

Review



Literature Review: Gender Research in Human Computer Interaction

Sabrina Burtscher

November 2019



"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 741128".

Table of Contents

Introduction.....	3
Definitions: Gender.....	4
Definitions: Human Computer Interaction (HCI).....	7
Gender and Design/Research: Pitfalls.....	8
I-methodology: “Everyone is like me”.....	8
One Size Fits All: “Everyone is the same”.....	9
Gender Stereotyping: “All women are alike”.....	9
Gender difference: “Women are inherently different”.....	10
Case Studies.....	12
Content Notes for Case Studies.....	14
Case Study 1: Blackwell et al. 2016.....	15
What can we learn from this?.....	15
Case Study 2: Karuei et al. 2011.....	17
What can we learn from this?.....	17
Case Study 3: Otterbacher 2015.....	19
What can we learn from this?.....	20
Case Study 4: Clarke et al. 2013.....	21
What can we learn from this?.....	23
Case Study 5: Ahmed et al. 2014.....	25
What can we learn from this?.....	26
Case Study 6: Haimson, Brubaker, and Hayes 2014.....	28
What can we learn from this?.....	28
Case Study 7: James DiSalvo et al. 2011.....	29
What can we learn from this?.....	30
Outside Academia.....	32
What can we learn from this?.....	33
Discussion and Recommendations.....	34
Include Gender Studies and Intersectionality Basics.....	34
Remember the normative impact of language.....	34
Provide rich information on context.....	35
Make your project accessible.....	36
Make methods accessible.....	37
Consider the position of the research team.....	37
Enabling thorough discussion and reflection.....	38
References.....	39

Introduction

The present report is the outcome of a literature research conducted as a part of the Horizon2020 project “Gender Equality in Engineering through Communication and Commitment” (GEECCO, grant agreement No 741128) on gender in Human Computer Interaction (HCI) research. The purpose of this document is to

- provide readers with a basic understanding of gender, and HCI
- present an overview of how gender research in various HCI contexts has been conducted and documented,
- and provide recommendations for making HCI research gender aware, and more inclusive.

“Gender” appears to be a characteristic that researchers tend to rely on for finding significant differences in their data, although this often occurs as a pure add-on to their actual research (Hines 2004; Maccoby and Jacklin 1974; in Fine et al. 2019). The widespread use of “gender” is illustrated, for example, by the two thousand-odd entries under the keyword “gender” in the ACM digital library (March 12, 2019). Since it would be very difficult to sift through and review them all, I decided to find a meta-analysis on gender and intersectional HCI, and review examples drawn from that report. A fairly recent meta analysis with a soundly sampled corpus resulted in the 2017 paper “Intersectional HCI: Engaging Identity through Gender, Race, and Class” by Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter.

For their report, Schlesinger, Edwards, and Grinter examined 140 papers presented at CHI, the ACM Special Interest Group on Computer-Human Interaction’s conference and mapped how “the user” is constructed in those papers. They concluded that “gender research” often still means “research concerning women” or “research about women”, while issues concerning men are posited as “neutral”, applying to everyone, and genders outside the binary construction are rarely present.

The possible effects of gender roles and stereotypes on projects varies greatly, which means that there is no “one size fits all” approach to making projects gender-inclusive. This can be seen in projects spanning various fields, ranging from architecture to industrial design (Bardzell 2010). The goal of the present paper is to offer recommendations to both researchers and funding organizations on where to look for gender issues, and how to discuss and handle them.

I selected seven papers from the 140-paper corpus that Schlesinger, Edwards, and Grinter discuss, in order to cover a wide range of topics and research fields, as well as various intersections of gender(ed) issues. As a whole, the papers show different ways of studying, constructing, and considering gender in research. I discuss possible pitfalls when trying to take gender into consideration (see Breslin and Wadhwa 2017), and show in the following sections where researchers have avoided them, and how. Additionally, some examples from industry illustrate impacts that occur when gender research is missing or has gone wrong.

Initially, I looked for ways to meaningfully engage with gender as a variable in HCI research. Looking at the selected case studies included in Schlesinger, Edwards, and Grinter (2017), however, my focus broadened to also include the structure and context of the research projects. Drawing from this, I provide recommendations on how research can be done considering the concept of intersectionality, because as we will see, gender is not the one and only influence on adoption and usage of technology.

Definitions: Gender

In order to talk about how gender research is done, and how it could or should be done, it is important to discuss what “gender” means, and to introduce a few important terms.

Breslin and Wadhwa 2017 (p. 74) present the following distinction:

“Sex: A person’s biological classification as male or female, including physical appearance, chromosomes, hormones, reproductive organs, and secondary sex characteristics.

Gender: A person’s self-identification and expression in relation to what is considered socially masculine, feminine, or other gendered category, within a given context.”

Note that the terms “male” and “female” apply to the biological classification, while when talking about gender, “masculine” and “feminine” are used.

Sex differences are biological differences, for example, concerning uteri or penises, prostrates, fat distribution, or bone density. When researching sex differences, it is important to make explicit *which* differences are being researched, instead of using the blanket terms “male” and “female”. That way, readers have a better understanding of what was included in the research/design, and outliers (for example male research participants with “atypical” body fat distribution) can be better explained. *In HCI/Computer Science*, these bodily differences can impact, for example, work on fitness tracking apps, health tracking apps, voice activation, etc. The question is which bodies the hardware works for, and if the software *really* has the ability to capture what is important? Real-life examples are smartphones that are too big for small hands, standard office temperatures in smart buildings based on the more active metabolism of male office workers, and voice recognition software for interactive children’s toys that does not recognize higher pitched voices. The underlying issues in these examples are different body sizes, variations in metabolism, and differences in voice pitch. The physical differences of users and study participants are not an issue – failing to account for them, on the other hand, is a serious oversight.

In contrast, **gender differences** are based on social and cultural factors affecting the ways in which people handle things. For example, social norms code certain tasks as “women’s work”, which impacts the number of men completing them, as well as the tasks’¹ overall image and importance. Looking at gender differences means scrutinizing assumptions about things, and their image, and how assumptions and image impact the choice people do (not) have. Often, gender differences can be found even in things assumed to be neutral, for example, in a town’s snowplowing pattern. One town in Sweden

¹ The choice of the word “task” is intentional: not all tasks are paid work, and the feminization of unpaid work, such as household chores, child rearing, home care, and emotional labor, has the effect that it is often not seen as “real work”.

changed their snowplowing pattern, turning from “main roads first, then sidewalks and smaller roads” to the exact opposite. The city council found that the original pattern had been planned with traffic behavior in mind that primarily represented men, who mostly use main roads to commute to and from work, while women “daisy-chain” various tasks, relying more on side-roads, and walking more. With the new system, walking accidents due to snow on sidewalks decreased significantly. Thus, the system change had positive impacts on women’s everyday ways (and lives) as well as on the city’s health costs, as the number of hospitalizations due to these accidents decreased, too (“Invisible Women” 2019).

In HCI/Computer Science, renowned work on gender differences includes, for example, Sherry Turkle’s “Computational Reticence” (1986), and Fisher and Margolis’ “Unlocking the Clubhouse” (2002). Both study gender issues, the former as related to the choice of field of study and hobbies, and the latter, continuance in CS careers. They look at what image women and men have of computers, in general, or of CS studies, where these images come from, and how they impact who enters a field, and who stays. Other works focus on how software can implement different features to become more usable for women, for example, to help them steer clear of harassment on social networks.

Thus far, we have talked about sex and gender in a very binary way, with male/masculine on one side, and female/feminine on the other. However, both biology and gender studies, have now accepted that gender is not binary. This is also recognized in various countries and cultures, and their respective legal systems – some have recently adapted, some have already long been aware (Breslin and Wadhwa 2017). Germany’s “dritte Option” (“third option”), is not a legal third gender, but rather, an “other” entry, for people who do not fit into the hitherto binary options for various reasons. Austria’s third option, on the other hand, is at the time of writing, explicitly limited to **intersex** people who are born with physical characteristics (anatomy, hormonal constellation, reproductive organs) that do not fit the constructed norms of male and female (UNFE 2015; van Lisdonk 2014; Ghattas 2019). As gender is assigned to people at birth based on their physical characteristics, it is possible that the assignment does not match their actual gender. The term **trans** or **transgender** describes a person whose gender does not match the one they were assigned at birth. In contrast, a person whose gender fits the one they were assigned at birth is called “**cisgender**” or “**cis**” for short. Finally, a person whose gender identity is neither masculine or feminine is described as **non-binary**.

However, people can hardly be described by only one characteristic. This is where the notion/methodology of **intersectionality** comes in: different socio-demographic factors interact with each other, forming an identity, and at the same time constructing how people are seen, perceived, and treated, and which resources they have access to (Haraway 1988; Harding 2015). Taking this congregation of attribution by oneself and others into account, it is obvious that these factors cannot be analyzed separately. As Schlesinger, Edwards and Grinter put it: the impact of identity on context and design in HCI must be a focus of intersectional analysis. For example, a white cis woman will have

different experiences at her work place than a Black² cis woman, but also different than a White trans woman or a First Nations non-binary person.

2 Following the recommendation in the *Publication Manual of the American Psychological Association* (Sixth Edition, 2012), racial and ethnic groups in this work are described by proper nouns and are capitalized. Thus, I use Black and White, instead of black and white. Furthermore, using adjectives would equal using a color to describe a person, which according to the APA manual is considered pejorative. I fully agree with this consideration and thus follow the recommendation to not use colors, but rather proper nouns to describe a person's heritage where it is important.

Definitions: Human Computer Interaction (HCI)

While Human Computer Interaction (HCI) is a broad field, the outline of the *Wiley Handbook of Human Computer Interaction* provides a good overview of what constitutes this research field (Kirakowski and Norman 2017, p. 4). In short, HCI studies the interfaces between humans and computers.

This involves the creation of interfaces (studying general design issues and the factual process of design) and their evaluation (including the question of which factors to evaluate, and different methods of evaluation), and the study of end users as both giving input, and receiving output. For the second part of the *Handbook*, Kirakowski and Norman pose the question of the effect technologies have on “us humans”, studying the interface itself, and the interactions taking place. Shackel (1997) also lists issues of accessibility and special needs, and social aspects of users and communities as important parts of HCI – already since the early days, when the integration of mainframe computers into organizations were the focal point of research.

