list of the top five words used to reference people within the dataset was extracted. To come up with this final list, words with the same lemmas were combined. For example we combined the words "participant" and "participants" into a single list item, as well as "person", "persons", and "people". Table 3-1 shows the results of the word analysis of the CHI literature. These five terms were chosen for further investigation in this study.

Word	Instances
user(s)	17663
participant(s)	14523
person(s)/people	2337
designer(s)	1092
researcher(s)	1044

Table 3-1: Frequency of "People" Words in CHI 2014 Literature

Note that the word "individual" was sometimes used to refer to people and actually has 1795 instances in CHI 2014 literature. We discussed including this word in our list; however, many of these instances were not referring to people (e.g. the phrase "each individual <object>" was often

[2] CHI 2014 website

[3] Google Scholar metrics showing sources of top HCI publications

used to describe artifacts, instances of data, etc.), and so we did not consider this to be in the top five words referring to people.

3.3.3 Drawing Application

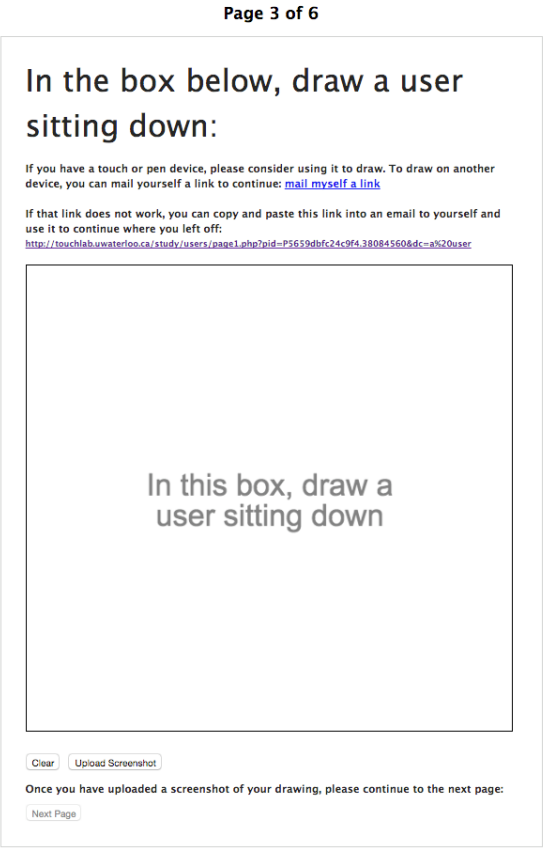

Figure 3-1: Drawing Tool Used in the Study

As discussed in Section 2.4, drawing was identified as a viable means to approach this topic in a novel manner. To collect drawings created by Mechanical Turk workers, a number of methods were explored, and a custom HTML5 paint canvas application was created, as shown in Figure 3-1.[4] This application allowed for black-and-white line drawings, balancing the need for an easy-to-use interface for participants of all skill levels with the flexibility to communicate sufficient details in the drawings. The drawing application was designed such that participants could paste their HIT URL that contained a unique participant ID into a tablet or other touch device if they felt more comfortable

[4] Drawing application created by Mark Hancock.

completing their drawing elsewhere. This was to allow for flexibility and accessibility as participants would not know what the drawing interface was like in advance of opening the HIT, and may find a mouse or touchpad more limiting than a pen, Wacom device, or tablet, for example. Respondents were able to supplement their drawings with text in their questionnaire responses to be clear about their artistic choices and any nuance, these questions are discussed further in Section 3.3.4.

3.3.4 Questionnaire

In this section I will describe the design of the questionnaire component of the study, which was created intentionally to function as a complement to the drawing data. By answering these detailed questions and providing information about the drawings, coding was not needed to interpret the visualizations (i.e. whether they drew a male or female) and data could be analyzed based on the responses submitted to questions, with drawings as supplementary data. After completing and uploading their drawing, the participants were first asked questions about the drawing, and then about the participants themselves.

Once the drawing was uploaded, a single question was presented:

"Did you draw a human?"

The options to answer here were yes or no. Upon answering "yes", participants were asked:

"Does the subject of your drawing have an age, and if so, what is it?"

The following question was:

"Does the person in your drawing look like you, or someone else?"

With a free-text box to collect participants' responses to this open-ended question. This question was designed to prompt participants to think about whether the drawing reflected themselves in some manner, or perhaps a family member, community member, someone in the media, or perhaps something else entirely. It also provided an opportunity to describe whether they had someone in mind when drawing.

Because our study was specifically targeted at collecting data about gender issues, we took care in how we asked participants about gender in the next question. Specifically, we asked:

"Does the subject of your drawing have a gender, and if so, what is it?"

With the options: Male/Female/Other, and a follow-up question:

"If you answered 'other', you may elaborate here".

Our decision to provide the options "Male/Female/Other" came after careful consideration of sources intended to support researchers needing to collect this type of data. We intentionally asked about gender as opposed to sex, using the terms "male" or "female" as opposed to "man" or "woman", since the former have been found to be less problematic than the latter which are socially constructed (Balarajan, Gray, & Mitchell, 2011, p. 9). In the interest of a feminist approach to data collection, it was an obvious choice not to restrict the options here to simply "male" or "female". However, since the participants were not aware that gender was something we would be analyzing, we kept this question a similar length to the others, therefore not drawing attention to it with a fully comprehensive list such as the 51 gender identity options currently available on Facebook (Herbenick & Baldwin, 2014). Due to this concern we also considered but ultimately did not choose to follow a 2-step approach by Balajaran et al. (2011) to first ask about sex or how they were described at birth followed by a question about current gender identity. Based on recommendations from the Human Rights Campaign ("Human Rights Campaign (HRC)," 2013) we chose to offer "male", "female", and "other" with the opportunity to specify or elaborate if the participant desired. While an open-ended text field similar to Question 2 would be most flexible, Enders (2013) notes that this situation is not ideal when statistics need to be done on the answers as they would all need to be coded, thus a compromise would be to offer "male", "female", and "other" with space for an optional description. A study about gender self-classification by Conron et al. (2008) showed that such an addition permits the identification of transgender participants without confusing cisgender participants, and would not be a source of discomfort leading to anyone avoiding the question.

For ethnicity, we asked:

> *"What is the race/ethnicity of the person in your drawing? (i.e., peoples' ethnicity describes their feeling of belonging and attachment to a distinct group of a larger population that shares their ancestry, colour, language or religion)"*

With options: Caucasian, Latino/Hispanic, Middle Eastern, African, Caribbean, South Asian, East Asian, Mixed, Other (clarify). At first, we referenced census practices to be consistent with how this data has been collected in the past. However, the categories for the United States census are very broad, for example, "Asian or Pacific Islander" encompasses an incredibly vast population, and "White" is meant to include persons of both European and Middle Eastern descent. Since Mechanical Turk workers reside worldwide, one country's way of classifying ethnic or racial groups would not

suffice. Morning (2004) performed a review of how censuses across the world asked about ethnicity, and we developed our question as a compromise of overlaps occurring in this list.

Participants were also asked:

"Were there any objects in your drawing? If so, what were they?" and *"What factors did you think about before drawing, if any? (Example: age, race, gender, objects, setting, sexual orientation, outfit, none…)"*

Both of these questions were designed to allow participants to describe the thought process behind their drawings, and to provide an opportunity to say things that they were not yet able to convey. They were also provided a final free-form text box to include any last comments about the drawing.

The questions used to ask participants about themselves followed the same order and format as those previously described, starting with age and gender. Prior to asking ethnicity, we gathered the following demographic data:

"In what country were you born?" and *"In which country do you currently live?"*

This data was intended to provide more context to the ethnicity information collected, and to help us understand where our participants were coming from.

Since participants were coming from all over the world and since this study was focussed on the interpretation of language, it was important to understand participants' English comprehension. We also collected data on first language. Participants were asked:

"Do you speak a language (or languages) other than English at home?"

With the options Yes/No. If they answered yes, they were asked to list which language or languages were primarily spoken. We then asked:

"What is your first or primary language?"

This was asked because some participants might speak English at home, but it may not be their first language. Or, they might not speak English at home, but have a different first language. Finally, they were asked:

"If English is not your first language, how well would you rate your ability to communicate in English?"

The options for this question ranged from “No knowledge at all”, “Limited knowledge”, “Functional knowledge”, “Fluent”, to “Not applicable (English is my first language)”. This sequence of questions was used in the 2000 United States Census and is used by other government agencies as well (Shin & Bruno, 2003). The answer options were modified from the original “Very Well”, “Well”, “Not Well”, “Not at all” to be more informative. By collecting this language data, we make it possible to investigate responses by different language groups after the study.

Previous efforts of the Draw-A-Scientist test have looked at results across ethnic groups (Finson, 2003) or socioeconomic status (Chambers, 1983). Some have tried the test in other countries. None of them, however, have taken a similar approach in asking so many questions about both the drawing and the participant. By collecting supplementary data we are able to make further comparisons that would not be possible with simply a drawing and some answers about the participant. We are also able to more accurately categorize the drawings received. Studies like Chambers (1983) and its derivatives rely heavily on specific coding to determine how many markers of a “scientist” stereotype were present in drawings, and they made comparisons between the number of female participants versus the number of female drawings. It’s unknown whether specific stereotypes exist to do with the HCI “people” words. Additionally, the images people draw aren’t translatable into specific motivations, but rather provide higher-level themes (Cohn, 1987). While we are interested in whether gender arises as a construct used to categorize the “people” words, we are also interested in “a deeper understanding of the psychological, organizational, social and ergonomic factors that affect the use of computer technology” (Abras et al., 2004, p. 10), including how others perceive the place of these gender-neutral terms in this wider landscape.

3.4 Study Design

There were five conditions in the study. The five possible conditions participants could be asked to draw were: user, participant, person, designer, and researcher. In each condition, the Mechanical Turk worker would be directed to our server and asked to think for 10 seconds about one of our “people” words. Workers would only be asked to draw one of these words, and then answer a questionnaire about their drawing and about themselves.

In the first analysis, the gender of the participant was the independent variable, and the gender of the drawing was the dependent variable. Chi-square tests were performed in each condition to determine whether the proportion of male or female drawings (DV) differed significantly from the

proportion of male or female respondents (IV). We also planned to compare ethnicity of participants (IV) to ethnicity of the drawings (DV), however there was not sufficient spread to do so. Paired t-tests were used to compare the age of participants (IV) to the age of the person drawn (DV). For a comparison between first languages of participants and the gender of their drawings, I first found whether the language had grammatical gender. I used a chi-square test to compare an expected proportion of male or female drawings (IV) to the actual number of male or female drawings (DV). I investigated answers to the "looks like me or someone else" question, the words that were used in participants' responses, and the presence of artifacts or markers in the drawings on a qualitative basis, observing patterns but not performing statistical analysis.

3.4.1 Hypotheses

We began thinking about this work because we thought that there was a possibility that the word "user" was acting as a false universal, which would result in it predominantly being thought of as male. Therefore, when participants were asked to draw the word, we expected the drawings would be mostly male. We expected that this bias would exist for participants of all genders. We thus hypothesized:

> H1. *Participants would draw "a user" and describe the drawing more frequently as "male".*

After an initial pilot, with the conditions "user" and "person", we observed the trend that both were thought of as male, and decided to include the five conditions described above. We thought that this might indicate that all of the words we use to describe or identify people within CHI papers are thought of as male. We thus can describe the same hypotheses for any of the words as follows:

> H1 (revised). *Participants would draw any of these words and describe the drawing more frequently as "male".*

We expected a similar bias about ethnicity descriptions, and thus had the following hypothesis:

> H2. *Participants would draw any of these words and describe the drawing more frequently as "Caucasian".*

For the age analysis, I hypothesized the following across all conditions:

> H3. *Participants would draw any of these words and the ages of the subjects in the drawings would differ from the ages of the participants.*

The additional data collected in the questionnaire is a valuable resource to contextualize the drawings. As a follow-up to the hypothesis tests performed above, qualitative analysis was carried out in the following four areas. First, the effect of first language on drawing gender has not previously been investigated in a similar context. I analyzed this data to see whether there was an influence of the participant's first language, hypothesizing that:

> H4. *Participants whose first language does not use grammatically gendered nouns would draw "male" and "female" drawings equally frequently.*

For the following three areas, I posed questions to guide the investigation:

1. *Do any of the conditions prompt more responses of "like me" than others?*
2. *What patterns of word use, if any, emerge around each condition?*
3. *Are there trends around inclusion of artifacts or markers in the drawings?*

Based on the literature, I would expect to see "person" emerging as more relatable than a word like "researcher". Determining which kinds of words are actually used across each condition will be a new contribution of this study. If the word "user" is skewed masculine and does imply dependency as the literature suggests, then I would expect more objects to be included in the "user" drawings than the "person" drawings.

3.5 Procedure

Mechanical Turk workers were given the opportunity to participate in our study via the Mechanical Turk platform. If they decided to participate in our Human Intelligence Task (HIT) then they would be directed to our server and presented with an information letter and consent form, which can be found in its entirety in Appendix A. Following acceptance of the ethics policy and consenting to participate, participants would be asked to think for 10 seconds about one of five words: "user", "participant", "person", "designer", and "researcher" using the exact phrase "for the next 10 seconds think about a …", as shown in Figure 3-2. Our software prevented them from continuing the questionnaire until these 10 seconds had elapsed.

Figure 3-3 depicts the drawing tool, where participants were then asked to draw that word with the phrase: "in this box, draw a … sitting down". In the image, it is shown with an example of what a participant could draw in the canvas. Figure 3-1 shows what the tool looked like when blank.

We specifically recommended that participants switch to a device that provided touch or pen input, if they had one available. Participants could take as long as they needed to do their drawing.

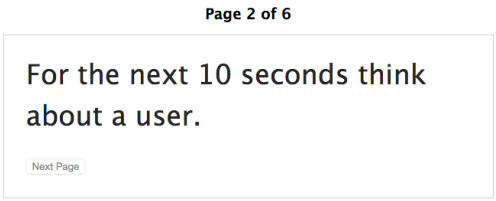

Figure 3-2: Visualization Prompt Prior to Drawing the Condition

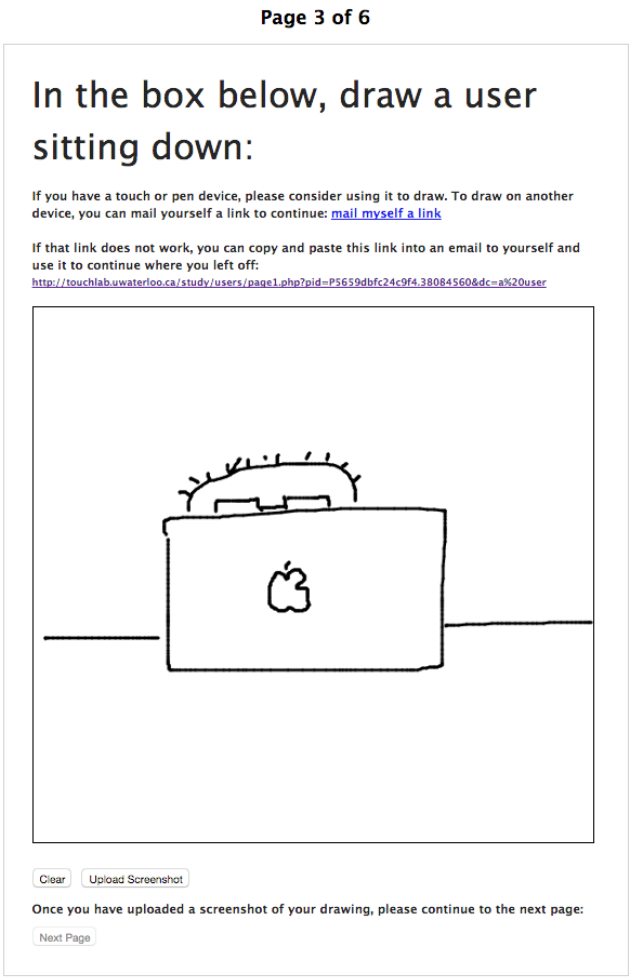

Figure 3-3: Mouse-, Pen-, or Touch-Input Drawing Tool in Non-Empty State with Example Drawing by a Participant (Figure 3-1 Shows Empty State)

Once they uploaded a screenshot of this drawing, they could review it and answer about whether it was a human before continuing (Figure 3-4).

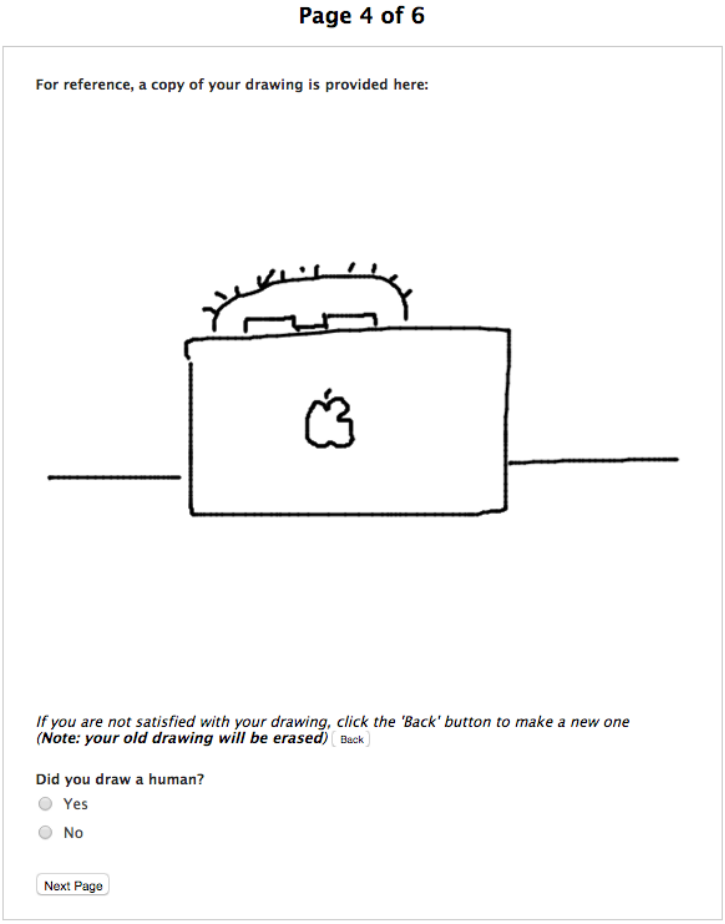

Figure 3-4: Verification of Satisfaction With Drawing

After confirming that they were satisfied with their drawing and indicating whether it was human or nonhuman, participants were given a questionnaire designed to gather data about the drawing (its gender, age, ethnicity, etc.). After answering questions about the drawing, participants were then presented with another set of similar questions asking to describe themselves (again, gender, age, ethnicity, etc.). On completion of the questionnaire, participants were given a code that could be pasted back into Mechanical Turk to confirm they had completed the whole task, and then would be paid. Note that it was only at this stage that participants were made aware that the intention of the study was to investigate how certain words may be gendered.

3.6 Data Collection

Two forms of data were collected: the drawings and the questionnaire responses, which were linked to each other using a unique participant identification number. The questionnaire responses were either answers selected from a predefined list (e.g., male/female/other, or the list of ethnicities) or free-form responses (e.g. any details participants wanted to add about what they drew). In our data cleaning process we used only complete responses with an image and corresponding questionnaire data, as some participants uploaded just a drawing, or left the drawing interface blank which resulted in a screenshot of the prompt being uploaded.

3.7 Results of Drawing and Participant Genders by Condition

Across the five conditions, 756 drawings were collected in total. Figure 3-5 shows the total number of drawings received per condition by and of each gender, as well as the proportion of gender of drawings received in the conditions.

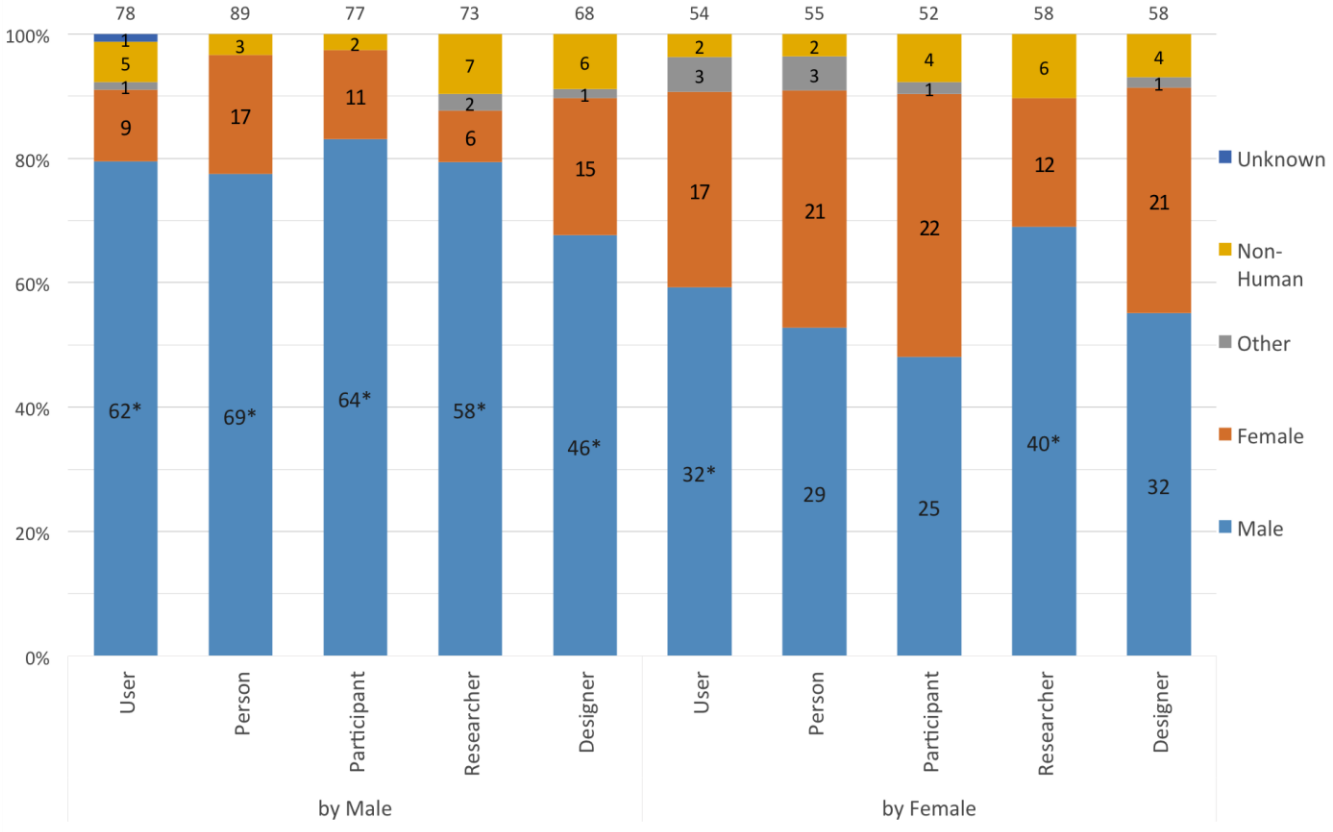

Figure 3-5: Number of Drawings by Gender

3.7.1 Analysis of Gender Results

Across all five conditions, only 15 responses described the participant's gender or the gender of the image as "other", therefore statistical analyses on data about this group were not conducted. Section 3.7.2 describes in further detail these instances where "other" was used. There were 134 instances when the gender of the drawing or participant was "unknown", which occurred when participants either described the drawing as being non-human or when they did not provide data about the image's or their own gender. A Pearson's Chi Square test was performed on the remaining drawings: those described as being either male or female by participants who described themselves as being either male or female. A summary of the results is found in Table 3-2.

Drawing Condition		User		Participant		Person		Designer		Researcher	
Participant Gender		M	F	M	F	M	F	M	F	M	F
Results	df	1	1	1	1	1	1	1	1	1	1
	N	71	49	75	47	86	50	61	53	64	52
	χ^2	39.6	4.6	37.5	0.2	31.4	1.3	15.8	2.3	42.2	15.1
	φ	0.71	0.29	0.70	0.06	0.59	0.15	0.48	0.20	0.76	0.51
	p	<.001*	0.03*	<.001*	0.66	<.001*	0.26	<.001*	0.13	<.001*	<.001*

Table 3-2: Summary of Gender Analysis for All Conditions

Our hypothesis H1 was that participants would draw any of these words and describe the drawing more frequently as "male". For analysis, the null hypothesis was accordingly that the drawings would be equally distributed across these two genders: that the word, if it was interpreted as being non-gendered, would be uniformly distributed independent of participant gender.

For all five words, images drawn by males were more frequently of males than of females (shown in Figure 3-5). This suggests that when men read these commonly used words, they tend to perceive these people to be men. Similarly, images drawn by females were more frequently of males than of females for the words "user" and "researcher". This finding suggests that when women read these two words, they tend to perceive these people to be male.

Promisingly, images drawn by females for the words "participant", "person", and "designer" were not more frequently of either men or women. While it is not appropriate to accept the null hypothesis (i.e., that this frequency is actually equal to 50%), this finding suggests that this gender bias, if it does exist, is far less pronounced for women for these words than it is for men. Based on

this data, H1 can be confirmed for drawings by male participants and all words, but H1 can only be confirmed for drawings by female participants for the words “user” and “researcher”.

3.7.2 Use of “Other”

Out of 756 unique responses, 15 study participants chose to use “other” to describe their gender or the gender of their drawing. Of these 15 responses, only one participant chose not to clarify their answer in the accompanying text box. 13 uses of “other” (87%) were by study participants whose first language is English, one was by a Tamil speaker, and one was unspecified. “Other” was used two times to describe the study participant’s gender, and 13 times to describe the gender of the image drawn. The two study participants who identified as “other” wrote that they were “genderqueer” and “ftm” (female-to-male), respectively. Both of these participants drew females. Of the study participants that elected to use “other” to describe the drawings, four were male, seven were female, and one is unknown (left blank).

Of the 12 drawings described as “other” that included accompanying reasons, nine were purposely unspecified or left ambiguous (one example being that they chose an alien to be a computer user, so that “anyone can identify”), one stated that they did not think about gender, one image actually contained two genders (a male and female in the image together), and one drawing (for “participant”) was transgender. Some participants who were purposely ambiguous about gender stated that they did so to deliberately resist putting their drawings into a box. One female participant who drew a designer chose “other” but wrote in the elaboration box that they “would think more female for designers, younger, white”. A male participant who drew a researcher wrote that the drawing was “generic”, “asexual”, and that they wanted to avoid giving it “stereotypical” attributes.

3.8 Ethnicity

Table 3-3 below presents a summary of participant ethnicities and the ethnicities recorded for all drawings. “Non-human” was not an option for respondents to describe themselves, as this would preclude their ability to participate in the study. Figure 3-6 and Figure 3-7 (page 48) summarize the frequency by which each ethnicity drew other ethnicities in each of the five conditions.

Ethnicity	Drawings by Ethnicity	Drawings of Ethnicity
Caucasian	223 (29.50%)	265 (35.05%)
Latino/Hispanic	27 (3.57%)	27 (3.57%)
Middle Eastern	7 (0.93%)	8 (1.06%)
African	17 (2.25%)	25 (3.31%)
Caribbean	4 (0.53%)	10 (1.32%)
South Asian	314 (41.53%)	265 (35.05%)
East Asian	35 (4.63%)	30 (3.97%)
Mixed	12 (1.59%)	34 (4.50%)
Other	14 (1.85%)	23 (3.04%)
Unknown	103 (13.62%)	21 (2.78%)
Non-Human	n/a	48 (6.35%)
Total	756	756

Table 3-3: Ethnicities of Study Participants and of their Drawings

3.8.1 Analysis of Ethnicity Results

Hypothesis H2 was that participants would draw any of the people words and describe them more frequently as “Caucasian”. The sample in the dataset was not spread sufficiently across all ethnicities to be able to conduct formal analyses of this data, with 223 Caucasian, 314 South Asian, 103 not specified, and the remaining 7 with all less than 40. Once spread across the five drawing conditions, there were ≤ 10 data points for each of these 7 ethnicities. However, as shown in Figure 3-6 and Figure 3-7, there was a tendency for Caucasian and South Asian participants to draw people of their own ethnicity in all conditions. It appears as though participants of other ethnicities tend to draw Caucasians in addition to their own ethnicities, though the samples of these populations are perhaps too small to draw conclusions about this data. Thus, H2 cannot be confirmed.

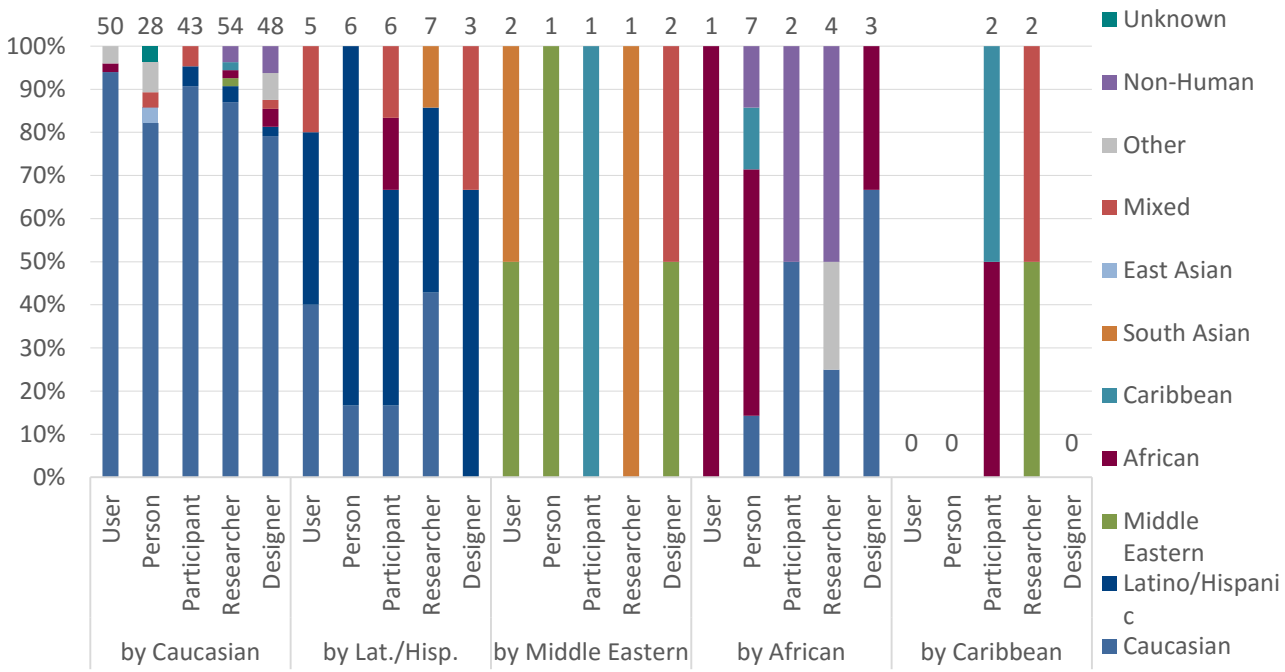

Figure 3-6: Distribution of Drawings' Ethnicities by each Ethnicity per Condition

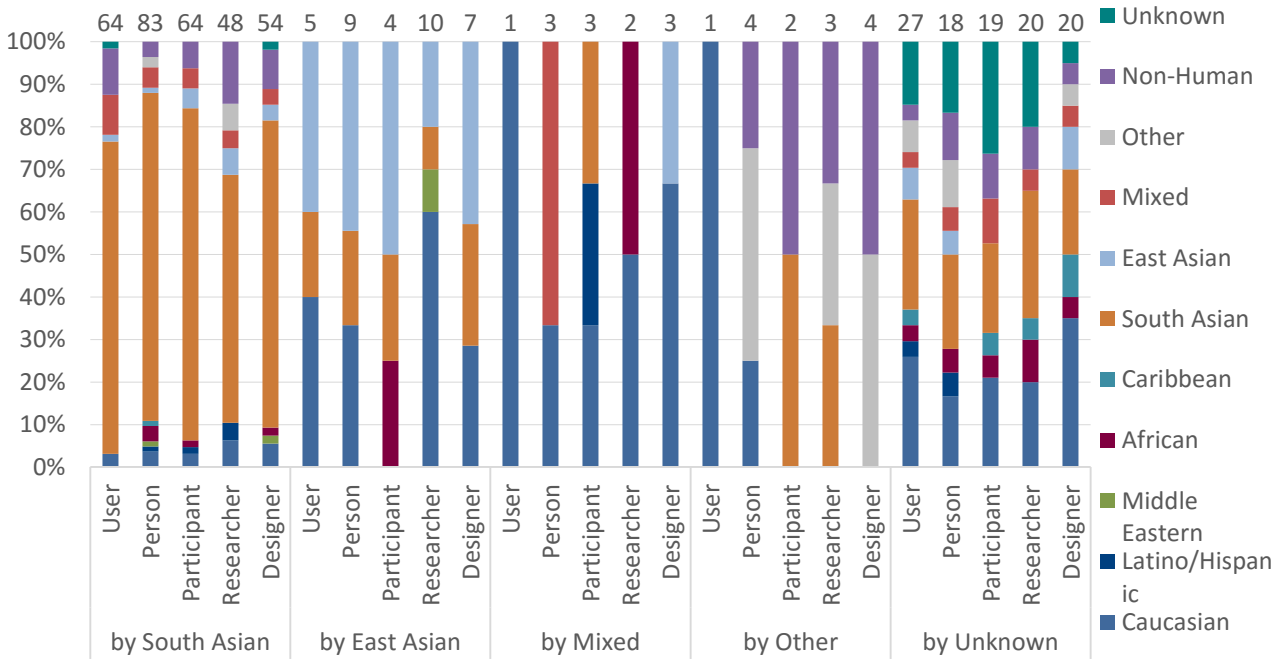

Figure 3-7: Distribution of Drawings' Ethnicities by each Ethnicity per Condition

3.9 Age

The age of the participant and an age of the person they drew were both collected using the questionnaire. Figure 3-8 shows the comparison of average age of the participant and the average age of the drawing in each condition, with asterisks denoting a significant difference between them.

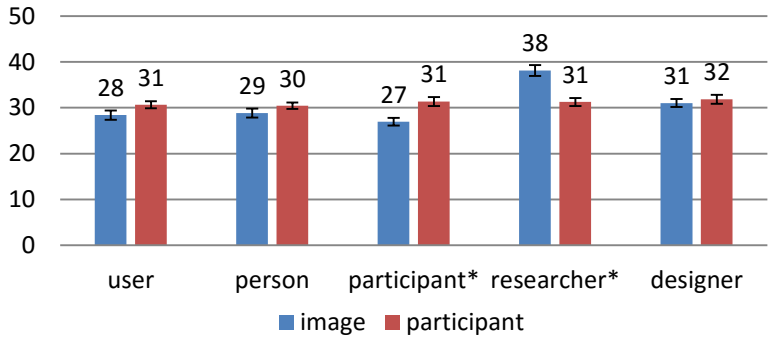

Figure 3-8: Average Age of Images Compared to Average Age of Participants

3.9.1 Analysis of Age Results

I hypothesized in H3 that participants would draw any of the HCI "people" words and the ages of the subjects in their drawings would differ from the ages of the participants. Paired t-tests were conducted to compare the age of participants to the person drawn. Table 3-4 contains the results of this analysis.

Condition	User	Participant	Person	Designer	Researcher
Average Age (Image)	28	27	29	31	38
SE(Image)	1.02	0.85	0.97	0.90	1.17
Average Age (Participant)	31	31	30	32	31
SE(Participant)	0.77	0.98	0.70	0.94	0.88
N	97	98	123	102	107
df	96	97	122	101	106
t	1.8	3.9	1.4	0.7	4.7
Cohen's d	0.19	0.39	0.13	0.07	0.46
p-value	0.07	<.001*	0.16	0.49	<.001*

Table 3-4: Analysis of Age Data for Images and Study Participants

Drawings were of people younger than themselves when they drew "participants" and older than themselves when they drew "researchers". For the remaining three words, the age of the drawings was not significantly different than the age of the participant. This finding suggests that our

participants perceive researchers to be older, participants to be younger, and the remaining words (user, person, and designer) to be of similar age.

3.10 Chapter Summary

A Mechanical Turk study was used to investigate how the most common words that have been used to refer to people in recent HCI literature are received by non-experts. The top five CHI 2014 people words are: user, participant, person, designer, and researcher. We asked participants to think about one of these words for ten seconds and then to draw an image of it. After the drawing was done we asked simple demographic questions about both the participant and the created image. The collected questionnaire responses were analyzed using a mixed method of qualitative and quantitative techniques. The initial results of the study suggest confirm H1 that a gender effect is present for men in all conditions and for women in the "user" and "researcher" conditions only. There was insufficient data to support H2, and the age difference of H3 was confirmed for "participant" and "researcher" only. More qualitative analysis of study data is presented in the following chapter.

Chapter 4
Qualitative Analysis

The hypotheses tested in Chapter 3 provide a first look into the perception of HCI's "people" words. There was more data collected in the questionnaire beyond the gender, age, and ethnicity data, which is intended to assist with interpreting and contextualizing the results and the drawings. In this chapter, I summarize the analysis of responses by different linguistic groups (Section 4.1); whether participants indicated their drawing was "like them" (Section 4.2); use of terms in responses across the five conditions (Section 4.3); and artifacts observed in the drawings (Section 4.4). The implications of these results are discussed in Chapter 5.