As stated later in the *Wiley Handbook* by Breslin and Wadhwa (2017, p. 71), gender is often present in HCI work – but often in an implicit way that has not been reflected on. Gender stereotypes and roles impact how people interact, not only, but also with computers. For example, Turkle (1986) discusses “the social construction of the computer as a male domain”, and criticizes the “culture of young male programming virtuosos [that] tends to dominate the computer cultures of educational institutions [...]” (p. 44).

Users’ needs and requirements are rather specific depending on the context, and their adaptation to and use of systems are transient and hard to foresee. Still, as the US Air Force learned in the 1950s, even in a seemingly homogeneous group like their fighter pilots, there is no average person. They took ten measurements of 4,000 pilots to define an average in order to create a cockpit to fit that average, as most people should fit into that average. Even using only three of the ten dimensions resulted in less than 3.5 percent of the pilots actually fitting into the cockpit. Obviously, there was no average, and no two pilots with long arms would have legs of the same length or torsos of the same size. Finally, instead of designing for the non-existent average pilot, USAF decided to design for extremes, and to make their designs adjustable (Hough 2015). Assumptions about what users want to do with a system, and about how they want to complete these tasks, create awkward situations – such as a gym system assigning a customer a locker in the wrong locker room, because her title (“Doctor”) could only mean she was a man (Wachter-Boettcher 2017, p. 6).

Gender and Design/Research: Pitfalls

As stated above, gender affects what we do, and how we do it – including the interactions we have with other people, the research we do, and the designs we create, and for whom. Breslin and Wadhwa (2017) talk about some pitfalls that come with the (gender) norms incorporated in our societies and technologies.

In the following, these pitfalls will be introduced and illustrated with examples of product design issues, such as speech recognition and image tagging software, but also more tangible, non-computer-related things such as protective gear. The explanation of these pitfalls tends to use examples with only binary genders. This is mostly due to the fact that if gender is considered, this mostly happens in a binary way (as seen for example in Schlesinger, Edwards, and Grinter 2017).

More detailed examples from areas of business and industry can be found starting on page 32.

I-methodology: “Everyone is like me”

Designers and developers often use themselves, or their immediate surroundings, as models for the things they create (Breslin and Wadhwa 2017, p. 72), and derive ideas for products from their own experience. A common adjacent is the **hallway testing** method, where designers/developers ask people walking down a hallway to test a design/product.

While these methods are not per se negative, they introduce a certain bias into the pools of ideas and test subjects. According to data quoted in Breslin and Wadhwa (2017, p. 72), 80–90 percent of position holders in tech are men, and 92–94 percent are Whites and Asians at, for example, Google, Facebook and Twitter. This bias is not an issue specific to Silicon Valley or the USA, but rather to the field of computer sciences in general: At TU Wien, my alma mater and Austria’s largest technical university, the numbers are similar: about 85 percent of the students in computer science are men³ who are likely to share a certain set of experiences such as having studied at a HTL⁴ (Grabher, Unger, and Zaussinger 2014).

These sets of experiences lead to the development of products that perfectly cover the represented majority group’s needs, but completely exclude others’ requirements and desires. For example, many health trackers in their first (publicly available) iterations could not recognize the workout/activity of someone pushing a baby stroller.⁵ Similar bias in training data sets can impact the basic functionality of software even making it completely dysfunctional (DeVries et al. 2019). While men’s voices can operate voice-controlled software without any issue, women have been reported struggling to do so in various settings (Criado-Perez 2019; Tatman 2016).

3 https://tiss.tuwien.ac.at/statistik/lehre/studien?filter%5B%5D=ohne+Mitbeleger&kategorien%5B%5D=Studienrichtung&kategorien%5B%5D=Geschlecht&prozent_ anzeigen=false, retrieved 2019-06-June 17, 2019

4 HTL = Höhere Technische Lehranstalt. A school comparable to a High School, training students specifically for technical jobs, for example programming, electrical engineering, or structural engineering.

5 <https://www.theguardian.com/lifeandstyle/2019/feb/23/truth-world-built-for-men-car-crashes>, accessed October 17, 2019-10-17

Drawing from their own needs, experiences, and requirements, many ideas and products developed by the well-situated white men in IT focus on and serve privileged lives. Ride-share services, food delivery and cooking services, laundry services with pick-up and delivery – many recently very successful companies of the share economy can be dubbed “technologies replacing [the developers’] mums.”⁶

One Size Fits All: “Everyone is the same”

The issues of “one size fits all” and “works for me” thinking derive from the above-mentioned situation that the work forces in IT and HCI are very homogeneous.

The issues stemming from this methodology can be seen in both hardware and software. When hardware is modeled with only some users in mind, phones, game controllers, and sensors might not work for the actual users. When people are meant to interact with something, it should fit – but often, products are too big for women’s bodies. Smartphones and game controllers that are hard to grasp hold of, and heart-rate chest-straps that cannot be tightened enough to fit one’s chest, or cannot be worn in combination with a sports bra, are possible consequences. Worse than these inconveniences, studies show that protective gear, no matter the area of application, is modeled to accommodate men’s bodies – thereby leaving women unprotected, occasionally even hindering them (Criado-Perez 2019).

In software, for example, web search term suggestions, and machine learning/artificial intelligence (AI) can also be seen as an application of the “one size fits all” approach. Here, large amounts of data are used to draw conclusions of what is “important” or “correct” for users. However, “big data” also has its biases – and can thus result in strange “majority votes”. For example, when Microsoft presented their AI “Tay” to twitter, users taught it how to be racist and sexist within mere hours. The auto-complete feature embedded in Google’s search function shows a similar bias towards sexist and racist stereotypes, asking, for example, if women are “attracted to money” (Criado-Perez 2019).

Gender Stereotyping: “All women are alike”

The pitfalls mentioned above are cases of implicit gendering, where people just did not think *that* gender has impacts on their work. Gender stereotyping, however, is a sign of explicit gendering – here, people did realize that there might be differences between women and men. However, reflection on *why* women buy and use certain things, in comparison with the products bought by men, did not necessarily take place. Making explicit the assumptions used and ways that decisions were made provides a way to improve design/research. However, explicit “design for women” can result in “shrink it and pink it” (SNP), or bad stereotyping. Literal SNP has been, for example, repeatedly applied to tool kits. While it is correct that many standard tools like wrenches are too big for women’s hands⁶, making them pink is an unnecessary instance of othering (Dervin 2015). Pinkification can also be seen in children’s toys, for example, toy medical kits, or campaigns that aim to make the STEM field more interesting to girls by showing, for example, that chemistry has many applications in cosmetics.

⁶ <https://www.businessinsider.com/san-francisco-tech-startups-replacing-mom-2015-5>, accessed 2019-10-17 October 17, 2019

Design approaches such as SNP cater to stereotypes and reproduce them rather than making technology more inclusive. They also make it hard for men to use the “pinkified” designs. Furthermore, they disregard the fact of inter-group diversity, such as the existence of tech-savvy women, and the many differences between women in relation to culture, sexuality, socioeconomic status and other categories of difference (Breslin and Wadhwa 2017, p. 73).

But not only are some products made “more feminine” in order to sell them to women: marketing and design also work the other way around, making products “more manly” in order to make “unmanly” products more appealing to men. Examples of this include, but are not limited to, chocolate,⁷ and hot sauce.⁸ Hot sauce in Central Europe and the US often has a masculine connotation, with labels depicting symbols of death and destruction (using skulls, fire, and the like) while in Latin America, labels tend to depict motherly figures, as cooking and seasonings have a female connotation, and hot sauces are nothing other than a seasoning. Similarly, preparing meat outdoors has been reframed as a “manly BBQ”.⁹

One famous, and often-used example of gender stereotyping in HCI/computer science is the prompt to “explain something so your grandmother can understand it”, which is often used in job interviews, or oral exams, to see if someone really understands a concept. This completely ignores the option that the testee’s grandmother might be an avid user of the concept in question.¹⁰

To avoid stereotyping, many institutions use personas, which are abstract representations of users. Using personas is meant to help developers/researchers identify and understand their target audience (Pruitt and Grudin 2003). Personas may include demographic information, goals, and scenarios involving the topic of research, for example, what a user might want to accomplish using a software. They can shed light on different needs and requirements of different user groups, when applied properly. For example, good personas would show that the usage of a washing machine differs strongly depending on household size, or if the washing machine is placed in a student dorm. However, creating good personas requires time, information, and insight. When inadequate time is put into creating the personas, one runs the risk of relying too heavily on stereotypes, thus introducing into the project what one had tried to avoid in the first place.

Gender difference: “Women are inherently different”

There are some things where men and women are different; some things, where male and female users are different; and some things where people are different. Gender is just one parameter in the big picture – but these differences should not be seen as natural or inherent. Instead, as stated above, sociocultural context plays an important role – gender roles vary, depending on when and where one

7 <https://twitter.com/freakingmuse/status/731057719403270145>

8 <https://www.youtube.com/watch?v=jitjTtWTHYQ>

9 <https://www.vox.com/the-goods/2019/6/28/18760073/barbecue-grilling-men-stereotype>, accessed 2019-10-17 October 17, 2019

10 <https://twitter.com/danrkports/status/1108083064104587264>, accessed 2019-10-17 October 17, 2019

looks (Ceci et al., 2009 in Breslin and Wadhwa 2017, p. 73). For example, while pink is rather strictly assigned to girls and blue to boys nowadays, these assignments were less strict until the first half of the twentieth century (Stimpson 1930; Paoletti 1987).