4.1 Language

Participants in the study were asked their first or primary language, whether they spoke a language other than English at home, and if so, to indicate the language. In the case that English was not their first language, they were asked to rate their ability from "no knowledge at all" to "fluent". Of the 756 study participants, 332 indicated that English was their first language, 301 indicated English was their second language, and 123 left this question blank. 328 participants indicated that they speak a language other than English at home, 97 of these speaking more than 1 language. Some participants left their first language blank but indicated their English ability as a 5, meaning "not applicable (English is my first language)" which raised the count of native English speakers to 344.

Participants' first languages other than English that had more than 10 speakers were: Tamil (108), Malayalam (44), Bengali (22), Hindi (15), and Telugu (15). Languages spoken primarily at home (if not English) with more than 10 speakers were: Tamil (127), Hindi (105), Malayalam (50), Telugu (29), Bengali (25), Spanish (16), Kannada (13), French (12), and Russian (11). The full variety of languages provided by participants is described on the next page in Table 4-1, where subcategories of the language families have been provided to assist with categorization. The final column denotes whether these languages are known to have grammatical gender, which is of interest for further analysis.

Participants' Languages				Speak at Home	First Language	Gendered Nouns
Afro-Asiatic	Chadic	Hausa		1	0	M F
	Semitic	Hebrew		1	0	M F
		Arabic		1	0	M F
Austroasiatic	Vietnamese			2	1	n/a
Austronesian	Tagalog			3	0	n/a
French Sign-Based	ASL			1	0	n/a
Tai-Kadai	Thai			5	4	n/a
Indo-European	Italic	French		12	1	M F
		Italian		6	1	M F
		Spanish		16	3	M F
	Balto-Slavic	Lithuanian		6	0	M F
		Russian		11	9	M F N
	Germanic	Dutch		3	3	M F N
		English		**n/a**	**332**	**n/a**
	Indo-Iranian	Pashto		1	0	M F
		Indo-Aryan	**Bengali**	**25**	**22**	**n/a**
			Hindustani (incl. Hindi and Urdu)	**110**	**21**	**M F**
			Gujarati	1	0	M F N
			Saurashtra	3	6	M F N
			Konkani	1	0	M F N
			Marathi	9	8	M F N
			Nepali	1	0	n/a
			Marwari	3	0	M F
Dravidian	Telugu			29	15	M F N
	Tulu			5	0	M F N
	Tamil			**127**	**108**	**M F N**
	Malayalam			50	44	M F N
	Kannada			13	3	M F N
Sino-Tibetan	Cantonese			1	0	n/a
	Mandarin			4	0	n/a
	Chinese			1	0	n/a
Japonic	Japanese			0	1	n/a
Other	Creole			1	0	n/a

Table 4-1: Languages Spoken at Home & First Languages of Participants

4.1.1 Analysis of Responses by Language Groups

Boroditsky, Schmidt and Phillips (2003) ask whether thinking for speaking a particular language has an effect on how people think when not thinking for speaking that same language. In other words, they are wondering how thinking for speaking one language might exert influence over other types of thinking. They cite past research showing that despite linguistic similarities, people have remarkably similar colour memory, but very different conceptualizations of time, suggesting a linguistic influence over more abstract domains not so reliant on the senses (2003, p. 63). Over a series of studies presented in *Sex, Syntax and Semantics*, the authors demonstrate that people's thinking about objects is influenced by the grammatical gender that their native language assigns to the objects' names (2003, p. 70). They even demonstrated that just differences in grammar, and no differences in culture, can be enough to influence how people think about objects. To go one step further, they studied the phenomenon through the use of images without words. Again, they showed that learning new grammatical categories influences how people think about objects (i.e. through perceived increase in similarity between pictures) (2003, p. 73). When they ask specifically whether linguistic categories, such as a noun's being grammatically masculine or feminine, can actually alter non-linguistic representations, the answer was *maybe*: they had done enough studies to preclude saying no, and the phenomenon is far too complex to say exactly the interaction between linguistic thinking and non-linguistic representations.

This raises an interesting question for the research presented here. While it has been said that Mechanical Turk displays a globalizing effect of imagery, thus being a reliable source for image-based research (McMaster, 2012), there are a large number of study participants with a non-English first language. For speakers of English as a first language, out of 306 drawings 74.18% were of males and 25.82% were of females. As English does not have gendered nouns, then language is not expected to play a significant role in how the HCI "people words" are perceived. A Chi-square test shows that the null hypothesis should be rejected as English speakers drew significantly more males than females ($\chi^2(1,N=306)=17.51$, $\varphi=0.24$, $p<.001$). As Boroditsky et al. wrote, it is very complex to draw the relationship between how we think of words and what we draw. In Hindi, the words user, participant, person, designer and researcher all have a male grammatical gender. The sample size of 15 drawings was too small for a successful investigation. In Tamil, the HCI "people" words have a common neutral gender unless referring to a specific person. A significant result in Tamil shows a tendency to draw males more frequently ($\chi^2(1,N=91)=10.02$, $\varphi=0.11$, $p<.01$). In Bengali, a language without grammatical gender, the drawings were submitted by males only. In this part of the sample,

there were 10 drawings of males and 10 of females. Though this was an even split, the sample size was too small for conclusive analysis, and there were no drawings received by women for comparison.

4.2 Looks Like Me or Someone Else

After completing the drawing and answering whether it was human, the questionnaire asked participants whether the drawing "look[s] like you, or someone else?" Of the 756 responses, 116 (15%) indicated that the drawing did look like them to some extent, for example: "a little bit/kind of like me"; "like me but younger/taller/thinner/behaves differently"; "like me in bad drawing version"; "similar to me in hair/skin/ethnicity"; "it's just like me"; "it is me"; "like me due to the setting". Since this question had a free-form answer field, participants could also write things like "it looks like my husband" or "it looks like someone else" without specifying whether the husband or someone else is "like them" or "not like them."

There were 78 cases (10%) of participants indicating the drawing looked like someone they knew personally: a brother/sister (8), son/daughter (5), husband/wife (3), girlfriend/boyfriend (5), mother/father (7), or other relation such as nephew/grandmother/grandfather (14). There were 27 instances where "friend" was used to describe what the drawing looked like, 11 of these being for the "person" condition (all drawings of males), the next highest instance being 6 for "user" (again, all males). Overall, "person" received the most responses indicating the drawing looked like someone they knew (25 overall, 21 male and 4 female).

There were cases where participants referred to unnamed celebrities ("someone I saw on a show"), however there were also several when specific people in the public eye were named. Coco Chanel, Calvin Klein, Vera Wang and Alexander McQueen were all named as designers. For the "person" condition, Arjun Kapoor (Indian actor) and Justin Bieber (Canadian pop star) were both named. Sachin (a famous cricket player) was named as a "user", Bharathiyar (famous Tamil activist) and Rajinikanth (Indian actor) were named as "participants". For the "researcher" condition, Albert Einstein had multiple mentions, accompanied by Stephen Hawking, Marie Curie, and Deborah Gordon (professor of biology at Stanford University). "Researcher" was the condition with the most references to famous names with 6 in total (2 of females and 4 of males, 3 of those for Einstein), "designer" is a close second with 5 references to famous names (3 of females, 2 of males). These are, however, just the explicit references made by participants, not including any implicit references made

through drawing likenesses, or instances where a participant failed to answer the "looks like you" question but indicated their drawing's inspiration in another question. An analysis of the words used in the "researcher" condition across the whole questionnaire in Section 4.3.5 shows that "Einstein" was used 6 times in total.

Looks like…	User		Participant		Person		Designer		Researcher	
	M	F	M	F	M	F	M	F	M	F
Me	20	10	21	8	15	12	9	4	11	6
	30 (20.4%)		29 (20.4%)		27 (18.2%)		13 (9.5%)		17 (11.9%)	
Generic person	11	1	1	0	5	4	2	0	4	0
	12 (8.2%)		1 (0.7%)		9 (6.1%)		2 (1.5%)		4 (2.8%)	
Someone I've seen	8	0	5	3	7	1	8	0	0	1
	8 (5.4%)		8 (5.6%)		8 (5.4%)		8 (5.8%)		1 (0.7%)	
Person in the public eye	1	0	2	0	2	0	2	3	4	2
	1 (0.7%)		2 (1.4%)		2 (1.4%)		5 (3.6%)		6 (4.2%)	
Someone I know personally	16	6	8	5	21	4	7	1	9	3
	22 (15.0%)		13 (9.2%)		25 (16.9%)		8 (5.8%)		12 (8.4%)	
No one in particular	10	0	8	7	10	3	7	5	9	4
	10 (6.8%)		15 (10.6%)		13 (8.8%)		12 (8.8%)		13 (9.1%)	
Typical version of the condition	0	0	2	0	0	0	9	8	25	3
	0 (0.0%)		2 (1.4%)		0 (0.0%)		17 (12.4%)		28 (19.6%)	
Someone else	38	11	40	15	34	15	35	17	40	3
	49 (33.3%)		55 (38.7%)		49 (33.1%)		52 (38.0%)		43 (30.1%)	
Definitely not me	1	0	2	0	0	0	0	1	1	0
	1 (0.7%)		2 (1.4%)		0 (0.0%)		1 (0.7%)		1 (0.7%)	
A non-human	6	2	4	3	4	1	6	4	5	6
	8 (5.4%)		7 (4.9%)		5 (3.4%)		10 (7.3%)		11 (7.7%)	
Indecipherable	4	2	5	3	8	2	8	1	6	1
	6 (4.1%)		8 (5.6%)		10 (6.8%)		9 (6.6%)		7 (4.9%)	
Total	115	32	98	44	106	42	93	44	114	29
	147		142		148		137		143	

Table 4-2: Summary of "Looks Like Me or Someone Else" Categories

4.3 Word Usage Across the Conditions

The benefit to the questionnaire data accompanying the drawings is that participants had several opportunities to describe and qualify what they drew. They could add context, and some participants even chose to include back-stories about the people shown in their drawings. The qualitative data analysis software ATLAS.ti was used to assist with discovering patterns in the questionnaire data received. The tool was used to determine occurrences of words used in questionnaire responses across

the five conditions. For each condition, ATLAS.ti generated a list of the most frequently occurring words in the participants' answers. Then, I categorized these words into basic parts of speech: object nouns, verbs, adjectives, proper nouns (i.e. brother, doctor, student), human-related nouns (i.e. hair, clothes, shirt), setting nouns (i.e. chair, table, park, room), and other (i.e. nothing, random). Stop words were not counted, nor were words that repeated text from the question (i.e. gender, someone, ethnicity, elaboration). After this exercise some words coalesced into themes of related words, which were verified in the participants' complete responses for context of use. The 1-2 themes with the greatest number of related words are shared in the following Sections 4.3.1 to 4.3.5. The extent of participants' responses to the questionnaire were not predictable before completing the study, as all questions were not required. The analysis done here is a first pass of the data received to determine whether patterns would begin to emerge. In order to speak more definitively about the meaning of these patterns and their frequency, further studies should include a second coder to assist with reliability of categorizing ATLAS.ti output.

4.3.1 User

In the 148 drawings of a human "user" by both males and females, the patterns emerging from the most frequently used words were an association with consumer electronics, and for a smaller subset, with substance abuse.

Words to do with computer/technology usage included: computer(s) (57), desk (21), laptop (19), table (19), work(ing) (16), keyboard (8), monitor (6), phone (6), iPhone (4), tablet (4), battery (2), headphone(s) (3), mobile (2), MTurk (2), PC (2), smartphone (2), tool(s) (4), app (1), flatscreen (1), gadget (1), game (1), graphics (1), innovation (1), machines (1), prototype (1), screen (1), silicon (1), software (1), website (1).

The "users" of consumer electronics were often either relaxing or entertaining themselves while solitary: "I draw a man playing video game in a smart phone", "I just tried drawing a user who is using his laptop and enjoying music with his head phone by sitting in the floor". Other depictions made a point of being generic: "I was just referencing instructional drawings if anything. Like the ones you see on vending machines, on instructions, or whatnot", "User is a general term so I wanted my drawing to signify it. That is why the drawing shows a faceless person with no distinct features other than the laptop he is using". Despite this person wanting to be as universal as possible, the use of the pronoun "he" shows that the faceless norm still defaults to masculine.

Words associated with drug/substance use were present: drug(s) (8), needle (5), syringe (3), cigar (2), shoot (2), cigarette (1), death (1), died (1), heroin (1), high (1), inject/injection/injects (3), junky (1), "marajuana [*sic*]" (1), tourniquet (1), vein(s) (2), wastoid (1). There were 9 drawings (5.8%) that were images of drug users, examples are shown in Section 4.4.2.

For the drawings of drug users, 2 participants indicated that the drawing depicted someone they knew, and 4 participants indicated that they drew an anonymous drug user. For those whose lives have been touched by addiction, the connotation between "user" and powerlessness was strongly negative: "I drew my father alone and clouded by the choices he made. The lightning strike symbolized his moment right before death. He realized it was too late". Another participant interpreted "user" in the context of their own personal problems: "I AM BEING USED BY A WOMAN".

4.3.2 Participant

There were 138 drawings received of a human "participant". Words used in this condition referred to *doing* things in particular *settings*, in contrast to the "user" condition, which referred more to objects than situating the person in an unfolding event. Themes emerged around "participant" as a competitor or a performer in a contest (such as in Figure 4-1), and "participants" in research studies such as our own.

Participants as performers or competitors: stage (3), booth (2), singing (6), song (2), music/musician (2), voting (3), ballot (1), cheering (1), clapping (1), dance(ing) (2), perform (1), competition (5), mic(rophone) (6), gameshow (1), game (3), quiz (3), contest (2), race (4), artist(s)/artistic (8), actor (2), athletic (1), competitive (1), publicly (1), pedestal (1).

Figure 4-1: Two Drawings of Competition Participants

To some participants, the “participant” condition immediately invoked some kind of competition: “I thought about what kind of contest he will be participating in”. Some drawings had specific connotations: a “90’s housewife” competing on a game show, and “I was thinking of a 30-year-old white girl with blonde hair, on a game show”. Another wrote “I immediately thought of a female quiz show participant”. Beyond the number of game shows, one participant depicted a boy participating in a drawing competition, and yet another in a spelling bee.

Participants in research/studies/technology: survey(s) (20), paper (9), computer(s) (8), newspaper (5), pencil (5), experiment (4), laptop (4), data (3), desktop (3), monitor (3), screen (3), software (3), device (2), interview (2), keyboard (2), MTurk (2), study (5).

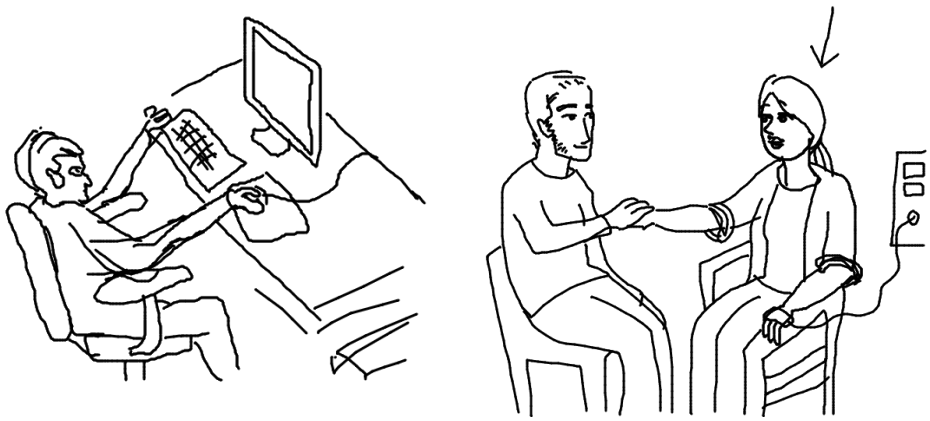

Figure 4-2: Two Interpretations of "Participant"

A small number of participants decided to make a meta-drawing to respond to this condition, showing themselves participating in online surveys (such as Figure 4-2, left). One participant wrote: “The intention was to depict an Mturk workaholic, glued to their screen with money on their mind”. Another said: “I mostly just thought of myself participating in this survey at my desk, though the likeness didn’t come out in the drawing”. Others, such as the drawing in Figure 4-2 (right), interpreted this as participating in medical research: “There is a monitor for heart rate and perspiration attached to my hand, and it is connected to a machine that measures and records the data”. Being a “participant” was not limited to the act of participating in something, but also consenting to do so, as one participant opted to show a girl signing forms to be a part of something.

4.3.3 Person

There were 152 drawings of a human "person". In this condition, there were frequent references to different types of people (family members, for example) as well as generic "people" words. While in the "user" condition participants often referred to objects to help define the person, in the "person" condition there were more frequent references to the physical appearance of the person in the drawing.

Some participants drew someone they know personally, "my grandpa watching tv", whereas others drew a person they may have seen in passing: "male saw in a restaurant cooking food". Words used to refer to known people include: me/myself (34), friend (16), brother (4), daughter (4), mother (3), sister (3), grandpa (2), uncle (2), husband (1), neighbour (1), colleague (1), nurse (2), artist (2), chef (1). When drawing someone familiar, participants sometimes considered their emotional state and used this to describe what we were seeing. For example, under "other factors", one participant wrote "aloneness" then elaborated: "My friend is always thinking how to lead life. Every time bad luck follows him".

Some participants were deliberately ambiguous: "I wanted something that did not represent any elements from my life", "I wanted to draw them like a gingerbread man … basically an upgraded version of a stick figure and featureless". Generic people words used included: man/male/guy/boy (44), girl/female/lady/woman (22), kid (2), human (4), people (2), sir (2), anyone (5), everyone (2), stickman (2).

A number of participants indicated that they thought about outfit before they drew, in addition to factors like age, gender, and position. Words used to describe physical appearance (hair, clothes, body) included: looks (25), outfit/clothes/clothing (13), hair(style) (11), race (6), sex (6), face (5), hat (3), shirt (4), style (4), hand(s) (5), legs (2), dress (2), pant (2), beard (1), bangs (1), cowboy hat (1), ponytail (1), mustash(e) [*sic*] (2), saree (1), shoes (1), suit (1), stylish (1), wearing/wears (3).

Two instances of "computer" were used in one participant's response where they drew their friend, another three came from a participant who drew himself and said that "I will always using [*sic*] computers". References to objects included: TV (10), computer(s) (8), book (4), paper (4), phone (2), keyboard (1), MTurk (1), software (1).

4.3.4 Designer

There were 133 drawings of a human "designer". In this condition, participants would sometimes indicate in the "does it look like you" question that it looked "like a designer" (17 cases). Correspondingly, there was heavy reliance on words that described designers doing designing, making designs, holding the design that they designed, and so on: designer(s) (57), designing/designed (10), design(s) (19). Someone "drew a person sitting down in a chair, thinking about design". Other themes generally showed designers professionally in different contexts.