Prevalent issues when looking at gender difference are: *binary division*, *mashing up sex and gender*, and *immutability of gender* (Keyes 2018; Schlesinger, Edwards, and Grinter 2017; Breslin and Wadhwa 2017). Binary division has been discussed in the chapter “Definitions: Gender” (page 4). Mashing up sex and gender is easily explained: you say “gender”, but look at physical things, like body fat distribution (see chapter “One Size Fits All: “Everyone is the same”, page 9). Immutability of gender describes the assumption that gender is something fixed – once assigned, this assignment cannot be changed. This disregards the fact that the same person may sometimes challenge gender norms, and sometimes adhere to them, depending on context, place, or time. For example, a person may very well feel content within their assigned role for a time, but later in life want to break it. This can also be temporary, or depending on a situation – a hacker who likes to create robots may sometimes just want to bake cookies. The former does not make her less of a woman, the latter does not make her less of a hacker.

Case Studies

In this section, discussing the case studies from Schlesinger, Edwards, and Grinter (2017), as well as a few examples from the industry, we will see different kinds of gender research. Some will be applications of the pitfalls that have been described in the previous section, while others show how to successfully avoid them.

Schlesinger, Edwards, and Grinter (2017) conducted a meta-review analyzing how users are identified or classified in HCI research. To that end, they collected 140 papers published between 1982–2016 by CHI¹¹ on gender, ethnicity, race, class, and sexuality. Findings include that identity-focused research tends to analyze one dimension of identity at a time (for example, gender OR race, but not both). In addition, the authors found that research on gender and socio-economic class is more common than on ethnicity and race. The paper contains helpful recommendations on including intersectionality in HCI research, such as clear reporting of context and demographic information, inclusion of author disclosures, and deeper engagement with identity complexities.

Each of the following case studies will be analyzed as follows: I start with a short description of the research and, when applicable, the associated project(s) and their connection to HCI, followed by gender aspects included in the research. Each analysis ends with answers to the question: “What can we learn from this?” A summary of the lessons one can learn from the discussed papers, together with specific recommendations targeted at researchers and research funding organizations is presented in the section “Discussion and Recommendations” on page 34.

The papers were selected in order to cover the wide range of topics included in HCI (see table below). Not all authors have made explicit where their research took place. Still, a strong bias towards works from the US (4) and the UK and Europe (3) is visible. One case study is rooted in Asia (ground work has been done there, some authors have US affiliations).

While I initially looked for ways to meaningfully engage with gender as a variable in HCI research, my focus broadened to also include the structure and context of the research projects while engaging with their presentations.

¹¹ The Association for Computing Machinery (ACM) annual Conference on Human Factors in Computing Systems (CHI) is considered to be the most distinguished in the HCI domain.

Case Study/Reference	HCI connection ACM Classification Keywords	Gender Aspects
1: Blackwell et al. 2016	studies how LGBT parents use social media H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.	gender as important part of LGBT parents' identity
2: Karuei et al. 2011	studies tactile interfaces H.5.2. User Interfaces: Haptic I/O — Tactile & Haptic UIs, User Interface Design, Handheld Devices and Mobile Computing, Multi-modal Interfaces	gender as independent variable
3: Otterbacher 2015	studies how bias may be introduced into data sets created with gamification systems I.2.6. [Learning]: Knowledge acquisition. H.5.2. [User Interfaces]: Natural language.	gendered differences in labels assigned to images
4: Clarke et al. 2013	studies how digital media can be used to help survivors of abusive relationships H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.	target group: women who left abusive relationships
5: Ahmed et al. 2014	user-centered design and creation platforms to share experiences of sexualized violence H.1.2. Human Factors	target group: women who have experienced sexualized violence
6: Haimson, Brubaker, and Hayes 2014	studies use of language in online personal ads H.4.3 Communication Applications; J.3 Life and Medical Sciences: Health; K.4.1 [Computers and Society]: Public Policy Issues: Computer-related health issues.	target group: men who have sex with men
7: James DiSalvo et al. 2011	created a system to introduce young African American men to computing as a profession H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.	target group: men who dis-identify with computing as a profession

Table 1: Overview of discussed Case Studies, HCI connection and Gender Aspects

Content Notes for Case Studies

I tried not to include potentially trauma-triggering content in this work. However, some of the cited works deal with trauma-related topics. In order for all readers to be able to make an informed decision on whether they can/want to deal with the respective contents when looking into the case studies, I provide the following content notes:

Case Studies 4 (Clarke et al. 2013) **and 5** (Ahmed et al. 2014) deal with issues of **gendered violence**, specifically domestic violence and sexualized harassment in public spaces. Especially the article by Ahmed et al., however, contains quotes from people describing harassment against women.

Case Study 6 (Haimson, Brubaker, and Hayes 2014) deals with sexual health related language, focusing on men who have sex with men. It also touches on the topics of **sexually transmitted infections** and the HIV/AIDS epidemic of the 1980s.

Case Study 1: Blackwell et al. 2016

Blackwell, Lindsay, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck. 2016. "LGBT Parents and Social Media: Advocacy, Privacy, and Disclosure During Shifting Social Movements." In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 610–622. CHI '16. New York, NY, USA: ACM. <https://doi.org/10.1145/2858036.2858342>.

For this paper, Lindsay Blackwell, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck¹² looked at how LGBT parents use social media. LGBT parents here refers to people who are parents and identify as LGBT. Human-Computer Interaction here is the usage of social media: what kinds of social media are used? What do participants share, with whom, and how (for example use of friend group features to restrict information dissemination to people who are considered safe). The term "gender" is not contained in the papers' keywords, but gender is an important aspect of the research.

They group their findings into three primary themes (p. 614):

1. *"Detecting disapproval and identifying allies: LGBT parents use social media sites to obtain social cues that allow them to evaluate their safety in relation to others."* Contacts' reactions to, for example, news relating to marriage equality can be used by LGBT parents to determine whether it is safe to be open about their own identity with the respective contacts.
2. *"Incidental advocacy: LGBT parents become incidental advocates when posting online about their daily lives is perceived to be advocacy work."* Some participants feel that their mere existence and visibility are acts of advocacy and resistance in times of anti-LGBT politics (in the US). Each and every posting on social media can become the starting point for discussions on equal rights for LGBT people.
3. *"Networked privacy management: for LGBT parents, online privacy is a complex and collective responsibility shared with children, partners, former partners and families."* Posting about oneself often includes information about others. This means that the posting person has the responsibility of assessing whether it is safe and okay for other people to be included in these postings. For example, there may be unwanted consequences for a teenager if their classmates found out that the teenager's parents are LGBT.

What can we learn from this?

Blackwell et al. give **detailed information** about who their study participants are, presenting data on their age, gender identity, sexual orientation, children's age(s), and region (Blackwell et al. 2016, p. 613). Blackwell et al. suggest that some of the work may also be applied to LGBT persons who are not

¹² I opted to fully list all authors of a given paper when I first introduce it. With this, I want to acknowledge all authors' contributions, and hope to improve visibility of non-male authors. Finally, I want to break up the hierarchy implied by ordering authors and reduction of references to the first author(s), as this may not reflect the actual amounts of contribution.

parents, but that most is specifically about LGBT parents. Also, they do not blanket LGBT people, and for example, point out differences of same-sex versus bisexual couples, and the effects a transition has on trans people and their friends and families. This way, Blackwell et al. can show the exact kind of influence a certain part of their participants' identities have on their social media usage. Rather than giving very general and crude information on "non-hetero parents", the readers are presented with **meaningful insights** into LGBT parents' lives and/on social media.

Blackwell et al. pay special attention to the **vocabulary used** throughout the paper. Talking about gender identity, sexual orientation, cis/trans, "chosen names" vs. "legal names", they explain what special terms mean. By explaining the term cis rather than trans, they avoid presenting cis as **the norm**. Furthermore, they make explicit that during interviews, interviewers paid attention to the language used by each individual participant regarding their identity, so that the interviewer could adapt their own language for the rest of the interview. So, rather than using pre-found vocabulary and categories, **participants' self-descriptions and preferred expressions** were used (p. 614).

One significant quote from a participant explains what intersectionality means: "Both are intrinsic parts of who I am. I'm always a parent and I'm always queer. The way I experience the world is based on and influenced by being queer, and by being a parent. I can't separate those things." (p. 615). In addition to sexual orientation and gender identity, the authors also present location (rural vs. urban), and socioeconomic status as **dimensions of intersectionality** (for example, it can be easier to be out and proud in big cities as opposed to rural areas), thus further enriching their descriptions.

Case Study 2: Karuei et al. 2011

Karuei, Idin, Karon E. MacLean, Zoltan Foley-Fisher, Russell MacKenzie, Sebastian Koch, and Mohamed El-Zohairy. 2011. "Detecting Vibrations Across the Body in Mobile Contexts." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3267–3276. CHI '11. New York, NY, USA: ACM. <https://doi.org/10.1145/1978942.1979426>.

Idin Karuei, Karon E. MacLean, Zoltan Foley-Fisher, Russell MacKenzie, Sebastian Koch, and Mohamed El-Zohairy examine which body locations are more sensitive to vibrations and which are more affected by movement, whether/how visual workload, expectation of location, or gender impact the performance of tactile displays and whether users have subjective preferences to any of these conditions. The HCI connection is the usage of vibration emitters placed on the body to receive notifications when the user is not using/holding their device or if it is not stowed close to the body. In their findings, Karuei et al. report that gender has an influence on the detection rate of signals and response time, but also that this influence is neither consistent nor significant.

What can we learn from this?

In their experiments, the team balanced gender (male/female) "to allow the consideration of its impact, which could arise through, for example, gender-linked differences in body fat composition" (p. 3270). Thus, the research team manages to avoid the "One Size Fits All" pitfall, by **actively including women** in their research. Body fat composition (or distribution) however, is not related to gender, but to sex. This is **inaccurate language**. In addition, the implied hypothesis that body fat composition may have an impact on detection of signals is not properly tested, as Karuei et al. did not measure their participants' body fat. This is an example of **gender stereotyping**.

Also, the experiment falls victim to the pitfall of gender difference, literally positing "male"¹³ as the norm: "Feet are the baseline for sites, *male for gender*, sitting for movement, no workload for workload, and first trial for trial number." (p. 3271). The **normative aspect of language** is very important in gender research and should be considered (compare, for example, to Blackwell et al.).

Overall, Karuei et al. did try to make their research gender aware, but stopped just shy of touching on truly fundamental gendered issues. This is especially visible when they discuss that "thigh was among the least effective and least preferred stimulus site we tested; and yet, front pocket is a common location to stow a mobile device, particularly for men" (p. 3275). The gendered issue here, of course, is clothing itself, and how pants (and dresses, and skirts) designed for women often do not have pockets big enough for smartphones, if they have any at all.

In order to reorient this specific research more towards gender issues, the impact of (often gendered) clothing and (possibly gendered) differences in acceptance of emitters in certain body locations would

13 This, again, is imprecise or even faulty language. As discussed in the chapter "Definitions: Gender", the term "male" is usually used when referring to sex, while [the appropriate term](#) when talking about gender [would be](#), "masculine". ~~would be correct~~.

be interesting points for discussion. In addition to this and the actual impact of fat distribution, for example, *more emitter locations* and *intensity of signals* could have been studied.

Especially with locations such as collarbone, stomach, and thigh, there may be many more reasons for poor user response than those mentioned in the hypotheses, for example, dense androgenic hair.

Case Study 3: Otterbacher 2015

Otterbacher, Jahna. 2015. "Crowdsourcing Stereotypes: Linguistic Bias in Metadata Generated via GWAP." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1955–1964. CHI '15. New York, NY, USA: ACM.
<https://doi.org/10.1145/2702123.2702151>.

Jahna Otterbacher looked into the results of ESP, a Game With A Purpose (GWAP), which was used to crowd-source textual data about images. In ESP, users are (randomly) paired up and challenged to agree on as many words (labels) to describe an image as possible within two and a half minutes. For every word the players agree on, they earn points, and a bonus when they were successful with fifteen images. In order to create specific labels instead of very general ones, the game shows a list of taboo words. Taboo words are selected based on the frequency of their use to describe an image (also by other players). So, for example, two players may be shown an image of a table laid out for dinner. In the beginning they will earn points for labeling it "table", and "dishes", but when enough players have used these terms, labels will have to become more specific, for example "Hanukkah" if matching decoration can be seen on the image.

This work connects to HCI as it deals with the use of gamification to improve natural language processing. It studies how bias may be introduced into data sets used to train machine learning algorithms, which in turn can cause biases in automated decisions. Otterbacher states that there is already a noteworthy body of research on how tasks can be gamified, and on the incentives the games provide to the players (p. 1956). Gender is the focus of the research, with the goal being to find if and how gender stereotypes enter into the labels generated for images via ESP.

To do so, Otterbacher compares characteristics of labels for 18,916 images of men and 14,628 images of women, looking at (original emphasis, p. 1958f.):

- The proportion of assigned labels that are *adjectives*.
- The proportion of adjective labels that are *strongly subjective*.¹⁴
- The proportion of subjective adjective labels that have *positive/negative prior polarity*.
- The most frequent strongly subjective adjectives.

Otterbacher's results provide evidence that players tend to describe *how* they perceive the women to be (labeling them with adjectives) while describing *what* the men are (using nouns, for example occupations). Looking at the strongly subjective adjectives most frequently used to describe either men or women, Otterbacher finds that all reference either "physical appearance (for example, sexy, ugly, cute) or disposition or character (for example, angry, happy, fun)" (p. 1960). In detail, they found impressive gender differences: of the top ten subjective adjectives used to describe men, only two specify their appearance. In contrast, of the top ten subjective adjectives used to describe women, five describe their appearance. Overall, 2,425 images of women (16.6 percent) were labeled using the most

¹⁴ Subjective adjectives, as the name suggests, are adjectives used to express one's subjective belief or opinion about something.

frequent, subjective label of “sexy”. The same label was used to describe only 20 pictures (0.1 percent) of men.

Another part of the analysis considers the differences in labels for images labeled as homosexual. Otterbacher reports a higher probability for adjectives if the players believed the depicted person was homosexual. Accepting that stereotypical language and beliefs are at the same time descriptive (how something is) and prescriptive (how something is supposed to be), the analysis of the labels suggests similar expectations of heterosexual women and gay men (p. 1961).

Otterbacher also includes some thoughts about the images included in the game and the analysis. Considering that they were collected from online resources and selected randomly, they probably represent a wide variety of depictions of people. However, it is also important to remember there are remarkable differences in the images of men and women shared on the Web – for example, there might be more pictures of physically attractive rather than “normal” women available, while for men this ratio might be even greater. In order to find out if the gender biases discovered in their first analysis hold true for less biased sets of images, Otterbacher designed also analyzed images in contexts where labels had been assigned related to occupations (p. 1961). One example resulting from this in-depth analysis is that players have distinct biases against women doctors: half of the images of women doctors are labeled “nurse”, while the same label never occurs with depictions of men. This indicates the stereotype that women are mostly likely nurses, while men are doctors (p. 1962).

What can we learn from this?

The first thing, and most easy to do, concerns the **pronouns** Otterbacher uses when talking about people of unknown gender: Otterbacher uses “she/her” pronouns, for example on page 1957 (my emphasis): “The more expectancy-incongruent a person and *her* behavior appears to us, the more likely we are to describe the person with more concrete language [...]”.

Otterbacher discusses the fact that there are **more than two genders**, and **more sexual orientations** than hetero- and homosexuality, and reasons why both features can only be studied in simplified terms: “[W]e identified the subset of images with one or more of the following labels: gay, homosexual, lesbian. This resulted in four sets of images, based on ESP players’ perceptions of the subjects’ genders and sexual orientations [...] We were not able to explore the labels used to describe images of people of additional genders / sexual orientations as they did not appear with adequate frequencies in the dataset” (p. 1959).

Finally, Otterbacher explicitly considers **where the data for their research comes from**, and what kind(s) of bias(es) it may contain, and designs a second analysis to test the results of the first. When testing results for women against results for men, and finding differences, **looking at subgroups** can provide further insight. In this specific example, expectations towards women seen as heterosexual and men seen as gay, were found to bear similarities. For other research, these subgroups could be related to age, location, or education.

Case Study 4: Clarke et al. 2013

Clarke, Rachel, Peter Wright, Madeline Balaam, and John McCarthy. 2013. "Digital Portraits: Photo-Sharing After Domestic Violence." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2517–2526. CHI '13. New York, NY, USA: ACM. <https://doi.org/10.1145/2470654.2481348>.

Rachel Clarke, Peter Wright, Madeline Balaam, and John McCarthy worked with six women who had left their abusive partners in the last one to six years. The goal of their research was to explore how existing photography technology could be used to help women after a life disruption, such as leaving an abusive relationship. Most often in this area, HCI research considers communication technologies, and how they can be used by abuse survivors to maintain their social life, while at the same time evade their abuser(s). The project described in the paper I discuss here can be seen as a variation in which assistive technology is developed – here, the technologies themselves already exist, and their application and impact on the users' lives come under scrutiny.

The project focused on the use of photo-sharing, "as a particular kind of digital media sharing that supports face-to-face social interaction across a number of flexible formats; tangibly through prints, digitally on screens and storied through creating video sequences" (p. 2518). Photo-sharing (as presentations, in digital/analog albums, as prints, etc.) as well as photography itself incorporate acts of storytelling (by choosing what to show, and how), which in turn helps construct identity and build relationships.

Often in HCI projects, focus is on researching existing designs and artifacts, or gathering information for new designs. Probes are typically short-term projects of several weeks. The work of Clarke et al., in contrast, spanned several months, and revolved around issues of social justice, re-building identity, group discussion, and relationships. They fostered an open and long-term engagement to work on (individual) collections and displays of images in various (non-)physical states: printed, displayed on screens, and incorporated into videos.

A probe was designed in order to encourage the women to reflect on the things they saw as forming parts of their selves. It consisted of a digital camera, a sound recorder, a portrait frame, and a set of "inspiration tokens" and instructions meant to assist the participants in photographing or recording the aspects of their lives they wanted to share and/or retain. The tokens represented four main themes the women could include in their collections: people, sensory experiences, places, and objects (p. 2520). The research done in the early stages of the project revealed that it would be important for the women to create something they could take with them, something they could share both in the workshops and at home. This could help in building confidence, experiencing a sense of achievement, and affirming their agency.

Clarke et al. conducted ten sessions of two hours per week between November and February, with breaks for Christmas and school holidays. Initially, videotaping the sessions was planned, but the

women were uncomfortable with that idea, so documentation of the process was done through anonymized fieldnotes (after each session), interviews with the women's center staff, an outreach worker and the center coordinator, and a recorded group discussion (p. 2521).

In terms of the findings, Clarke et al. report that three qualities around the sociality of photo-sharing emerged through the process (p. 2521ff):

- *embodied expressions of relationships*: images taken of, with, or by friends; for example, on joint trips, illustrate their friendship. Two participants presented nearly identical photo albums showing images of time spent together on outings with the center, and their children. The images presented slightly different perspectives of the trips, with each woman carefully posing for the camera.
- *Balancing coming-together-ness with independence*: The women reported that the time in the workshops was a valuable chance to spend time with each other. Sharing photos in the group was even described as therapeutic by Samiya, the outreach worker of the center. The women also described the project as a means to think about their moving on and becoming more independent as well as sharing this process. Using photography and technology also gave the women a sense of independence and agency.
- *Negotiated sharing practices*: Carelessly shared digital content, for example, on social media, can become a tool of continued control and abuse in a context of domestic violence. Acknowledging the digital gender gap, women often need special help with their devices and online activity; for example, to keep their location safe from their abuser. Even in a safe space such as the workshop group, the women felt self-conscious about some images and the stories they told when put in context. Photographs of themselves, their families, personal places and objects were sometimes shared only after great thought was put into it.

The gender aspect of this project is visible in that while domestic violence can occur in relationships no matter the partners' class, age, religion, ethnic groups, genders and sexualities, the vast majority of serious and recurring violence is perpetuated by men towards women (p. 2517). Additionally, there is the above-mentioned digital gender gap, meaning that many women are less proficient and confident using various kinds of technologies.

Clarke et al. use gender as an analytical focus to restrict the population of their research, acknowledging that women who have experienced and survived domestic violence have different needs¹⁵ as compared to people without these experiences.

¹⁵ It is important to see the intersection of characteristics: "women who have experienced and survived domestic violence". Not all women [do](#) have these needs; on the other hand, people of other genders might share these needs.