Accompanying the numerous references to existing fashion designers discussed in Section 4.2 were references to the Devil Wears Prada and other stereotypical, Cruella de Vil designer types as well as an emphasis on appearance: fashion (7), "female NY biotch" (1), chic (1), sewing (1), dress (5), outfit (5), shirt(s) (7), clothing/clothes (3), cloth/fabric (5), mannequin (1), Prada (1), black (3), artsy (1), opulence (1), Paris (2), model (3), Chanel (2), style/stylish (4), blonde (1), bracelet (1), dresses (1), hanger (2), earring (1), haircut/style (2), pant (1), scarf (1), shoes (1), sneakers (1), undercut (1). The following participant's response captures a particular mood: image objects "cigarette holder, dog, chair, scarf", image factors "wearing black", and additional comments "the dog's cigarette is also in a cigarette holder". Another drew their fashion designer with "a nose ring because most designers are quite creative and like to stand out".

Figure 4-3: Drawing of a Designer

Others drew designers or graphic designers in a more traditional artist style: "I drew glasses, a pen, and a computer. A good designer needs all three!" Words in this genre included: graphic(s) (4), creative (1), art/artist (4), printing/prints (2), computer (14), pen(s) (10), Adobe (1), Photoshop (2),

lorem ipsum (1), Wacom (1), paint (5), paper (5), mouse (4), brush (3), canvas (3). One designer was shown with Lorem Ipsum on her monitor as she designed a brochure (Figure 4-3).

"Designer" was interpreted alternatively by others still, with examples of interior designers, an automotive designer, and an architect. There were also two beggars asking for money.

4.3.5 Researcher

Figure 4-4: Example Drawings in the Researcher Condition

There were 137 drawings of a human "researcher". The researchers depicted spanned numerous domains: academia (1), chemistry (6), science (7), knowledge (2), medical (2), biological/biology (2), education (1), engineering (1), geology (1), math (1), pharmaceuticals (1), statistics (1), environment (2), politics (2), advertising (1). In addition to being researchers, they were: scientist(s) (9), professor (4), teacher (4), Einstein (6), academic (3), student (2), analyst (1), astronomer (1), technologist (1), nurse (3), director (2), neighbour (2), alumni (1).

The researchers were shown in settings involving: chair (70), table (32), desk (30), lab(oratory) (12), University (3), Stanford (2), office (5), college (1), sofa (2), stool (2). They had objects close by such as: (super)computers (21), book(s) (32), laptop (8), pen (7), pencils (1), paper(s) (9), beaker(s) (6), microscope (5), clipboard (4), tube(s) (5), data (4), mouse (4), board (2), bottle (2), camera (2), car (2), keyboard (2), money (2), cells (1), degree (1), diploma (1), iPad (1), notebook(pad) (2), PC (1), telescope(s) (3), animal (2), ant(s) (4), equipment (2), instruments (2), journal (2), rock (2), server (2), wheelchair (2).

These researchers had distinct appearances: (eye)glasses/specs/spectacles (19), outfit (12), coat (5), hair (5), hat (5), gloves (1), goggles (1), beard (2), clothes/clothing (4), posture (2). They were often up to something: doing (10), looking (10), reading (6), work(ing)(s) (27), research(ing)

(19), experiment (4), think/thought/thinking (13/27/5), test (5), writing (3), measuring (1), study/studying/studies (4), test(ing) (6), task (3), trying/tried (3/3), daydreamed (2), processing (2).

On occasion, they would have some adjectives if it didn't quite come across in the drawing: bionic (2), daydreamer (2), mad (2), asexual (1), graduate (1), nerdy (1), studious (1), undergraduate (1), employed (2), interested(ing) (4), relaxed (2), stylish (2), androgynous (1).

4.4 Artifacts and Markers in the Drawings

Since participants were asked to draw their condition sitting down, there were many drawings including chairs. In fact, "chair" was the most frequently occurring word across all five conditions. In counts of artifacts in the drawings I have exclusively used the declarations of participants as opposed to coding the drawings.

4.4.1 Computers

Many participants in the "user" condition indicated that they had drawn someone sitting down and typing at a computer; it was only indicated for 6 drawings (3.8%) that they included phones or tablets. This suggests that in addition to "user" being thought of predominantly as male, it also is thought of as a desktop computer user.

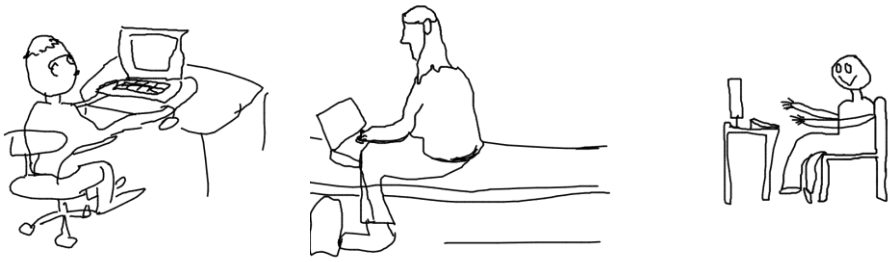

Figure 4-5: Users with Computers

"Computer" was used most frequently in the "user" condition with 55 instances, followed by "researcher" with 19 instances and "designer" with 14 instances. "Participant" and "person" both saw 7 instances of "computer", but this was a lower proportion overall in the "participant" condition. Variations on "computer" included "laptop", seen most frequently in the "user" condition (19 times), followed by the "researcher" condition (7 times).

4.4.2 Drug Use Equipment

There were 9 drawings (5.8%) in the "user" condition described as drug users, some examples are shown in Figure 4-6. Usually these images were described by the participants as either having needles or drug smoking paraphernalia within the image. This finding suggests that the word "user", despite being so ubiquitously used when referencing people who use technology, is still thought of in the context of drug use by a more general population.

Figure 4-6: Examples of Drug User Drawings

4.4.3 Inclusion of Markers

In Chambers' Draw-A-Scientist test, seven markers were used as indicators of the standard image of the scientist: lab coats, eyeglasses, facial hair (beards, moustaches, or long sideburns), symbols of research (scientific instruments and laboratory equipment of any kind), symbols of knowledge (principally books and filing cabinets), technology ("the 'products' of science"), and relevant captions (formulae, taxonomic classification, saying "eureka!", etc.) (1983, p. 258). In their control condition in which participants were asked to "draw a person", none of the indicators were drawn except for eyeglasses in a small number of cases (5, or 0.5%) and facial hair in 14 cases (1.5%) (1983, p. 258). This means that in the stereotyped condition there was a tendency for markers to be present, some of which were manifest in appearance and others in artifacts, whereas in the control condition these were very infrequent and, if anything, to do with physical appearance.

In this study, drawings of the "researcher" condition matched many of the seven markers defined by Chambers, including: lab coats (6); glasses (19); beard (2); symbols of research, such as computers/laptops (38), beakers (6), microscopes (5), test tubes (5), goggles (1), telescopes (3);

symbols of knowledge including books (32), degrees/diplomas (2), paper/clipboard/journal (15), pens/pencils (8), “studious” (1); technology such as supercomputers (1), “bionic” (1), “innovation” (2). Participants did not log captions such as the “eureka!” example however they did note things like the subject was reading (6), working (27), researching (19), thinking (45), writing (3), studying (4), and so on.

In the “person” condition of this study, there were no inclusions of lab coats nor glasses. There was one case of a beard. Artifacts in common include computers (8), symbols of knowledge include books (4), papers (4), and pens (2). There was no similarity in technologies. Common verbs for the “person” condition included sitting (71), working (12), doing (5), going (5), watching (7), smiling (3) – there was no overlap with concepts like reading, researching, writing or studying.

In analyzing the images we also observed many images with what could be considered gender markers, such as facial hair and ties for men, and dresses and bows for women. These markers are not exclusive of gender (women can wear ties, and men can wear dresses), but participants identified these images as being male or female as well. In the counts used in this study I have exclusively used the declarations of participants and merely note here that drawings often, but not always, also included gender markers.

4.5 Chapter Summary

A Mechanical Turk study was conducted to evaluate whether the HCI “people” words were perceived as gendered. Following the statistical testing on the gender, age, and ethnicity data in Chapter 3, qualitative analysis was performed to investigate effect of first language of the participants. Additionally, an analysis of the word usage across conditions and the inclusion of artifacts or markers in drawings draws attention to themes concurrent with the gendering of these “people” words. The following chapter summarizes and discusses the implications of these findings.

Chapter 5
Discussion

In Chapter 3 I discussed the study results regarding the three main hypotheses: that participants would draw males more frequently, that these drawings would be more frequently Caucasian, and that these drawings may be of different ages than the participants. In Chapter 4 I analyzed other factors that could influence participants' responses, namely the presence of gendered nouns in their first languages. I also reviewed their answers to the question of "does the drawing look like you or someone else", shared some initial analysis of the kinds of words being used to describe each condition, and discussed the inclusion of markers (for example, accessories or artifacts) in participants' drawings. In this chapter, I will summarize these results in the same order and discuss the implications of these findings in the wider context of this research and in terms of the implications for actual language use. Finally, the limitations of the study method and analysis are discussed.

5.1 Gender

The revised hypothesis H1 expected that participants would draw any of the "people" words (user, participant, person, designer, researcher) and describe the drawing more frequently as "male". Based on the data received, H1 can be confirmed for drawings by male participants of all words, but can only be confirmed for drawings of female participants for the words "user" and "researcher". The drawings by females for "person", "participant", and "designer" were not more frequently of either males or females.

5.1.1 Use of "Other"

The option to indicate "other" for gender was determined based on consultation of resources for researchers concerned with fostering inclusivity while meeting the overall aims of research (e.g. Balarajan et al., 2011; "Human Rights Campaign (HRC)," 2013; Conron et al., 2008). Due to the low number of responses using "other" (15 out of 757 responses) it was not possible to formally analyze any trends, but its use merits discussion nonetheless. As discussed in Section 3.7.2, "other" was used 13 times to describe the drawing, and only 2 times to describe the study participant. When used to describe the drawing, 10 of the participants indicated that they lived in the United States, 2 lived in India, and 1 was left blank. Both of the study participants who indicated "other" for their own gender lived in the United States. It is a time of high visibility for transgendered, genderqueer, and non-

cisgendered people in North American society and culture, and this may be reflected in Americans' higher use of the "other" option.

As most of the participants who used "other" indicated English as a first language (13, or 87%), it is unlikely that the option was misunderstood. Furthermore, participants provided qualifying words and comments such as "Neither, just a person, as an object, I guess", "just a nongendered stick(hu)man", "stick people are neutral", "I made is [*sic*] asexual [*sic*]. It could be a man or a woman", "Trangender [*sic*]", "I pictured no gender". The participants who themselves identified as "other" wrote "genderqueer" and "ftm".

It is important to note that while non-binary gender identities are gaining more acceptance, many people still feel stigmatized or at risk, and may not feel comfortable indicating their gender identity on an anonymous online questionnaire. Additionally, due to concerns of personal privacy, people may not feel comfortable sharing this type of information regardless. The option to indicate gender in the Mechanical Turk questionnaire was optional, and there is no way to verify the answers received. Participants had the opportunity to leave the question blank, with 17 cases for drawings, and 55 for participant gender. Additionally, there is no reason that someone (for example) born of male sex but with female gender would indicate "other" on a form, because their gender identity is legitimately "female". Our participant who wrote that they are "ftm" could have also written "male", for example. It is therefore not valid to take the "other" results as literal due to the number of outside factors that can influence responses, but it is valuable to understand the contexts in which "other" has been used. In the future, we would likely want to use a 2-step gender question as recommended by Balarajan et al. (2011), where participants can first indicate how they were described at birth, and then they can indicate how they think of themselves. In this particular questionnaire we did not want to draw attention to gender as something we would be focussing on, as we thought this may influence the weight participants give to their answers.

5.1.2 Non-Humans

While non-human results were not included in the overall data analysis, there were some notable contributions worth including in the discussion. There were 11 drawings received of animals. A dog was drawn in the "participant" condition, with the comment "I wanted to draw a dog because they can also be participants in something". In the "researcher" condition, a participant who drew a rat wrote "RAT IS USED IN NUMBER OF CASES OF RESEARCHES [*sic*]". In the cases of the animals, they did draw them "sitting down" as asked. Rats are present in many different areas of research,

though they are not researcher*s* doing any research*ing*, the participant may have chosen a rat as an abstract image symbolizing research as a whole. In the "designer" condition, one participant drew where the designer would work: "I tried to express what I feel is typical designer's workplace look like [*sic*]". This response suggests that to some people, such a role might be defined by the typical environment more than what the person may actually look like. For a "person", one participant chose to write what their name looks like in Tamil. This is an abstract, symbolic representation of a person. These non-human responses introduce alternative ways of representing the "people words", through mostly-valid readings of what participants were asked to do. In Section 5.7 I discuss further symbols and themes found through participants' inclusion of artifacts and markers in their drawings.

Some participants classified their drawings as "non-human" when other participants with similar drawings selected "human", leading to a slight increase in the number of "non-human" results. One non-human submission for "researcher" wrote "I drew a human" when asked if the image looked like them or someone else, and that they considered "age, mood, an overcoat, bottles and tubes to conduct experiment with". Others considered their cartoon or stick-person illustration to be non-human, though it represented a human: "like a person sitting", "a stick figure", "a cartoon character" as "a professor hanging out". Responses were not modified when inconsistencies like these occurred, so they were not counted as human.

5.1.3 Gender Patterns

Part of the motivation for this research was the concept that so long as the feminist subject is conceived within a representational system where they are outside of the norm, and in such a system abstract referents to people tend to be thought of as male, then the feminine cannot be well-represented (Butler, 1990). It was therefore important to determine whether the abstract referents to people used in HCI were perceived as gender-neutral terms. If they are not, then it would draw attention to potential concerns with the current way people are discussed in the literature, and the community can work towards using more inclusive language. If the referents are acting as false universals signifying a masculine worldview, then we would expect the drawings received to tend to be male. The results received were more complex, with males across all conditions drawing males more frequently, but with females drawing males more frequently in the "researcher" and "user" conditions only.

"Participant", "designer", and "user" have not been tested before in this manner, but previous incarnations of the Draw-A-Scientist test indicate to some extent what we might expect from the

"person" and "researcher" conditions. The "researcher" responses contained many markers consistent with the stereotypical image of the "scientist" from numerous other studies: lab coats, beakers, glasses, pens in the pocket, stacks of books, associations with academia, and so on. Finson (2003) indicates that it is the norm for people to draw an image matching themselves when asked to draw a "person", and that the same does not hold true for "scientist". Thus this may contribute to the pattern observed whereby drawings of "person" by females did not tend to be more frequently of males or females, whereas for "researcher" females did draw males more frequently.

It has been asked in the past whether what people draw is what they see in their head and actually indicative of some bias, or whether it is a reflection of what they see in the world. Drawings show us "not the thing actually seen but its representations in human consciousness", which is informed through history, culture, and codes (Türkcan, 2013, p. 601). Mass media is a source of gender stereotypes that primes people in subtle ways (Steinke et al., 2007, p. 41). Regardless of whether a tendency towards drawing male "users" is influenced by media stereotypes or some other source, we have already seen that gender stereotypes influence what designers produce based on assumptions they make about their audience, and that when a falsely universal term with male connotations is used, the assumptions made for a female audience no longer appear (Huff & Cooper, 1987). How people perceive these "people" words does make a difference as it does impact our work.

The results of this study are significant in that they demonstrate that males and females may not perceive HCI's people words in the same way. Stereotypes associated with roles and occupations form at a young age, as do understandings of "feminine" versus "masculine" values. These values become associated with roles and occupations, with STEM disciplines being typically viewed by children as stereotypically masculine pursuits (Steinke et al., 2007, p. 37). It has been shown that when considering future career paths and aspirations, young girls may not consider STEM disciplines as they appear incompatible with these feminine values (2007, p. 37). Females may therefore become less interested or disengaged if they feel excluded from the worldview of the domain. Dray et al. (2013) assert that it is "imperative that both masculine and feminine perspectives are included in shaping our modern day technologies" given that "technology has a profound mediating effect on the way we relate, obtain knowledge, and contribute to society." If women are unable to align themselves and their values with those that they perceive in HCI, we will fail to engage them in shaping new technologies.

In Chambers (1983)'s Draw-A-Scientist test, "person" was used as a control, which showed that "markers" tended to be absent from this condition whereas there were many markers (equipment, accessories, setting, etc.) present in the stereotyped condition. One could argue that the "researcher" condition has a lot of history, high visibility in the media, and therefore several markers shared with the "scientist" that create associations with a masculine domain. Participants indicated things in their "researcher" drawings such as (super)computers, books, beakers, microscopes, test tubes, diplomas, goggles, and experiments. They referenced roles such as: scientists, professors, teachers, academics, analysts, astronomers, technologists. They made wider references to domain such as: academia, chemistry, science, medicine, biology, education, engineering, geology, math, pharmaceuticals, statistics, environment, politics, and advertising. The role of the "researcher" is embedded in society. The role of the "user", on the other hand, is a bit more ambiguous. The word "user" was appropriated in the 1960's to refer to a computer user ("user, n.1.," 2014). Though it has less history, the results indicate that a pattern of association of objects to a "user" has also begun. Markers indicated in the "user" condition include: computers, desks, laptops, keyboards, monitors, mobile/smartphones, tablets, PCs, apps, gadgets, games, prototypes, and software. A more complete discussion of word usage across all the conditions can be found in Section 5.6, however it is important to highlight the convergence of markers that are appearing here. Objects are being used to assist with defining what a "user" is, which seems to be someone interacting with technology. The HCI community should be aware not only of the gender perception of the "people" words but the networks of objects and settings that define these words: gender and identity are relational and situational. The drawings received highlight contexts associated with masculinity that researchers should be aware of when situating the "people" words in the future.

Bodzin & Gehringer (2001) show that an intervention as simple as real-life exposure to someone in a stereotypical role can reduce the number of stereotyped images produced: after meeting a female engineer, students were more likely to draw female scientists, and after a scientist visited the classroom, drawings were less likely to include things like glasses and lab coats. It has, however, also been shown that as people encounter new information that may change or alter gender stereotypes, those with more rigid gender schemas tend to "alter or forget" information inconsistent with those schemas (Steinke et al., 2007, p. 41). This supports the idea that having a token minority member of stereotyped groups is not sufficient to alter perceptions: a more conscious, intentional effort is required to dissuade the association of "people" words that the HCI community would like to see as neutral with non-inclusive gendered value sets.

We must be careful in interpreting these results to check assumptions about heteronormativity: this narrative causes gender and technology coproduction in society in a chicken-and-egg manner, where masculinity causes technology that is masculine, and so on (Landstrom, 2007). When this assumption is also present in research, it stands in the way of adopting an approach that would actually support the goals of feminist research: it treats gender as stable and technology as malleable (2007, p. 7). The design of the study was intended to allow for non-heteronormative responses, but very few took that opportunity. The drawing of one of HCI's "people words" creates a snapshot in time, and cannot capture the changing self-concept as Landstrom suggests. It is a challenge to reverse this tendency in order to begin seeing gender as malleable and technology as stable. This project presents a first step in understanding how the general public perceives HCI's "people" words and does not tease out the relationship between subject and object extensively. Now that we have seen some of the prevalent themes associated with what it means to be a "user" we can move forward with understanding the effects of technologies on our gendered self-perception and vice versa.