What can we learn from this?

The first, most obvious lesson to learn from this project is: Remember that **there are no “standard users”**. Life disruptions, and a hopefully good life thereafter, exist. (Clarke et al. 2013; Wachter-Boettcher 2017). Designs and research touching on such life disruptions, be they domestic violence, or loss of loved ones, has to be handled carefully. Assumptions about a stable family or personal life, for example, must be thoroughly examined (p. 2525).

In order to handle life disruptions with the appropriate sensitivity, Clarke et al. got **in touch with the relevant experts for their participants**, in this case, survivors of domestic violence, early on. For example, together with staff at the women’s center, they examined the vocabulary they would use (p. 2519f), as well as the scope of the project (p. 2520), and the documentation of the project (p. 2521).

As seen in other case studies, and as recommended by Schlesinger, Edwards, and Grinter (2017), Clarke et al. provide readers with **rich details on their participants**, and how they got involved with the project: “[...] a self-selected group of six South Asian women who were aged between 25–38 and had been separated or divorced for between 1–6 years [...]” (p. 2520).

Again, as the topic at hand is very sensitive, Clarke spent three months volunteering at the women’s center before the project’s sessions started, in order to **get in touch with future participants**, and to inform the research questions (p. 2519).

In order to facilitate participation, a lot of thought went into the timing of the sessions, and breaks for Christmas and school holidays were observed in order to avoid clashes with the participants’ other commitments (p. 2520). Here, one should keep in mind the additional intersection of “women who are mothers and have experienced and survived domestic violence”. Additional intersections, but also personal preferences, and in the present case, content created, can be reasons for **diversity of needs within the pool of participants**. Clarke et al. report that they themselves and the staff at the women’s center had assumed the participating women would prefer similar levels of anonymity – but due to the different kinds of content created, different concerns were raised, and different types of anonymity were required (p. 2523). In short: even when people share many characteristics and experiences, their needs are possibly different. Needs may also include, for example, covering travel expenses and providing food during each session (p. 2520).

Knowing that photographs and the memories and emotions connected to them can present emotional challenges to those viewing them – especially in the context of domestic violence, the women’s center offered free **counseling for the participants** during and after the project sessions.

As described by Schlesinger, Edwards, and Grinter (2017), self-disclosure of the research team is important. Reporting on your own background as a researcher provides further information on why methods were chosen, and why questions were framed a certain way. Self-disclosure furthermore

highlights knowledge gaps between the researchers and research participants, who are the actual experts on their own situation. Finally, Clarke et al. (2013) use self-reflection to also reflect on the power imbalances and ethical challenges stemming from their position of power:

“[...] the importance of respecting diversity the women themselves would bring through their backgrounds and how these would differ from us as white middle-class researchers with limited experiences of domestic violence. Furthermore this meant reflexively acknowledging our positions in relation to power and ethics; who we are is not value neutral and being transparent about this position when approaching potential partners and participants was important in building relationships and trust” (p. 2519).

Case Study 5: Ahmed et al. 2014

Ahmed, Syed Ishtiaque, Steven J. Jackson, Nova Ahmed, Hasan Shahid Ferdous, Md. Rashidujjaman Rifat, A.S.M Rizvi, Shamir Ahmed, and Rifat Sabbir Mansur. 2014. "Protibadi: A Platform for Fighting Sexual Harassment in Urban Bangladesh." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2695–2704. CHI '14. New York, NY, USA: ACM. <https://doi.org/10.1145/2556288.2557376>.

The Protibadi (a Bangla word meaning "one who protests") project addressed a specific gendered issue, namely sexualized harassment and violence. The team, consisting of Syed Ishtiaque Ahmed, Steven J. Jackson, Nova Ahmed, Hasan Shahid Ferdous, Md. Rashidujjaman Rifat, A.S.M Rizvi, Shamir Ahmed, and Rifat Sabbir Mansur, created a platform where people who experienced sexualized violence could share these experiences. The paper documents the whole process of user-centered development of the web and mobile phone platforms. The project team conducted surveys, interviews, and focus group discussions at three different universities in Dhaka over the course of one year. The system designed to report, map, and share women's stories around sexual harassment in public places launched in August 2013. Three months later, the website had 110 registered users (twenty self-identified as men, the rest as women). The users had shared twenty-four reports of sexual harassment from different parts of Dhaka city.¹⁶ Three months after the launch, the team examined the strengths and limitations of the system by conducting user studies and monitoring of public responses to Protibadi.

Ahmed et al. reference several projects spanning the world, as well as localized projects, for example, from Egypt, whose insights and design approaches they were able to build on. In addition, they contacted female university students in Dhaka in order to find out the women's ideas about systems and applications that could help them. In the end, some features from different existing projects were combined into Protibadi: the user could quickly inform emergency contacts if they experienced a situation of harassment; they could document the location and nature of incidents, and they could post blog entries and thus share their experience with other users (p. 2697).

Ahmed et al. received 121 responses to their *online survey*, including 51 women, 42 men and 28 people who chose not to disclose their gender. Seven women and two men contacted the research team via email, phone calls, and personal encounters after completing the survey and provided further information. All of the 121 respondents stated that they had witnessed harassment of women in public places. All 51 woman participants additionally reported direct experiences of harassment in public places. Thirty-two women recalled experiencing either direct or indirect harassment more than once. The respondents agreed that public sexual harassment is a common and damaging experience of living in urban Bangladesh (p. 2697). The *focus group discussion* was conducted by a female faculty member at one of three participating universities. Thirteen women participated in the three-hour discussion. The women reported strong feelings of shame, sadness, and regret as well as defiance, anger, and a strong desire for change. In the *semi-structured interviews*, the women talked about their understanding and

¹⁶ A snapshot of the Protibadi web site from August 2013 can be viewed at <https://web.archive.org/web/20130815020151/http://www.protibadi.com/code/index.php/home>. The corresponding Facebook page ~~has been~~ was last active in 2016.

experiences of sexual harassment as well as their requirements and needs for design targeted at this issue.

The features requested and discussed included:

- *Help on the spot*: Many women said they wanted to alert bystanders to the harassment happening to them, in order to avoid, escape, or reduce the severity of the incident. However, the taboo surrounding sexual harassment also led some women to speak against this feature, as they anticipated feeling ashamed and embarrassed of receiving attention this way. Accordingly, in the user studies conducted three months after the launch, none of the interviewees had used the “Save me” button implementing this feature. Of the ten participants in the user studies, six said they would use it, while four said they would not in order to avoid embarrassment.
- *Reaching friends when needed*: In situations of harassment, it can be difficult to call friends and ask them for help, even if the victim has a mobile phone on them. However, all participating women said they would feel better if they had a way of letting their friends know where they were.
- *Sharing Experiences with Others*: Again, all participants agreed that it would be important for them to be able to share their experiences anonymously and to receive support. In addition to feeling better because of the support the poster would receive, readers could use the information on the platform and decide whether to take any precautions based upon others’ experiences.

Participants who had shared their experiences with others before (for example with family members and sometimes friends), experienced contradicting emotions of relief, but also deep embarrassment, anguish, and shame. These negative emotions were described as the most serious and pervasive consequences of their harassment experience.

The feedback collected in the *user studies* was positive, in general. Several interviewees noted the strong need for such a system, while being aware that it was only a drop in the ocean in relation to the omnipresence and seeming acceptability of public sexual harassment.

What can we learn from this?

As we have seen in other case studies, Ahmed et al. **give detailed insight into who their participants are**, and why they chose them: “Our choice of university women was dictated by three basic factors. First, most of them had access to technology like mobile phones and Internet, and so were more obvious first targets for an experimental system that made use of such tools. Second, in large part because of their education and socioeconomic standing, university women in Dhaka are often more attuned to problems of gender discrimination and sexual harassment, and more receptive to systems that combat it. Finally, as the inclusion of university-educated women in public life is often identified as a step towards gender participation and equality more generally, their exclusion from public space and participation through instances of harassment may be particularly insidious and damaging to the broader goals of gender equity and participation in public life” (p. 2697).

Ahmed et al. present several **dimensions of intersectionality** they deemed important for the success of the project. This occurs, for example, in the description of their participants, and in the discussion of future steps, where they mention an NGO planning to adapt Protibadi for rural areas with less mobile coverage, as well as a group wanting to extend it to include harassment in the workplace. Comparing Protibadi to other similar projects around the world, Ahmed et al. point out the need to consider a few localized differences, such as design metaphors, and assumptions, for example, about infrastructure (stable power supply, systems of law and governance), and the need for different modes of engaging locals. If work is done by non-local researchers, Ahmed et al. state that misunderstanding local culture and power structures could lead to non-adoption, suspicion, or plain indifference towards a project (p. 2696).

Having chosen a very specific target group, Ahmed et al. are aware that this focus creates **limitations** on the generalizability of their findings as well as the results.

Talking about the users of their platform, Ahmed et al. highlight that people self-identified as either women or men (see first paragraph). In their survey, in contrast, participants had the option to not disclose their gender. It is unclear whether users had the option to not disclose their gender on the platform, and whether there is **inaccuracy in reporting, or inconsistency in data collection**. It should go without saying that both inaccuracy and inconsistency should be avoided.

In their group discussion, Ahmed et al. tried to **make their participants feel comfortable** talking about sexual harassment. As some women did not feel at ease speaking about it, they were given the opportunity to write things down, and have them read for them by the woman who was conducting the discussion. This is also worth mentioning: all interviews and discussions with woman participants were conducted by women, and in rooms set aside for this purpose, to create a space where talking about this issue would not result in further embarrassment or shame. Audio-recording the interviews was the preferred documentation method for the researchers, but would only be done if all participants agreed. Furthermore, the participants could stop at any point of the interviews, request their data be destroyed, and walk out. One participant requested that her data should not be shared with male members of the research team, and in one case, an interviewee brought a companion who helped her share her story.

Case Study 6: Haimson, Brubaker, and Hayes 2014

Haimson, Oliver L., Jed R. Brubaker, and Gillian R. Hayes. 2014. "DDF Seeks Same: Sexual Health-Related Language in Online Personal Ads for Men Who Have Sex with Men." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1615–1624. CHI'14. New York, NY, USA: ACM. <https://doi.org/10.1145/2556288.2557077>.