5.1.4 Implications

While "other" was used intentionally by some participants in order to be neutral, studies show that blindly defaulting to neutral terms can lead to presumptions of masculinity. Researchers should pursue the motivation that we saw with the participants who endeavoured to show inclusive or non-binary subjects in their drawings: intentionally considering a spectrum, as opposed to using a neutral term to fit a universal. Future work should also consider the effect of objects and contexts on gender perception, and the situational as opposed to static nature of gender identity: is a drawing presumed to be masculine because its context influenced it to be that way, or is it drawn to that context due to its masculinity? The malleable nature of identity is not captured in static case studies but in interactions unfolding over time. The consistent influence of positive and diverse role models is shown to be a powerful way to change perceptions. Similarly, the use of a diverse set of personas could function to actively combat subtle assumptions during research and design.

5.2 Ethnicity

Finson (2003) demonstrated that tests constructed like the Draw-A-Scientist test (DAST) can be used with some confidence with participants of different racial groups when attempting to reveal stereotypical perceptions. Furthermore, the author found that images produced by different groups

were similar, with no significant difference between the groups' DAST scores (2003, p. 23). Considering this combined with the pattern of "globalizing" imagery found in drawing-based research on Mechanical Turk (McMaster, 2012), it was hypothesized that a similar trend in the ethnicity of the drawing would persist across all ethnicities in the sample. Specifically, hypothesis H2 posited that participants would draw any of the five "people" words and describe the drawing more frequently as "Caucasian". There was insufficient spread in the data to reach formal conclusions based on the responses received, however there was a tendency for Caucasians and South Asians to draw people of their own ethnicity, and for other ethnicities to draw Caucasians in addition to their own ethnicity: 29.50% of the overall drawings were by Caucasians and 41.53% were by South Asians, but 35.05% of the drawings received were of Caucasians, and 35.05% were of South Asians.

In an implementation of the Draw-A-Scientist test by Ford & Varney (1989), out of 1600 drawings by both girls and boys, only 20 of the scientists were people of colour. However, according to Finson (2003), it is the norm for people to draw an image matching themselves when asked to draw a "person", but the same does not hold true when asked to draw a "scientist". Thus we might expect that, given a larger and more diverse sample, there may be variability in the results across the five conditions, that is, whether the "people" words are drawn with similar or different ethnicities to the participants.

While there have been studies of race as a factor in the Draw-A-Scientist test, these studies have tended to be constrained within one country or continent, thus introducing racial diversity within the umbrella of a common culture (e.g. Chambers, 1983; Finson, 2003; Fort & Varney, 1989; Miele, 2014; Steinke et al., 2007). In an American study, Sumrall (1995) found that minority students tended to draw scientists who were not minorities themselves (African-Americans drew Euro-American images). There were no limits on the location of participants in the Mechanical Turk study, thus the sample spanned many countries across the world. The different ethnicities that partook in the study are not minorities in their home country, so Sumrall's findings are not transferable to this work.

Finson (2003) found that members of different ethnic groups produced drawings that were similar and attributed this in part to common exposure to media influences. Finson's racial groups were exposed to American media and the effects of pursuing STEM disciplines as African-Americans or Native Americans: they created similar drawings, even though their racial identity lead to different lived experiences in this domain. Due to the geographical diversity of the participants of the study presented in this thesis, a confounding effect is introduced whereby racial or ethnic identity exists

alongside different, not necessarily shared cultural narratives that may or may not frame HCI's "people" words similarly to their interpretation in other locations. In this study, participants reported living in India, Venezuela, Lithuania and beyond; it is not possible to control, account for, or predict the effects of media influences this widely-reaching. Additionally, cultural factors running deeper than media may affect responses: Monhardt (2003) received a higher-than-usual proportion of female drawings from Navajo children, and one possible explanation put forth was that traditional Navajo society is matriarchal, with females occupying positions of power.

While the original research question of this thesis was to determine whether words used in HCI papers are perceived as gender neutral, one of the outlined objectives was to evaluate the perception of these words overall based on the results received. Prior to this study, research around ethnicity and perception of words had been completed for "scientist", but not for the HCI people words (user, participant, person, designer, researcher). With the theoretical framework provided here, we can begin more directed efforts to understand the interaction between ethnicity and HCI's people words. It is already known based on past research that different ethnic groups perceive "people" words differently, and that cultural knowledge influences their interpretation of these words. In a global community of academics and professionals it is easy to forget that our words will be read under a variety of linguistic and cultural codes, thus purposefully vague gender-neutral terms may not elicit the same mental images as intended by their author.

From the associated research we now know that some racial groups may be more likely than others to insert themselves into a role, and some roles may be more likely to be identifiable than others. Role models, the media, cultural traditions, and exposure through everyday experiences all play a big part in helping different groups feel connected to "people" words.

5.3 Age

Drawings were of people younger than themselves when participants drew "participants", and older than themselves when they drew "researchers". For "user", "person", and "designer", the age of the drawing was not significantly different than that of the participants. This suggests that participants perceive researchers to be older, participants to be younger, and the remaining three conditions to be of similar age.

Steinke (2007) cites that older students were less likely to draw female scientists than younger students in the Draw-A-Scientist test; this is related to the establishment of more rigid

understanding of gendered occupations and values over time. The contents of drawings were not tested across different age groups in this study due to a lack of appropriate spread in the participants' ages, and the minimum age for participation on Mechanical Turk is 18, at which point these gendered concepts tend to be well established.

Miele (2014) found that the results of a similar test done among adults were very similar to when children did the same test, so it is not of value to dwell too much on the age of the participants and when certain gendered imagery starts to appear – many instances in the Draw-A-Scientist literature have established this. Instead, the ages of the drawings help us to understand what is implied, if anything. One participant wrote about their 35-year-old researcher drawing: "I thought about age simply because I am not going to draw a child when I am drawing what I am thinking of as scientists. I also did not want to give it an old stereotypical professor age".

The average age of the researcher drawings was 38, seven years higher than the average age of the participants, suggesting that there may not only be a gender bias associated with this word but age connotations as well. In the participant condition, the average age of drawings was 27, four years younger than the average age of those that drew them, indicating a different interpretation which may indicate a belief that this group is less mature than one's self. The average user was shown as 28, but this result was not significant. Age may be linked to a perception of control or agency, as a researcher is seen as independent but a participant is dependent on someone else for direction. Further study would help elucidate the meaning of these findings.

5.4 Language

Boroditsky et al. (2003) found that when people with a first language with gendered nouns were asked about nouns in English, their first language influenced how they categorized the objects in terms of masculine and feminine attributes. They even demonstrated that outside of cultural influences, grammatical gender was enough to influence people's thinking about objects. It was therefore of interest to investigate the first language of study participants to determine whether a similar effect could be present. Of the 756 participants, 344 cite English as their first language, the next most common language being Tamil with 108 native speakers. Native English speakers drew significantly more males than females (roughly 74% males and 26% females) despite a lack of grammatical gender. Among Tamil speakers, significantly more males were drawn than females once again, about 84% and 16% respectively. Tamil does have grammatical gender, but when not referring

to a specific person a neutral form is used. In Bengali speakers, a language without grammatical gender, there was an equal split between male and female drawings, but with a sample size too small to be conclusive. In Hindi, where all of HCI's people words are grammatically male, drawings were approximately 73% male and 27% female, but the sample size once again was very small.

Through languages with and without grammatical gender, the drawings of HCI "people" words show similar trends towards being male, with the exception of Bengali. Boroditsky et al. (2003) made statements about the influence of grammatical gender on representations after a series of many studies with various factors being controlled. There are insufficient controls in place here to come to similar conclusions after just one study, but it is possible that an effect could be present when conditions are tailored towards a study of grammatical gender, as past work has shown that grammatical gender does affect meaning (Konishi, 1993). The assignment of grammatical gender gets better and people become more sensitive to it with age (Magnan, 1983). As the average age of participants in this study was 31, patterns from Draw-A-Scientist tests cannot be used for comparison as it is often done with children. Additionally, while Draw-A-Scientist tests have been performed internationally before (with similar results in other countries), none have considered the interaction between first language and gendered representations. It would be advantageous to narrow the focus and study the effects of grammatical gender in isolated language groups rather than the large spread received here, to better control the study.

Grammatical gender is far more complex than whether a word is "male" or "female". Some languages include a neutral term that is used when the gender of the referent is unknown. In French, some words default to masculine (i.e. *un participant*) unless they are purposely made ambiguous to be inclusive (i.e. *un(e) participant(e)*). Some Dravidian languages have developed such that the feminine and neuter genders have coalesced so that masculine is opposed to non-masculine (Corbett, 1991, p. 10). Research on this topic requires not only a gender-diverse pool of participants and researchers, but also culturally diverse researchers that are aware of and can adequately detect these nuances.

Although the papers published at CHI, the premiere conference in HCI, are required to be in English, the audience is worldwide. Gross (2014) relates the internationalisation and subsequent diversification of HCI with the divergence of HCI and human factors. In other words, HCI's diversity helped it become the distinct field it is today. Gross also highlights that this diversity is not seen as a positive by everyone as its research strands are no longer as unified, leading some to call to action for "unifying theory and practice and standardising teaching curricula" (2014, p. 187). Different

languages have different grammatical genders associated with the words encountered in HCI papers, and will affect how these "neutral" words are interpreted by different groups of non-native English speakers, despite how much unification and standardization is done. It is therefore fallacious to assume that the ambiguity associated with the HCI "people" words carries across linguistic groups, because while English itself is non-gendered, it is first languages that have been shown to influence meaning.

5.4.1 Implications

Some authors call for establishment of more standardized theory, practice, and curricula in HCI due to how international and diverse the discipline has become. Past research has shown that grammatical gender in someone's first language can influence representations, and while an effect was not detected in this study, I found that the HCI "people" words do have different genders in other languages. Therefore, when we use these words with the intent of being neutral, they may not be interpreted that way by other readers in the wider HCI community. The HCI community should be aware of how its vocabulary may be (mis)interpreted so that communication can be clear among its diverse members. Additionally, context could be included in HCI writing (i.e. "male and female users") to negate some of the ambiguity of standalone words such as "user" or "participant" that are neutral and non-gendered in English but may not be in other languages.

5.5 Looks Like You or Someone Else

The third question that participants were asked about their drawings was "*If you drew a human, does the person in your drawing look like you, or someone else?*" Answers were given in free-form. As a result, responses were highly varied: participants could say it looked like them, someone else, nobody in particular, or someone they know (who may or may not be like them). For the "user" condition, responses were most likely to indicate "someone else" (33.3%) followed by "like me" (20.4%). For "participant", "someone else" was again most likely (38.7%) followed by "like me" (20.4%). For "person", answers were 33.1% for "someone else" and 18.2% "like me". For "designer", responses were most likely "someone else" (38.0%) followed by "like a typical designer" (12.4%). For "researcher", drawings were 30.1% of "someone else" followed by 19.6% of "like a typical researcher".

"Designer" and "researcher" were the only two conditions where participants tended to describe them based on what the condition was called. For example, none of the participants wrote

something like “it looks like a generic user” or “it looks like a typical person”, but 19.6% of the “researcher” respondents wrote that “it looks like a generic researcher” or indicated some stereotype (“it looks like a scientist”). For “designer”, these stereotypes tended to be things like “it looks like a fashion designer”. As “researcher” and “designer” are both professions, they have connotations unto themselves that “user” “participant”, and “person” do not.

Miele (2014) found that as students became more comfortable with and confident in STEM, they became more likely to identify their drawings as of “you,” “me,” “students,” “teachers,” and “my sister.” Previous evidence by Finson (2003) supports the finding that people were more likely to draw someone “like me” in the “person” condition (18.2%) than the “researcher” condition (11.9%). This does not, however, account for other ways in which participants could imply that their drawing was similar or different from themselves, for example by answering that it looks like someone from their family or like a friend. These people might be very much like them but, as someone else, act as a useful mental reference for the drawing. Furthermore, due to the wording of the question, participants may have different interpretations of what qualifies as “like them” or not. Some may feel that “like them” implies that the image must look exactly like them or be a picture of them, whereas others may interpret “like them” as someone who shares ethnicity, hairstyle, or background. These initial findings follow the pattern found in previous research of drawing someone closer to you when feeling more related to the role. Future research should refine how this question is asked alongside data collected to make it easier to gauge whether and how drawings are *similar to* or *different from* the participants.

According to past research, white males tend to have a higher percentage of self-image drawings than other ethnicities, which some have related to an internal locus of control (Finson, 2003, p. 17). There were 102 drawings of male humans by white males, and 13 drawings of females. 16 (15.6%) Caucasian male participants made an indication that their male drawing was “like them”, however 80 of these 102 drawings (78.4%) had both gender “male” and ethnicity “Caucasian”. Caucasian males drew 13 females in total, 10 of these were Caucasian. For one of the non-Caucasian female drawings, for the “person” condition, the Caucasian male felt the need to justify his choice: “The world today is largely populated with Asians. So if you choose one person randomly, good chances he/she will be Asian”. Caucasian females submitted 113 drawings in total, 72 of them were of males, and 5 of those were non-white. They drew 36 drawings of females of which 4 of them were non-white. Interestingly, 10 participants (27.8%) indicated that their drawings of females were “like

them", a much higher proportion than their male counterparts. This is a disagreement with the previous work, and one that should be investigated further.

The implication that someone feels a user, participant, person, designer or researcher could be "like them" was an important distinction to make. Interacting with technology is a performative act, and we know that "social reproduction and subject formation take place through (largely unquestioned, but not necessarily faithful) reiterations of existing forms" (Van House, 2011, p. 423). Furthermore, someone's perception of their ability to do something (or to be something) plays "a more important role in determining persistence than actual ability" (Knight & Cunningham, 2004, p. 8). A failure to identify, then, first suggests that the person is performing the role of another, and has a lower likelihood to engage with HCI altogether.

Self-perceptions are powerful influences on the choices we make (Finson, 2003, p. 24). For groups such as minorities and females, Hill et al. (1990) cite no factor more critical to their pursuit of STEM than their self-efficacy, strengthened through exposure to appropriate role models – even in comparison to other traditional barriers. Images in the media influence perceptions and self-perceptions in regards to ability to succeed in that role: it can be a source of positive inspiration, or one resulting in stereotype threat (Steinke et al., 2007, p. 41). In Chapter 2, I discussed the concept of self-efficacy and how some feminist scholars sought to redefine the trio of competence, autonomy, and relatedness, the cornerstones of self-determination theory (Ryan & Deci, 2000). The feminist concept of the self is both socially defined and relational; it is neither abstracted nor independent. Asking study participants to draw "a user" was asking them to portray something objectified and abstracted from the social context in which it is embedded; drawing them alone was to put the drawing into a social situation that was predetermined as more socially acceptable for a masculine worldview. Drawing "a user" leaves no room for understanding the drawing in terms of *relational autonomy*, or a new social arrangement that has room for dependency as something that can be valued, and identities that are socially and historically embedded, and shaped by factors such as race and class (Stoljar, 2014). Past studies show that when drawing multiple scientists, the stereotypes evolve from drawing to drawing, for better or for worse (Chambers, 1983). It would therefore be productive to start thinking of the "people" words not as one but as multitudes, so that the stereotypes begin to fall away or mature: Symington & Spurling (1990) suggest that if only given the opportunity for one drawing, people will try to include aspects of the public stereotype to make it recognizable, and Chambers (1983) found that as people produce additional drawings they reveal feelings about the

domain that were not apparent through the first (in that specific case, it was ambivalence about the social value of science). Further, thinking of more than one would create space for the relational concept of the self rather than the atomistic person. As a next step, adding a time dimension and drawing two users *doing* something will provide valuable insight into what extent the participants feel technology is defining the drawing.

In yet another instance of the Draw-A-Scientist test, Buldu (qtd. in Steinke et al., 2007) found that when children said they wanted to be like the scientist they drew, the person they drew was someone from TV or someone they knew. In this study, participants weren't asked whether they *wanted* to be like their user/participant/person/designer/researcher – just whether they thought they were. As much as exposure to role models is crucial in developing women's self-efficacy so that it is easier to see oneself in a role (with stereotype threat as its complementary dark side), the recent very-public vilification of people like former Reddit CEO Ellen Pao and media/video game critic Anita Sarkeesian not only driving women to not want to be "like them" but to be afraid of being like them.

5.5.1 Implications

A feedback loop exists where participants are more likely to identify with one of the "people" words when they have increased confidence in that area and increased identification raises feelings of self-efficacy. There is no hard-and-fast way to determine whether participants identified with their drawings due to the open-ended nature of the question, so further research would do well to be explicit about this in the future. Imagining the "people" words in situations where they are alone could influence how gender is perceived as well as the emotional state of the drawing, which may make participants feel more or less inclined to identify with them. Researchers should be aware that social situations perceived as positive for males may not have the same implications for females, and should imagine both solitary and social contexts of use. Deliberately imagining more than one "user" or "participant" may help researchers when trying to account for potential differences in a spectrum of people rather than an aggregate of many which would mute those differences.

5.6 Word Usage across the Conditions

In the "user" condition, there were frequent references to objects in relation to the person drawn, especially to consumer electronics such as computers, laptops, phones and video games. A portion of drawings depicted drug users. The "participant" condition elicited references to game shows, competitions, contests, and participating in medical or other research. The "person" condition

included references to people encountered in everyday life, and aspects of physical appearance such as hair and outfit. The drawings submitted for "designer" were described with words associated with fashion, graphic design, the arts, creativity, or more industrial applications of design such as architecture or cars. The "researcher" condition elicited strong connotations to academia, science, and the pursuit of knowledge. This condition, like the "user", also had references to technology.

When I investigated more closely the context surrounding the use of these words, participants were found to express frustration at the (purposely) vague drawing task. One wrote: "I got really frustrated trying to work out what KIND of 'user' you meant, so the person in the picture is the one that came to me the most strongly thinking of that word in use: a heroin user". This happened more than once: "What kind of user did you mean? Aha! A drug user. A junky!". When the participants themselves tried to be purposely vague, their responses started to form a pattern as well. There were numerous cases where participants stated that they started out drawing something generic, but ended up with a white male.

To begin to explain possible underlying causes for these patterns, I return to the idea of schemas. People use extensive networks of gender schemas developed at an early age to organize and understand information, and apply them to things that they encounter in everyday life (Steinke et al., 2007, p. 40). This organization of experiences into causal relations is the process by which we establish a semiotic view of the world (Gaines, 2006, p. 175); organizing our drawings and words into thematic relationships paints a bigger picture of our semiotic understanding of what it means to be a user, participant, person, designer, or researcher. This is precisely why both drawings and words were collected, they show us "not the thing actually seen but its representations in human consciousness", which is informed through history, culture, and codes (Türkcan, 2013, p. 601).

The original research question was whether the people words used in HCI were perceived as gender neutral. We know now that there is a complex interaction of longstanding schemas as well as exposure to prior examples or role models that may influence what one individual draws, however the goal of the study was to uncover broader themes that speak to a cultural tendency as opposed to a specific bias. As our language contains more stereotypes than the things we draw (Sontag, 2001), it was important to analyze words from this higher level rather than individual cases.

The English words analyzed take into account all responses by participants in a condition regardless of first language, although the "people" words could be gendered in their first language. Past research on grammatical gender and meaning by Konishi (1993) shows that grammatical gender

of an object in someone's first language affects the group of words that the person chooses to attribute to it, words perceived to be in its semantic network. The language effect therefore does not necessarily stop at whether they perceive their drawing as being more masculine or feminine, but extends to the types of words they attribute to it in their description as well, which rolls into the objects, artifacts, markers or setting chosen in the depiction.

Hiltz et al. (1986) found that the anonymity of people in a group online caused them to act in a more "normative" manner based on the group composition. It is not a requisite for participation that Mechanical Turkers communicate with each other, as the platform is based on them independently completing micro tasks. Communication between Turkers does happen however, but not to the degree that we would expect to see normative communication habits emerging similar to Hiltz. Emerging normativity in anonymized communities is, however, something to consider in future studies when looking at language use. Though I did not expect a particular "style" of communication from Mechanical Turkers and therefore that they would respond in a way that felt authentic to them, I chose not to analyze language use separately across genders, for example whether males and females used different clusters of words to describe their drawings in the conditions. In Section 2.3 I discussed the implications of classifying language patterns as "male" or "female" when conducting research. It goes against the feminist research goal of treating gender identity as malleable and treads into essentialist thinking. Furthermore, as discussed in Section 5.1.1, self-reported gender may not be consistently reliable. Characteristics of "masculine" and "feminine" language use vary widely across cultures, and the participant population for this study also was very diverse, thus no valid claims can be made about how males or females chose to describe their "people" words. Patterns cannot be identified to describe the *participants'* behaviour, but patterns in the words have been highlighted in relation to the *drawings* as a whole and not split by gender.

5.6.1 Implications

The results of this study allow us to begin understanding what the general population associates with HCI's "people" words. Specifically, we have started to see that words like "participant" and "person" have more ambiguity and diverse themes than words like "designer", "user", or "researcher", and that the meanings that the general public associates with these words do not necessarily map to or reflect those that the HCI community may be implying when using the "people" words in its own writing. The community should be aware of unintended implications when using these terms; even if the general public is not the primary audience, before each HCI researcher was a specialist they were a

part of the general public too. This writing sets the tone for those developing schemas that could be brought into the community in the future: social arrangements are "sedimented out of particular practices that we have a role in shaping" (Barad qtd. in Suchman, 2007, p. 11), and critical engagement with language practices in HCI is one way to evaluate the effects engendered by these practices.

Evaluating the spectrum of words recorded about participants' drawings revealed a network of meanings in relation to HCI's "people" words. These broad networks are part of how participants establish a semiotic view of the world, and they are a method by which gender schemas become associated with certain roles and behaviours. Even if the "people" words do not appear to figure definitively into gender schemas (i.e. "participant" was not drawn as more frequently male or female by female participants), we have a rich set of terms that provide more context around how these words are perceived. These *terms* may well fit into participants' gender schemas surrounding the "people" words. The HCI community should be aware of how these terms have emerged as schemas used by participants to contextualize typically "male" or typically "female" drawings, so that we can extend our understanding to not only how the "people" words are perceived but how the incorporation of these networked terms can influence *how* they are perceived. Perhaps the fact that "user" is perceived as more frequently male by both males and females is not the issue, but that the schema that "user" finds itself a part of has been typified as "male" – thus future contexts that a "user" finds themselves in make them more likely to be perceived as masculine.

Finally, a phenomenon was observed where participants defaulted to drawing a Caucasian male after it became too tricky to depict something else. Such drawings are false universals in action, failing to acknowledge the presence, influence, and oppression of those absorbed into its self-definition: they use circular logic that "reduces differences to sameness and universalizes and totalizes by substituting part for the whole" (Carnegie, 1998, p. 7). Through identification with the falsely universal term, bodies become "sexually undifferentiated" and cannot be interrogated outside of a model where they are unwoven and shared with "man" (1998, p. 7). That participants default to the normative state when all else fails speaks to an uncomfortable tendency that we should all be aware of.

5.7 Artifacts and Markers in the Drawings

Of all the artifacts that participants said they included in their drawings, chairs and tables were among the most popular, which was to be expected as they were asked to draw the condition sitting down. While chairs were the seat of choice for many drawings, the only time someone was seated in a wheelchair was in a drawing of Stephen Hawking for the "researcher" condition. No other drawings in any of the conditions had any descriptions of disability. Computers and other items such as smartphones, tablets and laptops were present most frequently in the user condition, followed by researcher, designer, person, and participant. Drug use equipment was found in a portion of the "user" condition, suggesting that the definition established in 1935 of the word "user" to describe a drug taker is still very much alive and may be imagined by readers. The word "user" was appropriated in the 1960's as a go-to noun for the emerging computer revolution, and this study shows evidence that the older meaning is still implicated by the use of this word. The inclusion of markers was compared between the "person" condition and the "researcher" condition to evaluate consistency with previous literature. It was also observed that in many cases of female drawings, participants included markers to indicate non-male gender.

Theories about "the nature of biological existence, about language, and about the nature of human action" deeply influence what we build and how we use it (Winograd, 1986, p. xii). We begin to associate certain things with masculinity by grouping them with other things encountered in causal relationships. This is how gender schemas are formed over time and, unintentionally, these concepts that are perceived as related begin to represent the role itself as being, for example, masculine. Images become metaphorical, transcending their origins, equating one concept with another: although the "concepts are theoretical, the implications are concrete" (Knight & Cunningham, 2004, p. 2). When these ideas become normative, the language formation begins to produce subjects who are presumed to be masculine, and it becomes more challenging if not impossible to make feminist progress within this framework (Butler, 1990, p. 2). It is critically important to understand how the repeated use of HCI's "people" words may naturalize certain assumptions about who they represent. Exploring the artifacts or markers associated with these words helps us grasp the gendered contexts that these "people" words imply.

Chambers' (1983) Draw-A-Scientist test found that the "person" control tended to contain few-to-no markers indicating that this was a "person", whereas the "scientist" contained on average 4-5 markers when tested with adults. This finding was consistent with the results found in the "person"

and “researcher” conditions of this study. Given the rich set of data and strong themes associated with the “user” condition, we may be able to determine through further study what markers of a modern-day “user” may be. The users shown in the drawings frequently were alone, indoors, with at least one piece of technology, usually a desktop or laptop computer. They were more frequently male when drawn by males or by females. 39 out of 148 “user” drawings (26%) were explicitly described as having a computer, a phone, a tablet, or some combination of these, and 9 (6%) were of drug users. For comparison, in the “researcher” condition there were 64 drawings (47%) that were described as having some combination of the “scientist” markers. Some of the variation in the drawings could be attributed to the fact that participants were not told what the drawings were for. Symington & Spurling (1990) suggest that if people think the task is to draw something that is recognizable as that thing, they will include items in the drawing that they believe to be part of the public stereotype of that thing. Without getting specific about what we wanted from the drawings in order to get participants’ first reactions, we were able to mitigate some of the immediate stereotyping through asking participants to draw their condition sitting down. Imagining the condition sitting down for 10 seconds before drawing helped to ground the image from a concept.

In Section 2.1.2, I wrote that the design of technologies can either support or blur traditional gender roles, and that researchers need to think about the ways their own work does this and be accountable for it. In the example of the Hoosier cabinet, I illustrated how technologies meant to liberate women from certain roles ended up, through use over time, actually identifying them with it. In the same way, objects can become associated with masculinity over time: the *Significant Screwdriver* project sought to transgress social norms around the gendered division of labour (Bardzell, Gross, et al., 2011). There is a strong pattern of markers in the “researcher” condition, which was also perceived as masculine by both males and females, congruent with the previous work in this area. This study shows that a pattern of markers may also be emerging around the term “user”, which was also perceived more frequently as masculine by both males and females. When the role of technology is emphasized in the identity of a “user”, it follows that the technology is providing the means to “constitute social reality, through language, gesture, and all manner of symbolic social sign” (Butler, 1988, p. 1). The HCI community should be aware of the gendered associations growing between terms like “user” and the fruits of their labour, new technologies, before social norms around the use of these technologies become sedimented such that they must also be transgressed in a CHI paper of the future.

Approaching this problem through feminist theory means that "user" does not have to mean a person (subject) using technology (object), in fact, a goal is to understand our relations to the material other than via the subject/object distinction (Suchman, 2007, p. 9). Such boundaries as those between human and machine are constructed and are "necessary for the construction of meaning" but are also "never innocent" because they are positioned by an agent rather than naturally occurring (2007, p. 11). Following Donna Haraway's intervention regarding the 'cyborg', feminist scholars have "begun to embrace the inseparability of subjects and objects, 'natural' bodies and 'artificial' augmentations" (2007, p. 2). Instead of rejecting the cyborg as a symbol of masculinist and military projects, she suggested reclaiming it for its life-affirming and enabling possibilities (2007, p. 2). Coleman (2011) argued that "user" implies "dependency, addiction, and, ultimately, objectification where one is conscripted into the logic of whatever is being used, be it a technology, drug, etc." We did see a number of drug users who had lost control of their lives. We also saw a large number of technology users, but were they as helpless? As long as we read interactions through the lens of masculine rationality, that is, with an implied subject/object distinction, then one side will continue to be perceived as the privileged reference. This research sought to understand "the effects of particular assemblages" and assess "the distributions, for better and worse, that they engender" (Suchman, 2007, p. 11). It showed that through the traditional lens, yes, "users" are perceived as being predominantly masculine, and within that logic, they are disempowered. Recall that "the conceptual frameworks that guide the organization of official knowledge in our culture tend to reflect the experience and interests of men, particularly racially and economically privileged men" (Sprague & Hayes, 2000, p. 674). Having technology play a role in self-identification is a hit to the traditional masculine concept of autonomy, which highly values independence. A feminist concept of the self *accepts* dependency and revalues it rather than allowing it to devalue the self. Today this is an alternative interpretation of what it means to be a "user" rather than the dominant reading. The more solidified that the schemas become that we saw associated with "user", the more entrenched its inherent masculinity will become.

5.7.1 Implications

Drawings submitted by participants showed artifacts and markers consistent with the development of schemas used to classify a condition. The relationship between a "user", masculinity and technology continues to imply a lack of agency so long as it is read using the masculine frame where a loss of independence has a negative impact on our perception of value. Within the feminist frame,

dependency is valued as a part of relational autonomy. If being a "user" continues to be associated with masculinity and a certain manner of constituting social reality, then it will become more difficult for anyone outside the normative group to identify with new technologies targeting a neutral, universal "user", as this term is no longer universal (if it ever was). In systems where abstract referents to people tend to be thought of as male, the feminine will never truly be adequately represented. Judith Butler argues that the relation between masculine and feminine cannot even be represented in a signifying economy in which the masculine "constitutes the closed circle of signifier and signified" (1990, p. 11). Demonstrating that produced subjects are presumed to be male problematizes the discursive structures in place in HCI, invalidates those terms being thought of as universal, and should motivate linguistic change to become more inclusive. The term "user" in its current form evokes connections to a masculine network of concepts and objects that the community should be aware of in order to explore how a "user" may be situated differently in the future.

5.8 Limitations

As with any user or experimental study, there are limitations. The following is a discussion of the limitations of the study method and the methods of analysis available to apply to this study.

5.8.1 Limitations of the Method

The study asked a wide-ranging group of participants about five of HCI's "people" words (user, participant, person, designer, researcher). While HCI may have relatively more women than other branches of computer science (Dray, Peer, et al., 2013), HCI, computer science, and STEM in general remain male-dominated fields. The literature on stereotype threat states that when members of a marginalized group perceive that they are stereotyped, "they run the risk of confirming that stereotype, and this can negatively influence performance on tasks related to the stereotype" (Steinke et al., 2007, p. 41). In other words, if female participants already feel excluded from one of these "people" words, they may end up confirming that through what they choose to draw.

Participants may have had multiple concepts or more complex understandings of a word than one drawing could permit. Similarly to Knight & Cunningham (2004), designing the study to allow for a combination of drawn and written responses allowed participants to express additional ideas that they wanted to be captured.

One of the final questions in the study was voluntary disclosure of the industry in which the participant works. The question was: "*Sometimes people who work or study in a certain industry will*

think about concepts in a similar way. In what field/industry do you work/study?" The motivation behind this question was that if there happened to be a large number of participants who worked in software development, for example, they may be more likely to draw a computer user. To cover a broad spectrum of possibilities, there were 52 options to answer this question (one being "Other"), which made it difficult to analyze. Studies have shown that people from both biology and liberal arts backgrounds tend to exhibit the same gender biases when completing the Draw-A-Scientist test (Steinke et al., 2007, p. 38), therefore an evaluation based on participant field may not be valid. A more appropriate way to explore how occupation may affect perception of the "people" words is to compare samples from two specific fields in the future, now that a general broad baseline has been set.

Finally, one of the challenges with the chosen method is determining whether using words from a CHI paper out of context can lead to insights that are meaningful back in the setting whence they came. It would not be possible to evaluate the mechanisms behind HCI's "people" words solely within the HCI community without some sort of baseline for comparison in order to understand what was being observed (do HCI experts perceive "user" as being predominantly male more or less than some average?). This study provides that context to compare. Additionally, assuming that these results do not apply in the HCI community simply because the participants were not limited to HCI practitioners is fallacious. HCI's language does not exist inside a bubble free of outside influence and with its own unique meanings divorced from longer-standing etymology. The HCI community is as diverse as the participants in this study, and it is relatively young as a discipline. Furthermore, the general public is often affected by new technologies put forth by the HCI community, and may choose to engage with the community through participation in research, through study, professionally, or on a consumer level. Divorcing the public's perception from the community's does not afford them the respect they deserve; the psychology community's choice to avoid this disregard contributed to their decision to elevate the wider population by referring to them as *participants* rather than *subjects*. In Section 1.1 I stated as a motivation that language producing subjects presumed to be masculine would show the discursive structure is problematized. By producing subjects presumed to be masculine, I have problematized HCI's discursive structure. Next, we can work in a more community-focussed way to determine how our language can be more inclusive.

5.8.2 Limitations of the Analysis

The participants in the study all used a basic drawing tool, where they could draw only black lines on a white background. They had the option to switch to a tablet or other input device if they felt more comfortable drawing in a different way. Since the drawings were in black and white, things such as ethnicity or other factors could not be inferred beyond what participants wrote in their questionnaire. Finson (2003) discusses how although some studies have allowed participants to use colours, it still has not resulted in all minority participants drawing minorities, therefore it was probably most efficient to perform analysis the way we did on the questionnaire responses.

The participants of the study were extremely diverse. There was no control for factors like country of residence, first language, or field of occupation. As a result, there were many different kinds of people, but not enough of any consistent profile to do more targeted analysis, as we ran into in Section 3.8. One of the benefits of Mechanical Turk is having access to such a large and diverse group of participants, but one of the trade-offs is the lack of control needed for more fine-grained analysis. Now that these results have been collected, however, it will be possible to design more specific experiments.

The primary analysis of the study data was the statistical analysis of participant compared to drawing gender, age, and ethnicity. These questions had limited options for a response as they were multiple choice, a number, and chosen from a list. The participant first language, whether the drawing "looks like you or someone else", use of terms in the responses, and prevalence of artifacts were all questions that had more freedom for participants to respond as they wished. A second rater was not used for this analysis and I relied on counts as much as possible, as opposed to coding. However for more thorough analysis, establishing inter-rater reliability would be beneficial if the goal was to move beyond identifying themes and into drawing conclusions. Related to this limitation, it was not possible to discern from the data whether the drawings provided had positive or negative connotations among the participants. In some cases, such as the drug users, they were clearly negative. In Section 2.1.5 I discussed how feelings of autonomy and engagement differ from masculine and feminine perspectives – a solitary user may be acceptable for those who value independence, but for others this may signify loneliness since a lack of interdependence means being cut off from others. We do not know whether participants were depicting their interpretation of HCI's "people" words in a positive or negative light.

There are a lack of prescriptive outcomes from this research: I cannot suggest that one of the HCI "people" words can or should be used over another. The goal, however, has been achieved, since we now know the imagery the general public associates with these words, that there are themes emerging, and that the community needs to exercise caution since neutral or universal terms risk signifying exclusively masculine subjects, therefore any limitations did not impede the success of this research.

5.9 Chapter Summary

This chapter discussed the findings presented in this thesis in the larger context of feminist theory and implications for language use in HCI. Additionally, the limitations of the method and analysis employed in this study were discussed. Based on deeper analysis of the study results, it was shown that the gender-neutral terms do evoke specific connotations to different degrees across the conditions. Implications of the study can be summarized in the following three categories:

Context of Use:

- Further context around use of the "people" words, avoiding blanket universality, can assist with the unpredictable ambiguity in how they are interpreted across different gender, ethnic, age, linguistic, and other groups.

Awareness of Gender Schemas:

- The HCI "people" words carry connotations to networks of words that work together to form gender schemas over time. Being aware of these and critically, deliberately challenging these through exploring different situational aspects of the person's identity through the use of tools such as the deliberately inclusive GenderMag persona kit (M. Burnett, Stumpf, Beckwith, & Peters, 2015) and by increasing visibility of non-normative role models can assist with falling back into a gendered, normative default that fills a space meant to be inclusive.

Reading Gender Identity:

- A rationalist reading of the benefits and pitfalls of being one of HCI's "people" words (for example, a user) offers a different perception on autonomy, relatedness and competence than a feminist reading would. As long as HCI continues to read interactions through this frame, gendered schemas will be embedded under this interpretation. More rigorous adoption of a Feminist HCI approach will allow for researchers and designers to challenge or shift

assumptions about what a “user” is, rather than relying on standardized universal terms that also standardize how we form assumptions about our audience.

These implications intend to motivate the use of more inclusive language and practices in HCI that avoid neutral references privileging the status quo which exclude or impede others from the growing community and those within the far-ranging reach of its influence.

Chapter 6
Conclusions

Feminist HCI emphasizes a strong commitment to people's everyday experiences, which play a large part in the construction of gender and identity. In order to emphasize the experiences of the marginalized as one of the key components of a feminist epistemology, we must be able to conceptualize the subject in all its various shapes and sizes. Existing research shows that there is not a good understanding of how we perceive the supposedly gender-neutral terms from HCI when they are used. The research presented in this thesis was motivated by this lack of understanding and the implications for representing those outside the norm when the universal language was presumed to be masculine.