Oliver L. Haimson, Jed R. Brubaker, and Gillian R. Hayes studied the language used by men who have sex with men (MSM) to pass on information relating to their sexual health in online personal ads.

This research can be placed in the area of HCI, as linguistic analysis can build a foundation for system architecture, and it can help in understanding how users (re)present themselves, and how they communicate. Haimson, Brubaker, and Hayes claim that similar methods of linguistic analysis could be applied to inform designers about how users "represent their health conditions, preferences, and activities" (p. 1622f). Their work provides a means to gain insight into HIV epidemiology as well as the discourse among specific communities.

In addition to the HCI scope, Haimson, Brubaker, and Hayes studied the temporal changes in SHR language used, for example, the (dis)appearance of phrases and words, and the increase of SHR language present in ads. They argue that the content of online personal ads may be used in STD research and prevention efforts, by making visible the local prevalence of STDs (p. 1615, p. 1623).

The team used open-coding techniques to find the sexual health related (SHR) language that MSM used on Craigslist personal ads. They collected 252,786 "men seeking men" (m4m) ads within a two-week period in August and September 2013 from all over the United States. A team of four coders, including two "gay-identified men" (p. 1618), coded a sample of 500 ads in total to build the dictionary for the study.

What can we learn from this?

The language used by Haimson et al. to describe their target group ("men who have sex with men", or MSM for short) is exact, and concise, while **not excluding/erasing** anyone. The terms "homosexual men" or "gay men" on the other hand would exclude bisexual men.

Again, we see an instance of **self-disclosure**, describing the coding team: "[f]our coders, including two gay-identified men" (p. 1618).

Haimson et al. **explain in detail how and why age groups were constructed, taking into account the specific history** of HIV/AIDS (p. 1620). Categories are defined, for example, based on the user's phase of life during the 1980s, when most people learned of HIV/AIDS (and thus, whether or not they were most likely sexually active at that time). Other groups consist of those too young to remember the HIV/AIDS epidemic, who may display more risky behaviors than older men (p. 1620).

Case Study 7: James DiSalvo et al. 2011

James DiSalvo, Betsy, Sarita Yardi, Mark Guzdial, Tom McKlin, Charles Meadows, Kenneth Perry, and Amy Bruckman. 2011. "African American Men Constructing Computing Identity." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2967–2970. CHI '11. New York, NY, USA: ACM. <https://doi.org/10.1145/1978942.1979381>.

Betsy James DiSalvo, Sarita Yardi, Mark Guzdial, Tom McKlin, Charles Meadows, Kenneth Perry, and Amy Bruckman developed Glitch Game Testers, a program to introduce low-income high school students with an interest in video games to work in computing. They state that in general, although African American men are passionate about video games, they are less likely to take an interest in the technology behind these games. This is possibly connected to how African American identity is constructed in stark contrast to being a "geek", "nerd", or "tech person": while African American identity is strongly connected to the body, athletics, and sexuality, the stereotypical hacker is depicted as ignoring the body and appearance. Comparable to how women cannot or only rarely see themselves as hackers, this mismatch in physical identity can lead to the disidentification, and thus rejection of (in this case) computing as a field of interest and future occupation.

To tackle this issue of participation and socioeconomic equity, James DiSalvo et al. developed a job training program to educate young African American men to "break open the black box" of games to learn about computing. Participants receive an apprenticeship as game testers working full-time during the summer, and part-time during the school year. Their task encompasses quality assurance for pre-release games from game companies such as Yahoo!, Electronic Arts, and Cartoon Network.

To measure effects of the program, pre- and post-test surveys were conducted with twenty-one Glitch participants. Students joining the program in 2009 were categorized as "oldtimers", those joining in 2010 as "newcomers" in the study. The survey measured the participants' perceptions of technical competency for both their peers and themselves, and the access to technical resources in their social groups. The pre-test was completed in the first week of the summer 2010 program, the post-test was administered eight weeks, about 280 working hours, after the pre-test in the last week of the summer. In addition, interviews and focus groups were conducted, and researchers spent over 800 hours finding, observing, and engaging with participants.

For the survey, participants were asked to list close friends and close family members, and to rank each person based on perceived technical expertise as well as four people they would go to for help if they had a question relating to computers and technology. For both lists (technical experts, and technical resources) they were asked to report relationship, closeness, and technical expertise of each connection. DiSalvo et al. report that participants were more likely to see their peers as technical resources after participation in Glitch than before, and that participants' overall access to technical expertise increased significantly. Results for participants' perception of their own technical expertise showed a much higher difference between oldtimers' pre-test and post-test self-ratings than those of the newcomers. DiSalvo et al. also report that oldtimers' self-ratings increased while newcomers decreased slightly (both not significantly).

In past evaluations, qualitative and quantitative results also suggested an overall positive impact on participants: five of the seven participants who graduated from high school in 2010 went on to attend college, and selected computer related majors. Of these students, only one had thought of going into computing before working with Glitch.

In conclusion, James DiSalvo et al. suggest that peer influence (for example influence oldtimers have on newcomers) can have positive impacts on technology adoption and identification. They make the following proposals for projects similar to Glitch Game Testers (p. 2970):

1. In order to motivate adoption and desire to learn among newcomers, pair them with oldtimers.
2. Changing the image of a technology within a social group can make knowing it more socially desirable (“cooler”, so to speak). Thus, introducing new technologies, especially those that require learning and acceptance of new users, may become easier.
3. The technical capital survey developed by DiSalvo et al. may offer an approach for HCI researchers to measure social stigmas among users or individual and group attitudes in various domains.

Their advice that “future research should consider social norms and group identification in the design process” (p. 2970) is an important lesson to take into any design process, no matter the domain.

What can we learn from this?

James DiSalvo et al. state that when recruiting participants, they did not explicitly include gender and race in their **selection criteria**. Instead, the entrance criteria were that students come from low-income households (measured by Free and Reduced Lunch eligibility) and express an interest in video games. All applicants who met these criteria happened to be African American young men – think about how intersections can change the direction of your project. Rather than being very strict and focused from the beginning, instead, think of “marginalized groups”, look at factors other than gender first!

James DiSalvo et al. highlight **how identities are constructed**, and what happens when one’s self-image and the image of a hobby or activity do not fit together (disidentification). They explicitly account for how masculinity differs by culture, contrasting White men who can easily identify as nerds, and Black men, who cannot.

Combining the knowledge of constructed identities and the intersectional lives of their participants, James DiSalvo et al. found an appropriate way to make computing interesting to them. In order to pave the way for newcomers into computing, projects often try to make it “fun” and “easy” and use elements of gamification. With Glitch Game Testers, participants had the opportunity to see computing as a desirable future occupation, and as something they CAN do.

Some details in the paper could be improved. For example, the term “male” is used when talking about gender, rather than “men”. As discussed above, when talking about gender, “man” and “masculine” are the more **appropriate terms**. Furthermore, **additional information on the project’s context** would

help readers gain a better understanding. For example, information on the student population at the participants' school(s) would underline the reported impact on college enrollment.

Outside Academia

As Human Computer Interaction is not solely an academic field, a great deal of research and development also happens in the industry. Thus, it is only reasonable to also look into the industry to see how gender factors into design processes and decisions, and how disregard of gender dimensions can cause product failures.

For this insight into the industry, I have found examples in the work of Sara Wachter-Boettcher (2017), Caroline (Criado-Perez 2019), and Carol Reiley (2016).

Ignoring the gender dimension in product design and development can lead to strange outcomes. While it has been possible for users to track all kinds of aspects of their life and body functions since Apple first introduced the Apple Watch, tracking one's menstrual cycle was only added as a feature in 2019.¹⁷ Reiley (2016) reports that Mattel's "Hello Barbie" was initially not able to understand its prime target group (little girls). Similar issues of voice recognition failing to recognize feminine voices are reported by Criado-Perez (2019) in the setting of voice-controlled, on-board computers in cars. Both can be traced back to the fact that voice recognition software is often trained only on adult's masculine voices. Another, rather infamous product of slanted training and test data happened to Google Pictures, when some users realized it tagged Black people as gorillas – which is a symptom of the test and training data sets basically not containing enough pictures of Black people to distinguish them from gorillas (Wachter-Boettcher 2017). Google's "solution" to this issue was to simply delete the tag "gorilla" from the software.¹⁸ This issue is not limited to Google alone: Microsoft and HP software for facial recognition had difficulties recognizing non-White people,¹⁹ and Apple's FaceID had issues telling apart Asian women.²⁰

While these failures sound pretty much like "first world problems", the data sets and software packages causing them are not used only for harmless things like toys and handy features like tagging images – they are also being used to teach self-driving cars, and security/surveillance systems with facial recognition. These surveillance systems often have their best results in exactly one category: White men; the worst performance has been reported for Black women (Buolamwini and Gebru 2018).

Disregarding gender dimensions affects not only software, however. Physical systems have similar issues, based on the history of design where men have long been centered in research, and their bodies accepted as a default. Instances of physical designs disregarding women are smartphones and other hand-held devices that are too large for small hands, or fitness trackers that cannot track movement/workouts when users are pushing a stroller (Criado-Perez 2019).

17 <https://www.cnet.com/how-to/apple-watchos-6-cycle-tracking/>, accessed ~~October 29, 2019~~ 2019-10-29

18 <https://www.theverge.com/2019/6/11/18661128/ai-object-recognition-algorithms-bias-worse-household-items-lower-income-countries>, ~~2019-10-17~~ October 17, 2019

19 https://www.pcworld.com/article/209708/Is_Microsoft_Kinect_Racist.html, accessed ~~2019-10-17~~ October 17, 2019

20 <https://www.theinquirer.net/inquirer/news/3023199/apples-face-id-tech-cant-tell-two-chinese-women-apart>, accessed-~~2019-10-17~~ October 17, 2019

What can we learn from this?

The most important lesson to be learned from the incidences mentioned above is to **be critical of the things you use**, whether data sets, parts of software, or full applications. Try to find out what the people behind your resources had in mind, and where they might have overlooked something. Think outside your (and your colleagues') box, find stress cases (Wachter-Boettcher 2017) and try to handle them well. Ask yourself whether things have to be done the way they have been done.