Within this thesis, the five most frequently used words to describe people in CHI papers were evaluated using a Mechanical Turk study to determine how they are perceived by the general public. Analysis of the study results revealed a tendency towards perceiving the subject as male. These findings are summarized below. As a little explored phenomenon, perception of gender in HCI as well as the influence of human-computer interactions on construction of gender and identity warrant further work beyond this initial study. Recommendations for future research directions are also discussed.

6.1 Research Objectives

The objectives of this research were to determine a theoretical framework suitable for evaluating the implications of gendered language in HCI; design a research study to investigate how the most common words that have been used to refer to people in recent HCI literature are received by non-experts; and evaluate the results on the perception of gender in HCI language. The first objective was addressed by reviewing literature related to feminist issues in HCI, how linguistic analysis can be used in HCI research, and feminist theory as a theoretical frame (Chapter 2). This review identified a gap in the literature pertaining to the existing methods used to approach feminist issues in HCI. It also provides a theoretical basis to read the results and implications of the study.

The second research objective was addressed by designing the research study (Chapter 3). The design of the study brought together the psychology-based methodology familiar within HCI, and the rhetorical theory necessary to motivate a focus on language and appropriately interpret the results.

Finally, the third objective was addressed through analysis and discussion of the study results (Chapter 4 and Chapter 5). Study findings led to the following three contributions:

- **Determining the theoretical framework for evaluating the implications of gendered language in HCI.** The literature review in Chapter 2 draws attention to the areas in which the HCI community has had difficulty encouraging equality or diversity in the past, and puts into context why the reception of "universal" terms as masculine is not an acceptable standard for the community.
- **Designing a research study to investigate how the most common words that have been used to refer to people in recent HCI literature are received by non-experts.** Based on a review of past work, an interdisciplinary approach was used to collect imagery elicited by HCI's "people words" along with demographic questionnaire data. We developed a new study method to engage with the broader community in order to collect and evaluate language used in HCI research papers and we found that this was an effective way to conduct research studies of this type.
- **Evaluating the results on the perception of gender in HCI language.** We found that the five HCI "people words" (user, participant, person, designer, and researcher) were predominantly perceived as male by our participants, with the exception of females being just as likely to draw "participant" and "person" as male or female. We offer a better understanding of how these words commonly used in HCI literature are received in contrast to the inclusive or neutral intentions of the authors using them.

6.2 Future Work

The results and conclusions presented in this thesis draw attention to a number of areas that warrant further study.

First, findings from this study suggest that the context in which a word is used can influence its perception: the context within a sentence (i.e. a "desktop user") or the context of the person listening (who they are). Dray et al. (2014) suggest making gender *sensitive* rather than gender *neutral* things. In order to be responsible and be sensitive, we need to understand what we should be sensitive to: future research should look at the commonly occurring concepts constituting gender schemas around the "people" words in a systematic way to become more aware. To investigate

effects of context for the people perceiving these words it will be necessary for future research to do more targeted and controlled follow-ups on the broader patterns found in this study.

Second, the HCI community needs to more firmly establish better ways to think about, talk about and investigate gender. Stepulevage (2001) recommends moving away from the type of discussions about "gender" and "technology" or "computing" that highlight differences, instead using "relation" to highlight social practices constituting difference, and identifying the practices or processes needing change within those. We need to move away from snapshot understandings of gender in HCI to understand what or how social relations enhance or diminish these differences, so that we can build and act on what we learned from this study.

Finally, an important part about remaining progressive as a community is regular exercise of critical engagement with our own assumptions, theories and practices. A trend map by Padilla et al. (2014) showing topics of discussion at CHI over the past five years demonstrates that discussions around women were most often linked to public or community life, peaking about three years ago before declining. Knight & Cunningham (2004) reinforce the fact that it's difficult to change our preconceived notions about roles unless we address them directly through discussion: this study brings those notions to light, and now it's time to engage with them.

Appendix A

Study Questionnaire & Instructions

Page 1: Ethics Consent Form

Participants were directed from the Mechanical Turk HIT to our web page when they decided to participate in the study. When they arrived at the website, they were presented with a page titled **Information Letter and Consent Form**, which read as follows:

> Dr. Mark Hancock, Dr. Sheelagh Carpendale, Dr. Neil Randall, Adam Bradley, Cayley MacArthur
>
> You are invited to participate in a research study conducted by Cayley MacArthur and Adam Bradley, under the supervision of Dr. Mark Hancock, Management Sciences Department of the University of Waterloo, Canada. The objectives of the research study are to better understand language and HCI. The study is part of a Master's thesis and ongoing graduate work.
>
> If you decide to volunteer, you will be asked to draw a picture and then fill out a short online survey that is completed anonymously. Survey questions focus on the picture that you have drawn and some demographic information. Participation in this study is voluntary. You may decline to answer any questions that you do not wish to answer and you can withdraw your participation at any time by not submitting your responses. There are no known or anticipated risks from participating in this study.
>
> This survey uses a UWaterloo system saved on Canadian servers. It is important for you to know that any information that you provide will be confidential. All of the data will be summarized and no individual could be identified from these summarized results. Furthermore, the web site is programmed to collect responses alone and will not collect any information that could potentially identify you (such as machine identifiers).
>
> You may decline to answer any questions that you do not wish to answer and you can withdraw your participation at any time by ceasing to answer questions, without penalty or loss of remuneration. To receive remuneration please proceed to the end of the questionnaire, obtain the unique code for this HIT, and submit it.
>
> This HIT is expected to take on average 20 mins and you will be paid $0.50.

> The data, with no personal identifiers, collected from this study will be maintained on a password-protected computer database in a restricted access area of the university.
>
> Should you have any questions about the study, please contact either Cayley MacArthur (csimacar@uwaterloo.ca) or Mark Hancock (mark.hancock@uwaterloo.ca). Further, if you would like to receive a copy of the results of this study, please contact either investigator.
>
> I would like to assure you that this study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Committee. However, the final decision about participation is yours. If you have any comments or concerns resulting from your participation in this study, please feel free to contact Dr. Maureen Nummelin in the Office of Research Ethics at 1-519-888-4567, Ext. 36005 or maureen.nummelin@uwaterloo.ca.
>
> Thank you for considering participation in this study.
>
> **Consent to Participate:**
>
> With full knowledge of all foregoing, I agree, of my own free will, to participate in this study.

Participants would then select "I agree to participate" or "I do not wish to participate (please close your browser now)".

Page 2: Visualization Task

Participants were given the following instruction:

> For the next 10 seconds think about a [*condition*].

After 10 seconds had elapsed, a button labelled "Next page" would become available to proceed to the next page.

Page 3: Drawing Task

On this page, participants were instructed to draw their condition. The page read as follows:

> In the box below, draw a [*condition*] sitting down:
> If you have a touch or pen device, please consider using it to draw.
> To draw on another device, you can mail yourself a link to continue: [*mailto link to send the HIT URL to yourself*]
>
> If that link does not work, you can copy and paste this link into an email to yourself and use it to continue where you left off: [*link*]

Below this text was the drawing interface. Written inside the drawing box were the instructions:

> In this box, draw a [*condition*] sitting down

Below the box were two buttons to clear the canvas or upload the drawing, with the instruction below:

> Once you have uploaded a screenshot of your drawing, please continue to the next page:

Followed by a button to go to the "Next page" which would only be active once a screenshot of the drawing had been uploaded.

Page 4: Drawing Confirmation

This page displayed a copy of the drawing for the participants' reference, and said:

> If you are not satisfied with your drawing, click the 'Back' button to make a new one (Note: your old drawing will be erased)

This was accompanied by a "Back" button. Underneath, it asked:

> Did you draw a human? [*Yes/No*]

Page 5a: Questionnaire about Human Drawings

This page also provided a copy of the drawing for reference and asked the following questions:

1. Does the subject of your drawing have an age, and if so, what is it? [*Text box*]
2. Does the person in your drawing look like you, or someone else? [*Text box*]
3. Does the subject of your drawing have a gender, and if so, what is it? [*Male/Female/Other (you may elaborate below)*]

3a. (If you answered 'other', you may elaborate here) [*Text box*]

4. What is the race/ethnicity of the person in your drawing? (i.e. peoples' ethnicity describes their feeling of belonging and attachment to a distinct group of a larger population that shares their ancestry, colour, language or religion) [*Caucasian, Latino/Hispanic, Middle Eastern, African, Caribbean, South Asian, East Asian, Mixed, Other (please clarify below)*]

4a. (If you answered 'other' above, please clarify here) [*Text box*]

5. Were there any objects in your drawing? If so, what were they? [*Text box*]

6. What factors did you think about before drawing, if any? (Example: age, race, gender, objects, setting, sexual orientation, outfit, none…) [*Text box*]
7. Please use this space to include or clarify any additional details about what you drew. [*Text box*]

Page 5b: Questionnaire about Non-Human Drawings

A copy of the drawing was provided for reference here as well, proceeded by the following questions:

1. Since your drawing is not a human, what did you draw? [*Text box*]
2. Does the subject of your drawing have an age, and if so, what is it? [*Text box*]
3. Were there any objects in your drawing? If so, what were they? [*Text box*]
4. What factors did you think about before drawing, if any? (Example: objects, setting, context, location, age, gender, outfits, attributes, none…) [*Text box*]
5. Please use this space to include or clarify any additional details about what you drew. [*Text box*]

Page 6: Questionnaire about the Participants

This page contained questions about the participants as opposed to their drawing. As such, a copy of the drawing was not needed for reference. Questions about the participants were optional to comply with ethical standards regarding collection of personal information.

1. What is your age? (Optional) [*Text box*]
2. What is your gender? (Optional) [*Male/Female/Other (you may elaborate below)*]

2a. (If you answered 'other', you may elaborate here) [*Text box*]

3. In what country were you born? (Optional) [*Text box*]
4. In which country do you currently live? (Optional) [*Text box*]
5. Please indicate your ethnicity (i.e. peoples' ethnicity describes their feeling of belonging and attachment to a distinct group of a larger population that shares their ancestry, colour, language or religion) (Optional) [*Caucasian, Latino/Hispanic, Middle Eastern, African, Caribbean, South Asian, East Asian, Mixed, Other (please clarify below)*]

5a. (If you answered 'other' above, please clarify here) [*Text box*]

6. Do you speak a language (or languages) other than English at home? (Optional) [*Yes (please answer 6a)/No*]

6a. If you answered 'yes' to question 6, please list the language(s) primarily spoken (i.e. French, Spanish, etc.) [*Text box*]

7. What is your first or primary language? (Optional) [*Text box*]

7a. If English is <u>not</u> your first language, how well would you rate your ability to communicate in English? [*No knowledge at all, Limited knowledge, Functional knowledge, Fluent, Not applicable (English is my first language)*]

8. Sometimes people who work or study in a certain industry will think about concepts in a similar way. In what field/industry do you work/study? (Optional) [*Agriculture, Accounting, Advertising, Aerospace, Apparel & Accessories, Arts, Automotive, Banking, Broadcasting, Brokerage, Biotechnology, Customer Service/Call Centres, Cargo Handling, Chemical, Computer: Hardware, Computer: Software, Consulting, Consumer Products, Cosmetics & Hygiene, Defense, Department Stores, E-Commerce, Education, Electronics, Energy, Entertainment & Leisure, Financial Services, Food & Beverage, Grocery, Healthcare, Humanities, Internet, Internet Publishing, Journalism, Legal, Literature, Manufacturing, Motion Picture & Video, Music, Newspaper/Magazine, Pharmaceuticals, Publishing, Real Estate, Retail & Wholesale, Service, Sports, Technology, Telecommunications, Television, Tobacco, Transportation, OTHER*]

8a. If you answered "other" above, or if you would like to clarify your selection, please use the space below. [*Text box*]

9. Please take this opportunity to clarify or elaborate on any of your answers here. [*Text box*]

Participants would then click a button labelled "Submit Data".

Page 7: Confirmation of Study Completion

This page contained the following text:

> **Thank You**
>
> Your Amazon Mechanical Turk Code is: QRxYg23

> Please copy the above code and paste it into the available area in the HIT.
>
> **Task Completed**
>
> Thank you for participating in this ongoing study. If you are interested in reviewing the results of this research, please email me (adam.bradley@uwaterloo.ca), and I will be sure to contact you when the study has been completed.
>
> The purpose of this study is to learn how people gender certain words.
>
> Once again, I remind you that the data collected as part of this study, with the exception of your Mechanical Turk User ID (MTurk ID), will be made publicly available.
>
> If you have concerns about this study, you may also contact the Director of the Office of Research Ethics at the University of Waterloo at (519) 888-4567 ext. 36005. This project was reviewed by, and received clearance through a University of Waterloo Research Ethics Committee.
>
> Sincerely,
>
> Cayley MacArthur and Adam Bradley
>
> If you have any comments regarding this HIT, or found the instructions hard to understand, we would love to hear from you. Please enter such details in the box below. We will use your feedback to improve this, and other HITs, in the future.
>
> [*Text box*]

The final button read “Click here to finish”. This submitted any final comments and took the participant to the next page which only contained a final reminder.

Page 8: Termination

After any final comments were completed, this page contained a simple reminder:

> Thank you for participating in our study. In order to be compensated, please remember to copy & paste the following code into the HIT on Mechanical Turk: QRxYg23 You may now close this window.

Appendix B
Ethics Approval

Dear Researcher:

This is to advise that the ethics review of your application to conduct research:

Title: The Gendered User
ORE #: 19634
Collaborator: Sheelagh Carpendale (sheelagh@ucalgary.ca)
Faculty Supervisor: Mark Hancock (mark.hancock@uwaterloo.ca)
Student Investigator: Adam Bradley (adam.bradley@uwaterloo.ca)
Student Investigator: Cayley MacArthur (csimacar@uwaterloo.ca)

has been completed through a University of Waterloo Research Ethics Committee. Based on the outcome of the ethics review process, I am pleased to advise you that your project has received ethics clearance.

Note 1: This ethics clearance from a University of Waterloo Research Ethics Committee is valid for one year from the date shown on the certificate and is renewable annually. Renewal is through completion and ethics clearance of the Annual Progress Report for Continuing Research (ORE Form 105).

Note 2: This project must be conducted according to the application description and revised materials for which ethics clearance has been granted. All subsequent modifications to the project also must receive prior ethics clearance (i.e., Request for Ethics Clearance of a Modification, ORE Form 104) through the Office of Research Ethics and must not begin until notification has been received by the investigators.

Note 3: Researchers must submit a Progress Report on Continuing Human Research Projects (ORE Form 105) annually for all ongoing research projects or on the completion of the project. The Office of Research Ethics sends the ORE Form 105 for a project to the Principal Investigator or

Faculty Supervisor for completion. If ethics clearance of an ongoing project is not renewed and consequently expires, the Office of Research Ethics may be obliged to notify Research Finance for their action in accordance with university and funding agency regulations.

Note 4: Any unanticipated event involving a participant that adversely affected the participant(s) must be reported immediately (i.e., within 1 business day of becoming aware of the event) to the ORE using ORE Form 106. Any unanticipated or unintentional change which may impact the research protocol, information-consent document or other study materials, must be reported to the ORE within 7 days of the deviation usng ORE Form 107.

Best wishes for success with this study.

Bibliography

About. (n.d.). *Art+Feminism*. Retrieved October 14, 2015, from http://art.plusfeminism.org/

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, *37*(4), 445–456.

Arrabito, G. R. (2009). Effects of Talker Sex and Voice Style of Verbal Cockpit Warnings on Performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *51*(1), 3–20.

Astur, R. S., Ortiz, M. L., & Sutherland, R. J. (1998). A characterization of performance by men and women in a virtual Morris water task: a large and reliable sex difference. *Behavioural Brain Research*, *93*(1-2), 185–190.

Balarajan, M., Gray, M., & Mitchell, M. (2011). *Monitoring equality: developing a gender identity question*.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215.

Bandura, A. (2006). Toward a Psychology of Human Agency. *Perspectives on Psychological Science*, *1*(2), 164–180.

Bardzell, S. (2010). Feminist HCI: taking stock and outlining an agenda for design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1301–1310). ACM. Retrieved March 3, 2014, from http://dl.acm.org/citation.cfm?id=1753521

Bardzell, S., & Bardzell, J. (2011). Towards a feminist HCI methodology: social science, feminism, and HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 675–684). ACM. Retrieved September 17, 2014, from http://dl.acm.org/citation.cfm?id=1979041

Bardzell, S., Churchill, E., Bardzell, J., Forlizzi, J., Grinter, R., & Tatar, D. (2011). Feminism and interaction design. *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1–4). ACM. Retrieved September 17, 2014, from http://dl.acm.org/citation.cfm?id=1979587

Bardzell, S., Gross, S., Wain, J., Toombs, A., & Bardzell, J. (2011). The significant screwdriver: care, domestic masculinity, and interaction design. *Proceedings of the 25th BCS Conference on Human-Computer Interaction* (pp. 371–377). British Computer Society. Retrieved February 25, 2014, from http://dl.acm.org/citation.cfm?id=2305379

Beckwith, L., Kissinger, C., Burnett, M., Wiedenbeck, S., Lawrance, J., Blackwell, A., & Cook, C. (2006). Tinkering and gender in end-user programmers' debugging. *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 231–240). ACM. Retrieved July 12, 2015, from http://dl.acm.org/citation.cfm?id=1124808

Bodzin, A., & Gehringer, M. (2001). Breaking Science Stereotypes. *Science and Children*, *38*(4), 36–41.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, Syntax and Semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind ; advances in the study of language and thought* (pp. 61–79). Cambridge, Mass: MIT Press.

Boynton, P. M. (1998). People should participate in, not be subjects of, research. *BMJ (Clinical research ed.)*, *317*(7171), 1521.

Bradley, A., MacArthur, C., Hancock, M., & Carpendale, S. (2015). Gendered or neutral?: considering the language of HCI. *Proceedings of the 41st Graphics Interface Conference* (pp. 163–170). Canadian Information Processing Society. Retrieved October 18, 2015, from http://dl.acm.org/citation.cfm?id=2788919

Brafman, A. H. (2012). *The Language of Drawings: A New Finding in Psychodynamic Work*. London: Karnac Books. Retrieved March 3, 2015, from http://public.eblib.com/choice/publicfullrecord.aspx?p=878529

Burgstahler, S., & Ladner, R. (2006). *Increasing the participation of people with disabilities in computing fields: From research to practice.* AccessComputing: The Alliance for Access to Computing Careers. Seattle: University of Washington. Retrieved August 4, 2015, from http://www.washington.edu/accesscomputing/get-informed/accesscomputing-research-practice

Burnett, M. M., Beckwith, L., Wiedenbeck, S., Fleming, S. D., Cao, J., Park, T. H., Grigoreanu, V., et al. (2011). Gender pluralism in problem-solving software. *Interacting with Computers*, *23*(5), 450–460.