Try to **consider the situatedness of knowledge** (Haraway 1988). For data sets, whether they are statistical data for mining information, or survey data from user studies, regard how they came to be. Where did the user study take place? When? Who was interviewed? What was the context of the survey? What cultural, temporal differences might change the outcome of a similar survey if it took place now and where you are located?

Discussion and Recommendations

This section consists of a summary of the lessons one can learn from the case studies presented above, and specific recommendations for both research funding organizations (RFO) and researchers. I illustrate each detailed recommendation with examples from the case studies above, and where available, provide further interesting reading material.

The most basic, but probably hardest, thing to do is to **be aware of the gaps in your idea, data, and plans, and try to mend them**. Finding those gaps can be difficult if you talk only to people from your own domain; finding experts able to look at whatever you are trying to do from a different perspective is helpful. For starters, have a look at the GenderMag method (Burnett et al. 2016), which offers help finding gender inclusivity bugs in software, or read about how forms and databases may exclude or include people who do not fall into a cis gender binary (Spiel, Keyes, and Barlas 2019).

Based on the recommendations below, a Gender Inclusiveness Checklist along the lines of ethics checklists such as the Horizon 2020 ethics self-assessments may be established. I will investigate this path in the context of future research.

Include Gender Studies and Intersectionality Basics

Both researchers and RFOs should learn about, and could mention in project documentation, some basics from the fields of Gender Studies and Intersectionality. This would clarify for readers what your work is based on, and some may even learn something new. In addition, it would contribute to the normalization of gender studies in CS.

Cues to follow are, for example, the point made by Otterbacher (2015) that there are more than two genders, and more sexualities than just hetero- and homosexuality;²¹ or the short discussion on the construction of identity by James DiSalvo et al. (2011).

Remember the normative impact of language

Use inclusive language to avoid perpetuating exclusionary stereotypes. For example, use singular they when referring to “abstract” users, state all options and choices when presenting statistics (“45 % identified as women, 44 % as men, 7 % as non-binary, 4 % chose not to disclose their gender”), and consider not using “traditional” defaults (for example, order sexuality by alphabet, ranking “asexual” first). Further, use participants’ self-descriptions and preferred expressions (Blackwell et al. 2016), refer to experts to check one’s vocabulary and assumptions (Clarke et al. 2013), and use inclusive terms (Haimson, Brubaker, and Hayes 2014). Finally, use accurate language, and the correct terms to refer to sex and gender respectively. For details, see the discussions on Karuei et al. (2011), Ahmed et al. (2014), and James DiSalvo et al. (2011). When talking about participants’ gender, make explicit what gender(s) you talk about – and do not assume your participants’ gender, but explicitly ask them.

²¹ “We were not able to explore the labels used to describe images of people of additional genders / sexual orientations as they did not appear with adequate frequencies in the dataset.” (Otterbacher 2015, p. 1959)

Research funding organizations (RFO) should, when reviewing calls and applications, pay attention to details. Being a linguistics specialist is not necessary; looking for the small things is a first step. For example, in Haimson, Brubaker, and Hayes (2014), there is no indication whether they found ads posted by transgender persons, whether they were included, or why they would have been excluded from the corpus.

Provide rich information on context

With a profound understanding of the context, readers and people building on your work can better understand your choices. Explain why you chose to include gender in your research,²² and why you decided (not) to break the binary construct of gender. Discuss the origin of pre-collected data you use for your study – for example, workplace environments probably have changed over the past thirty years. Make explicit what you base your work on, so your readers can contextualize it, or even learn something new (as seen, for example, in James DiSalvo et al. 2011 discussing the construction of identity).

For both **researchers** and **RFOs**, this means: keep your eyes open, widen the horizon of your project. Specific men's issues exist, for example, inaccessibility of emotional/medical/mental health issues, taboos surrounding certain jobs or topics, etc.

No matter why you did it: using gender without explaining the motivation is not acceptable. **RFOs** should request further information from projects that use gender without any explanation. It should be clear why gender was included, and why the researcher(s) decided to build on a certain definition (binary, biological, self-identified, etc.). Basing research only on gender, without any further information about the participants, is highly questionable, and should thus not be accepted without further explanation.

As a **researcher**, in order to be sure that one does not overlook any important influences on users and their usage of technology, detailed information on one's participants as well as a study's context, the source of the data used and the binning of variables are essential.

Discussing the intersections making up the participants' identities and influencing their access to and use of technology can be seen, for example, in Blackwell et al. (2016), Clarke et al. (2013), and Ahmed et al. (2014). Reporting the context in which the project has been/will be conducted helps in understanding the project, its development, opportunities, and limitations. When using pre-collected data for a study, considering the context this data was created in, as described by Otterbacher (2015) and Haimson, Brubaker, and Hayes (2014), can be very revealing. Finally, making clear how results

²² For example, Schlesinger, Edwards, and Grinter (2017, p. 5415) categorized motivations they found in their corpus as follows:

1. differentiating subjects for statistical analysis based on gender
2. restricting the population of study to subjects of one or more specific gender groups
3. focus on HCI interventions or innovations around a gendered issue

were grouped, and why, affects statistical analysis outcomes and interpretation (see Haimson, Brubaker, and Hayes 2014).

As a researcher, writing out these things will help you recognize if you, or the people whose work you build upon, missed anything of interest.

In the papers discussed in this work, some of these intersections were:

- the selection criteria of liking video games and stemming from a low-income home, which resulted in mostly African American young men being recruited for the project (James DiSalvo et al. 2011)
- self-identifying as LGBT and being a parent (Blackwell et al. 2016), who also discuss differences for LGBT parents in rural vs. urban settings
- the specific subgroups of women who are (perceived to be) heterosexual, and men who are (perceived to be) gay in Otterbacher (2015)
- the very diverse needs and personal groups within the rather specific group of domestic violence victims, which led to differences in technologies used, and contents shared in Clarke et al. (2013)
- in their work on Protibadi, Ahmed et al. (2014) note great differences in the perception and (non-) acceptance of sexual harassment in different cultures, but also within Bangladesh depending, for example, on the rural/urban, and workplace/public space dimensions.

As the last two points make especially clear, taking intersectionality into account means that there are no standard users. Finally, another important thing to do is to report on a project's limitations. Make absolutely clear who you are talking about. Which men, which women do your findings apply to?

Make your project accessible

Making research more inclusive is not always easy, and comes in many different shapes and sizes. However, it will make it possible to work with people who were not enabled by the circumstances until now, even if they had wanted to be part of research and development in HCI.

Researchers should remember to think about life outside their project and their institution. It is important to rank all hindrances on the same level, or at least make clear why one facilitation was deemed more important than another.

This begins by actively including people who are not the ubiquitous standard (mostly white, able-bodied men), and goes all the way to offering lunch during workshop days and tweaking the project schedule to accommodate for care responsibilities (Clarke et al. 2013), as well as making counseling available for participants (Clarke et al. 2013; Ahmed et al. 2014). Ahmed et al. (2014) also made various efforts to make their participants comfortable talking about a sensitive topic: participants could choose their interviewer, they could stop their interview at any time, and were also informed that they could subsequently withdraw their consent to data use.

Of course, in order to make your participants feel at ease, you need to find participants in the first place. For example, a team member of Clarke et al. (2013) spent some time volunteering at the women's center to gain the trust of the women spending time there. James DiSalvo et al. (2011) contacted schools and put up flyers there, and Ahmed et al. (2014) used the social media groups of local colleges. In short, as a **researcher**, one must find places the target group frequents. It should be clear from the start that one is looking for participants, and that the research team will try their best to accommodate the participants' needs. Contacting advocates, representatives, and other experts, such as the women's center's staff will provide initial insight into the needs and requirements your participants might have, allowing you to plan for their accommodation early on.

Funding organizations should take on the role of enablers for inclusion. For example, the availability of funding for childcare for study participants or counseling for participants and researchers could be made explicit in calls, forms, and other documentation. This might even encourage and inspire researchers to look at new target groups to do research on. Having a more inclusive mindset can also have positive impacts on the diversity of research teams, which in turn can improve the teams' research outputs.

Make methods accessible

Often, we build our work on great existing concepts and research. However, all predecessor technologies have been developed in and for certain contexts, so sometimes it will be necessary to make adaptations. In short, you need to be aware of the participant's perspective on the method(s) you want to use.

In the case of Clarke et al. (2013), the preparation time with the center's staff also led to adaptations to the original research plan. Clarke et al. (2013) decided to stick to the concept of cultural probes, but to rename them "digital portraits" to make the concept more easily approachable for their participants.

And while many projects that aim to include underrepresented groups in programming try to do so by making it more fun, James DiSalvo et al. (2011) managed to find a way to show African American teenagers that coding was something they could pursue as a career.

Consider the position of the research team

As pointed out by Haraway (1988) and Harding (2015), one's identity affects the information and knowledge one has access to, or how we see the world, and what we get to see of it. Thus, it is important, and also recommended by Schlesinger, Edwards, and Grinter (2017), to disclose some information about the author(s) of a paper or the project team. As seen in Clarke et al. (2013), this means reflecting power imbalances, and recognizing that one might not know as much about the user's desires as one thinks. Or, as seen in Haimson, Brubaker, and Hayes (2014), it can mean disclosing that a certain topic hits "close to home". As discussed by Schlesinger, Edwards, and Grinter (2017), self-disclosure should not be forced upon individuals, as disclosing information about oneself might have

negative consequences – the world we live in, after all, follows rules that are often (among others) ableist,²³ sexist, and racist. All in all, it is important to consider one’s position towards the research topic. If, how, and to whom this position is disclosed, should be decided individually. Sharing some information with participants can help build trust, while the same information may not be necessary or appropriate for a write-up.

Brulé and Spiel (2019) discuss the influence of the researcher’s position in the context of participatory design. Their idea of systematic reflexivity, however, could be applied in other contexts.

Enabling thorough discussion and reflection

To do all, or even some of the above, takes time. People need the time and space to ask and discuss (sometimes rather difficult) questions. I hope that the present work makes it easier to find the right questions to discuss. **RFOs** can make this work (and the card deck I plan to create based on it) available with their calls, and request applicants to apply it.

One means that can be used to facilitate discussion and reflection is the deconstructive “mind scripting” technique described by Allhutter (2012). The technique is used to encourage people to try and remember why things were done how they were done, especially in settings where gender and technology “co-materialize” (p. 684).

I am aware that some of the recommendations can be difficult to implement, and that these difficulties vary between projects. My own experience from discussions with researchers at TU Wien have revealed that often intersectionality is not a known framework, and people do not even know where to start when they have to describe the “gender dimensions” of their projects, as is often requested by RFOs. I hope that this work will make it easier to get started with the topic.

All of the recommendations listed above can, of course, also be applied in **industrial research**. It may be an unpopular opinion, but I think that giving people more time to develop their ideas, to discuss their motivations, and to reflect upon their work, will ultimately make for better artifacts.

23 Ableism is defined by Merriam-Webster as “discrimination or prejudice against individuals with disabilities” (“Ableism” 2019).

References

- “Ableism.” 2019. Merriam-Webster.Com. June 26, 2019.
<https://www.merriam-webster.com/dictionary/ableism>.
- Ahmed, Syed Ishtiaque, Steven J. Jackson, Nova Ahmed, Hasan Shahid Ferdous, Md. Rashidujjaman Rifat, A.S.M Rizvi, Shamir Ahmed, and Rifat Sabbir Mansur. 2014. “Protibadi: A Platform for Fighting Sexual Harassment in Urban Bangladesh.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2695–2704. CHI ’14. New York, NY, USA: ACM.
<https://doi.org/10.1145/2556288.2557376>.
- Allhutter, Doris. 2012. “Mind Scripting: A Method for Deconstructive Design.” *Science, Technology, & Human Values* 37 (6): 684–707. <https://doi.org/10.1177/0162243911401633>.
- Bardzell, Shaowen. 2010. “Feminist HCI: Taking Stock and Outlining an Agenda for Design.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1301–1310. CHI ’10. New York, NY, USA: ACM. <https://doi.org/10.1145/1753326.1753521>.
- Blackwell, Lindsay, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck. 2016. “LGBT Parents and Social Media: Advocacy, Privacy, and Disclosure During Shifting Social Movements.” In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 610–622. CHI ’16. New York, NY, USA: ACM.
<https://doi.org/10.1145/2858036.2858342>.
- Breslin, Samantha, and Bimlesh Wadhwa. 2017. “Gender and Human-Computer Interaction.” In *The Wiley Handbook of Human Computer Interaction*, 71–87. Wiley-Blackwell.
- Brulé, Emeline, and Katta Spiel. 2019. “Negotiating Gender and Disability Identities in Participatory Design.” In *Proceedings of the 9th International Conference on Communities & Technologies - Transforming Communities*, 218–227. C&T ’19. New York, NY, USA: ACM.
<https://doi.org/10.1145/3328320.3328369>.
- Buolamwini, Joy, and Timnit Gebru. 2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*.” *Proceedings of Machine Learning Research* 81, 1–15.
- Burnett, Margaret, Anicia Peters, Charles Hill, and Noha Elarief. 2016. “Finding Gender-Inclusiveness Software Issues with GenderMag: A Field Investigation.” In .
<https://doi.org/10.1145/2858036.2858274>.
- Clarke, Rachel, Peter Wright, Madeline Balaam, and John McCarthy. 2013. “Digital Portraits: Photo-Sharing After Domestic Violence.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2517–2526. CHI ’13. New York, NY, USA: ACM.
<https://doi.org/10.1145/2470654.2481348>.
- Criado-Perez, Caroline. 2019. “The Deadly Truth about a World Built for Men – from Stab Vests to Car Crashes.” *The Guardian*, February 23, 2019, sec. Life and style.
<https://www.theguardian.com/lifeandstyle/2019/feb/23/truth-world-built-for-men-car-crashes>.
- Dervin, Fred. 2015. “Discourses of Othering.” In *The International Encyclopedia of Language and Social Interaction*, 1–9. American Cancer Society.
<https://doi.org/10.1002/9781118611463.wbielsi027>.
- DeVries, Terrance, Ishan Misra, Changhan Wang, and Laurens van der Maaten. 2019. “Does Object Recognition Work for Everyone?” *ArXiv:1906.02659 [Cs]*, June.
<http://arxiv.org/abs/1906.02659>.
- Fisher, Allan, and Jane Margolis. 2002. “Unlocking the Clubhouse: The Carnegie Mellon Experience.” *SIGCSE Bull.* 34 (2): 79–83. <https://doi.org/10.1145/543812.543836>.

- Ghattas, Dan Christian. 2019. "Protecting Intersex People in Europe: A Toolkit for Law and Policy Makers." OII Europe. May 13, 2019. <https://oiieurope.org/protecting-intersex-people-in-europe-a-toolkit-for-law-and-policy-makers/>.
- Grabher, Angelika, Martin Unger, and Sarah Zaussinger. 2014. "Sonderauswertung Der Studierenden-Sozialerhebung 2011 Für Die Technische Universität Wien." <http://www.equi.at/en/projects/focus/8/Sonderauswertung+der+Studierenden-Sozialerhebung+2011+f%FCr+die+Technische+Universit%E4t+Wien>.
- Haimson, Oliver L., Jed R. Brubaker, and Gillian R. Hayes. 2014. "DDF Seeks Same: Sexual Health-Related Language in Online Personal Ads for Men Who Have Sex with Men." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1615–1624. CHI '14. New York, NY, USA: ACM. <https://doi.org/10.1145/2556288.2557077>.
- Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (3): 575–99. <https://doi.org/10.2307/3178066>.
- Harding, Sandra. 2015. *Objectivity and Diversity: Another Logic of Scientific Research*. Chicago: University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/O/bo19804521.html>.
- Hough, Lory. 2015. "Beyond Average." Harvard Graduate School of Education. August 2015. <https://www.gse.harvard.edu/news/ed/15/08/beyond-average>.
- James DiSalvo, Betsy, Sarita Yardi, Mark Guzdial, Tom McKlin, Charles Meadows, Kenneth Perry, and Amy Bruckman. 2011. "African American Men Constructing Computing Identity." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2967–2970. CHI '11. New York, NY, USA: ACM. <https://doi.org/10.1145/1978942.1979381>.
- Karuei, Idin, Karon E. MacLean, Zoltan Foley-Fisher, Russell MacKenzie, Sebastian Koch, and Mohamed El-Zohairy. 2011. "Detecting Vibrations Across the Body in Mobile Contexts." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3267–3276. CHI '11. New York, NY, USA: ACM. <https://doi.org/10.1145/1978942.1979426>.
- Keyes, Os. 2018. "The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition." *Proc. ACM Hum.-Comput. Interact.* 2 (CSCW): 88:1–88:22. <https://doi.org/10.1145/3274357>.
- Kirakowski, Jurek, and Kent L Norman. 2017. "Introduction: Human-Computer Interaction Yesterday, Today, and Tomorrow." In *The Wiley Handbook of Human Computer Interaction*, 1–6. Wiley-Blackwell.
- Lisdonk, Jantine van. 2014. "Living with Intersex / DSD - SCP English." The Hague: Netherlands Institute for Social Research | scp. https://www.scp.nl/english/Publications/Publications_by_year/Publications_2014/Living_with_intersex_DSD.
- Mars, Roman. n.d. "Invisible Women." 99% Invisible. Accessed August 2, 2019. <https://99percentinvisible.org/episode/invisible-women/>.
- Otterbacher, Jahna. 2015. "Crowdsourcing Stereotypes: Linguistic Bias in Metadata Generated via GWAP." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1955–1964. CHI '15. New York, NY, USA: ACM. <https://doi.org/10.1145/2702123.2702151>.
- Paoletti, Jo B. 1987. "Clothing and Gender in America: Children's Fashions, 1890-1920." *Signs: Journal of Women in Culture and Society* 13 (1): 136–43. <https://doi.org/10.1086/494390>.
- Pruitt, John, and Jonathan Grudin. 2003. "Personas: Practice and Theory." In *Proceedings of the 2003 Conference on Designing for User Experiences*, 1–15. DUX '03. New York, NY, USA: ACM. <https://doi.org/10.1145/997078.997089>.

- Publication Manual of the American Psychological Association, Sixth Edition*. 2012. VandenBos, G. R. (Ed). Vol. 6. Washington, DC: American Psychological Association.
<https://apastyle.apa.org/manual/index>.
- Schlesinger, Ari, W. Keith Edwards, and Rebecca E. Grinter. 2017. "Intersectional HCI: Engaging Identity Through Gender, Race, and Class." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5412–5427. CHI '17. New York, NY, USA: ACM.
<https://doi.org/10.1145/3025453.3025766>.
- Shackel, Brian. 1997. "Human-Computer Interaction—Whence and Whither?" *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE* 48 (11): 970–986.
- Spiel, Katta, Os Keyes, and Pinar Barlas. 2019. "Patching Gender: Non-Binary Utopias in HCI." In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, alt05:1–alt05:11. CHI EA '19. New York, NY, USA: ACM.
<https://doi.org/10.1145/3290607.3310425>.
- Stimpson, George William. 1930. *Popular Questions Answered*. G. Sully, Incorporated.
- Tatman, ~ Rachael. 2016. "Google's Speech Recognition Has a Gender Bias." July 12, 2016.
<https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/>.
- Turkle, Sherry. 1986. "Computational Reticence: Why Women Fear the Intimate Machine." In *Technology and Women's Voices*, Pergamon Press, 41–61. New York, NY, USA: Cheris Kramarae (ed.).
https://www.academia.edu/3129915/Computational_reticence_Why_women_fear_the_intimate_machine.
- UNFE. 2015. "Fact Sheet Intersex." In . United Nations Office of the High Commissioner for Human Rights. <https://www.unfe.org/wp-content/uploads/2017/05/UNFE-Intersex.pdf>.
- Wachter-Boettcher, Sara. 2017. *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*. 1 edition. New York, NY: W. W. Norton & Company.