Burnett, M., Stumpf, S., Beckwith, L., & Peters, A. (2015, November). The GenderMag Kit: To Find Usability Issues Through a Gender Lens. Retrieved December 22, 2015, from http://eusesconsortium.org/gender/genderMag-flyer.pdf

Butler, J. (1988). Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre journal*, 519–531.

Butler, J. (1990). *Gender trouble: feminism and the subversion of identity*. Thinking gender. New York: Routledge.

Butler, J. (2004). *Undoing gender*. New York ; London: Routledge.

Carnegie, T. (1998). The rhetoric of experience, explorations in experience as a key term in feminist discourse. Retrieved September 13, 2015, from https://uwspace.uwaterloo.ca/handle/10012/366

Cassell, J. (2002). Genderizing HCI. *The Handbook of Human–Computer Interaction. Mahwah, NJ: Erlbaum*, 402–411.

Cassell, J., & Jenkins, H. (1998). *From Barbie to Mortal Kombat gender and computer games*. Cambridge, Mass.: MIT Press. Retrieved July 22, 2015, from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=393

Chalmers, I. (1999). People are "participants" in research. Further suggestions for other terms to describe "participants" are needed. *BMJ (Clinical research ed.)*, *318*(7191), 1141.

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw-a-scientist test. *Science Education*, *67*(2), 255–265.

Cockburn, C., & Ormrod, S. (1993). *Gender and technology in the making*. London ; Thousand Oaks, Calif: Sage.

Cohn, C. (1987). Sex and death in the rational world of defense intellectuals. *Signs*, 687–718.

Coleman, B. (2011). *Hello avatar: rise of the networked generation*. Cambridge, Mass: MIT Press.

Collecting Transgender-Inclusive Gender Data in Workplace and Other Surveys. (2013). *Human Rights Campaign: Working for Lesbian, Gay, Bisexual and Transgender Equal Rights*. Retrieved June 27, 2015, from http://www.hrc.org/resources/entry/collecting-transgender-inclusive-gender-data-in-workplace-and-other-surveys

Conron, K. J., Scout, null, & Austin, S. B. (2008). "Everyone has a right to, like, check their box:" findings on a measure of gender identity from a cognitive testing study with adolescents. *Journal of LGBT health research*, *4*(1), 1–9.

Corbett, G. G. (1991). *Gender*. Cambridge textbooks in linguistics. Cambridge [England] ; New York: Cambridge University Press.

Czerwinski, M., Tan, D. S., & Robertson, G. G. (2002). Women take a wider view. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 195–202). ACM. Retrieved September 21, 2014, from http://dl.acm.org/citation.cfm?id=503412

De Beauvoir, S. (1948). *The ethics of ambiguity*. Citadel Press.

Dray, S. M., Holtzblatt, K., Murray, D., Busse, D. K., Brock, A. M., Peters, A. N., Bardzell, S., et al. (2014). Perspectives on gender and product design (pp. 53–56). ACM Press. Retrieved July 22, 2015, from http://dl.acm.org/citation.cfm?doid=2559206.2559218

Dray, S. M., Peer, A., Brock, A. M., Peters, A., Bardzell, S., Burnett, M., Churchill, E., et al. (2013). Exploring the Representation of Women Perspectives in Technologies. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13 (pp. 2447–2454). New York, NY, USA: ACM. Retrieved February 24, 2014, from http://doi.acm.org/10.1145/2468356.2468799

Dray, S. M., Peters, A. N., Brock, A. M., Peer, A., Druin, A., Gitau, S., Kumar, J., et al. (2013). Leveraging the Progress of Women in the HCI Field to Address the Diversity Chasm. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13 (pp. 2399–2406). New York, NY, USA: ACM. Retrieved February 24, 2014, from http://doi.acm.org/10.1145/2468356.2468785

Edworthy, J., Hellier, E., & Rivers, J. (2003). The use of male or female voices in warnings systems: a question of acoustics. *Noise & Health*, *6*(21), 39–50.

Enders, J. (2013, July 14). Sex and Gender. *Formulate Information Design*. Retrieved November 28, 2015, from http://www.formulate.com.au/blog/sex-and-gender

Finson, K. D. (2003). Applicability of the DAST-C to the images of scientists drawn by students of different racial groups. *Journal of Elementary Science Education*, *15*(1), 15–26.

Fort, D., & Varney, H. (1989). How students see scientists: Mostly male, mostly white, and mostly benevolent. *Science and Children*, *26*(8), 8–13.

Gaines, E. (2006). Communication and the Semiotics of Space. *Journal of Creative Communications*, *1*(2), 173–181.

Gasson, S. (2003). Human-centered vs. user-centered approaches to information system design. *Journal of Information Technology Theory and Application (JITTA)*, *5*(2), 5.

Gelber, S. (2000). Do-it-yourself: constructing, repairing, and maintaining domestic masculinity. In J. Scanlon (Ed.), *The gender and consumer culture reader*. New York, NY: New York University Press. Retrieved from http://trove.nla.gov.au/work/35226558?selectedversion=NBD21462199

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, *118*(3), 1049–1074.

Griggs, B. (2011, October 21). Why computer voices are mostly female. *CNN*. Retrieved from http://www.cnn.com/2011/10/21/tech/innovation/female-computer-voices/

Gross, T. (2014). Human-Computer Interaction Education and Diversity. *Human-Computer Interaction. Theories, Methods, and Tools* (pp. 187–198). Springer. Retrieved November 15, 2015, from http://link.springer.com/chapter/10.1007/978-3-319-07233-3_18

Heikes, D. K. (2010). *Rationality and feminist philosophy*. Continuum studies in philosophy. London ; New York: Continuum.

Hemphill, L., Erickson, I., Ribes, D., & Mergel, I. (2014). Feminism and social media research (pp. 319–322). ACM Press. Retrieved September 17, 2014, from http://dl.acm.org/citation.cfm?doid=2556420.2558864

Hemphill, L., & Otterbacher, J. (2012). Learning the lingo?: gender, prestige and linguistic adaptation in review communities. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 305–314). ACM. Retrieved July 11, 2015, from http://dl.acm.org/citation.cfm?id=2145254

Herbenick, D., & Baldwin, A. (2014, February 15). What Each of Facebook's 51 New Gender Options Means. *The Daily Beast*. Retrieved November 28, 2015, from

http://www.thedailybeast.com/articles/2014/02/15/the-complete-glossary-of-facebook-s-51-gender-options.html

Hesse-Biber, S. N. (Ed.). (2012). *Handbook of feminist research: theory and praxis* (2nd ed.). Thousand Oaks: SAGE.

Hill, O. W., Pettus, W. C., & Hedin, B. A. (1990). Three studies of factors affecting the attitudes of blacks and females toward the pursuit of science and science-related careers. *Journal of Research in Science Teaching*, *27*(4), 289–314.

Hiltz, S. R., Johnson, K., & Turoff, M. (1986). Experiments in Group Decision Making Communication Process and Outcome in Face-to-Face Versus Computerized Conferences. *Human Communication Research*, *13*(2), 225–252.

Huff, C., & Cooper, J. (1987). Sex Bias in Educational Software: The Effect of Designers' Stereotypes on the Software They Design. *Journal of Applied Social Psychology*, *17*(6), 519–532.

Huffman, A. H., Whetten, J., & Huffman, W. H. (2013). Using technology in higher education: The influence of gender roles on technology self-efficacy. *Computers in Human Behavior*, *29*(4), 1779–1786.

Irigaray, L. (1987). Translating French Feminism: Luce Irigaray's Ce Sexe qui n'en est pas un. (C. Porter, Tran.)*Translation perspectives*, *3*, 40–52.

Johnson, M. (2015, July 8). The Computer Science Pipeline and Diversity: Part 1 - How did we get here? *Google Research Blog*. Retrieved July 22, 2014, from http://googleresearch.blogspot.ca/2015/07/the-computer-science-pipeline-and.html

Kant, I. (1785). *Groundwork of the Metaphysics of Morals*. (H. Paton, Tran.). New York: Harper & Row.

Kimura, D. (1999). *Sex and cognition*. Cambridge, Mass: MIT Press.

Knight, M., & Cunningham, C. (2004). Draw an engineer test (DAET): Development of a tool to investigate students' ideas about engineers and engineering. *ASEE Annual Conference and Exposition* (Vol. 2004). Retrieved November 15, 2015, from http://engineering.nyu.edu/gk12/amps-cbri/pdf/Draw%20an%20Engineer%20Test%20(DAET)%20-%20Development%20of%20a%20Tool%20to%20Investigate%20Students%E2%80%99%20Ideas%20about%20Engineers%20and%20Engineering.pdf

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of psycholinguistic research*, *22*(5), 519–534.

Kotamraju, N. P. (2011). Playing stupid, caring for users, and putting on a good show: Feminist acts in usability study work. *Interacting with Computers*, *23*(5), 439–446.

Lam, S. T. K., Uduwage, A., Dong, Z., Sen, S., Musicant, D. R., Terveen, L., & Riedl, J. (2011). WP: clubhouse?: an exploration of Wikipedia's gender imbalance. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (pp. 1–10). ACM. Retrieved July 11, 2015, from http://dl.acm.org/citation.cfm?id=2038560

Landstrom, C. (2007). Queering feminist technology studies. *Feminist Theory*, *8*(1), 7–26.

Lim, S., & Kwon, N. (2010). Gender differences in information behavior concerning Wikipedia, an unorthodox information source? *Library & Information Science Research*, *32*(3), 212–220.

Lugones, M. C., Spelman, E. V., Lugones, M. C., & Spelman, E. V. (1983). Have we got a theory for you! Feminist theory, cultural imperialism and the demand for "the woman"s voice'. *Women's Studies International Forum*, *6*(6), 573–581.

Magnan, S. S. (1983). Age and Sensitivity to Gender in French. *Studies in Second Language Acquisition*, *5*(02), 194.

Mandiberg, M. (2015, February 1). The Affective Labor of Wikipedia: GamerGate, Harassment, and Peer Production. *SocialText*. Journal, . Retrieved July 11, 2015, from http://socialtextjournal.org/affective-labor-of-wikipedia-gamergate/

Margolis, J., & Fisher, A. (2002). *Unlocking the clubhouse: women in computing*. Cambridge, Mass.: MIT Press.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1–23.

McMaster, S. (2012). New Approaches to Image-based Research and Visual Literacy: A Pilot Project Report. *New Horizons in Visual Literacy: Selected Readings of the International Visual Literacy Association*, 122–132.

Meyer, E. (1997). The Expanded Field of Landscape Architecture. *Ecological Design and Planning* (pp. 45–79). New York: John Wiley & Sons. Retrieved July 8, 2015, from http://www.academia.edu/9082239/Expanded_Field_of_Landscape_Architecture_1997_

Meyers, D. T. (2002). *Gender in the Mirror*. Oxford University Press. Retrieved July 12, 2015, from http://www.oxfordscholarship.com/view/10.1093/0195140419.001.0001/acprof-9780195140415

Miele, E. (2014). Using the Draw-a-scientist Test for Inquiry and Evaluation. *Journal of College Science Teaching*, *43*(4), 36–40.

Monhardt, R. M. (2003). The image of the scientist through the eyes of Navajo children. *Journal of American Indian Education*, *42*(3), 25–39.

Morning, A. (2004). ETHNIC CLASSIFICATION IN INTERNATIONAL CONTEXT: A CROSS-NATIONAL COMPARISON OF 1995-2004 CENSUS ITEMS. *New York University, New York*. Retrieved November 29, 2015, from http://www.researchgate.net/profile/Ann_Morning/publication/228894288_ETHNIC_CLASS

IFICATION_IN_INTERNATIONAL_CONTEXT_A_CROSS-NATIONAL_COMPARISON_OF_1995-2004_CENSUS_ITEMS/links/0deec5307bd80bff23000000.pdf

Morrison, A. (2013). Facebook and Coaxed Affordances. *Identity technologies: constructing the self online*, Wisconsin studies in autobiography (pp. 112–131). Madison, Wisconsin: The University of Wisconsin Press.

Muller, M. (2011). Feminism asks the "Who" questions in HCI. *Interacting with Computers*, *23*(5), 447–449.

Nass, C., Moon, Y., & Green, N. (1997). Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers With Voices. *Journal of Applied Social Psychology*, *27*(10), 864–876.

Nicholson, J., Gelpi, A., Sulzby, E., & Young, S. (1998). Influences of Gender and Open-Ended Software on First Graders' Collaborative Composing Activities on Computers. *Journal of Computing in Childhood Education*, *9*(1), 3–42.

Padilla, S., Methven, T. S., Corne, D. W., & Chantler, M. J. (2014). Hot topics in CHI: trend maps for visualising research (pp. 815–824). ACM Press. Retrieved September 17, 2014, from http://dl.acm.org/citation.cfm?doid=2559206.2578867

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision making*, *5*(5), 411–419.

Posard, M. N. (2014). Status processes in human-computer interactions: Does gender matter? *Computers in Human Behavior*, *37*, 189–195.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology*, *14*(2), 154–166.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Sandstrom, N. J., Kaufman, J., & A. Huettel, S. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research*, *6*(4), 351–360.

Savicki, V., Foster, D., & Kelley, M. (2006). Gender, Group Composition and Task Type in Virtual Groups. In M. Barrett & M. Davidson (Eds.), *Gender and Communication at Work* (pp. 270–282). Ashgate Publishing Limited. Retrieved from http://www.wou.edu/~fosterd/papers/cmc.pdf

Shin, H. B., & Bruno, R. (2003). Language Use and English-Speaking Ability: 2000. Census 2000 Brief. Retrieved November 29, 2015, from http://eric.ed.gov/?id=ED482418

Sontag, S. (2001). *On photography*. New York: Picador USA : Farrar, Straus and Giroux.

Spence, J., Frohlich, D. M., & Andrews, S. (2013). Performative experience design. *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 2049–2058). ACM. Retrieved March 21, 2014, from http://dl.acm.org/citation.cfm?id=2468723

Sprague, J., & Hayes, J. (2000). Self-determination and empowerment: A feminist standpoint analysis of talk about disability. *American journal of community psychology*, *28*(5), 671–695.

Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., & Kuchibhotla, S. (2007). Assessing Media Influences on Middle School Aged Children's Perceptions of Women in Science Using the Draw-A-Scientist Test (DAST). *Science Communication*, *29*(1), 35–64.

Stepulevage, L. (2001). Gender/Technology Relations: Complicating the gender binary. *Gender and Education*, *13*(3), 325–338.

Stoljar, N. (2014). Feminist Perspectives on Autonomy. (E. Zalta, Ed.)*The Stanford Encyclopedia of Philosophy*. Retrieved July 11, 2015, from http://plato.stanford.edu/archives/spr2014/entries/feminism-autonomy

Suchman, L. (1987). *Plans and situated actions: the problem of human-machine communication*. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Suchman, L. (2007). Agencies in technology design: Feminist reconfigurations. *Unpublished manuscript*. Retrieved August 4, 2015, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.468.9781&rep=rep1&type=pdf

Sumrall, W. J. (1995). Reasons for the Perceived Images of Scientists by Race and Gender of Students in Grades 1-7. *School Science and Mathematics*, *95*(2), 83–90.

Symington, D., & Spurling, H. (1990). The "Draw a Scientist Test": interpreting the data. *Research in Science & Technological Education*, *8*(1), 75–77.

Takayama, L., & Nass, C. (2008). Driver safety and information from afar: An experimental driving simulator study of wireless vs. in-car information services. *International Journal of Human-Computer Studies*, *66*(3), 173–184.

Tan, D. S., Czerwinski, M., & Robertson, G. (2003). Women go with the (optical) flow (p. 209). ACM Press. Retrieved July 9, 2015, from http://portal.acm.org/citation.cfm?doid=642611.642649

Tischler, L. (2009, September 21). Introducing the Femme Den: Going Beyond "Shrink it and Pink it." *Fast Company*. Retrieved July 22, 2015, from http://www.fastcompany.com/1361553/introducing-femme-den-going-beyond-shrink-it-and-pink-it

Tupper, S. (2014, February 16). Kim Crow, The Original Bitching Betty - Audio Episode Show Notes. *Airspeed*. Aviation and Aerospace blog, . Retrieved October 18, 2015, from

http://airspeedonline.com/2014/02/kim-crow-the-original-bitching-betty-audio-episode-show-notes/

Türkcan, B. (2013). Semiotic approach to the analysis of children's drawings. *Educational Sciences: Theory and Practice*, *13*(1), 600–607.

Turkle, S., & Papert, S. (1991). Epistemological Pluralism. *Constructionism: research reports and essays, 1985-1990*, 161.

user, n.1. (2014). *OED Online*. Oxford University Press.

Van House, N. (2011). Feminist HCI meets facebook: Performativity and social networking sites. *Interacting with Computers*, *23*(5), 422–429.

Vasilescu, B., Capiluppi, A., & Serebrenik, A. (2014). Gender, Representation and Online Participation: A Quantitative Study. *Interacting with Computers*, *26*(5), 488–511.

Wajcman, J. (1991). *Feminism confronts technology*. University Park, Pa: Pennsylvania State University Press.

Wallet, G., Sauzeon, H., Rodrigues, J., & N'Kaoua, B. (n.d.). Transfer of spatial knowledge from a virtual environment to reality: Impact of route complexity and subject's strategy on the exploration mode. *JVRB - Journal of Virtual Reality and Broadcasting, 6(2009), no. 4.*, *6*(4). Retrieved from https://www.jvrb.org/past-issues/6.2009/1757

Ward, A. (2012, July 4). A bad case of jetnag? Fighter pilots to get cockpit instructions from female voice "because it relaxes them more" (but they've already nicknamed her Nagging Nora). *Daily Mail*. Retrieved October 18, 2015, from http://www.dailymail.co.uk/news/article-2168713/Fighter-pilots-nagging-Nora-female-voice-commands-cockpit-Typhoon-jets.html

Weiser, M. (1991). The computer for the 21st century. *Scientific american*, *265*(3), 94–104.

Wikipedia Editors Study: Results from the editor survey, April 2011. (n.d.). . Wikimedia Foundation. Retrieved July 11, 2015, from

https://upload.wikimedia.org/wikipedia/commons/7/76/Editor_Survey_Report_-_April_2011.pdf

Willett, C., Anderson, E., & Meyers, D. (2015, June 21). Feminist Perspectives on the Self. (E. Zalta, Ed.)*The Stanford Encyclopedia of Philosophy*. Retrieved July 8, 2015, from http://plato.stanford.edu/archives/fall2015/entries/feminism-self/

Winograd, T. (1986). *Understanding computers and cognition: a new foundation for design*. Language and being. Norwood, N.J: Ablex Pub. Corp.

Woolf, V. (1966). Three Guineas. *Reprint, San Diego: Harvest/HBJ*. Retrieved from http://gutenberg.net.au/ebooks02/0200931h.html

Wright, S. (2014). “I Mean, the Queen”s Fierce and the King’s Not’: Gendered Embodiment in Children’s Drawings. *International Journal of Early Childhood*, *46*(3), 391–406.