



“But Where Would I Even Start?”

Development of a (Gender) Sensitivity Knowledge Base and Discussion Starter for HCI Research and Practice

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Media and Human-Centered Computing

by

Sabrina Martina Burtscher, BSc

Registration Number 01125029

to the Faculty of Informatics
at the TU Wien

Advisor: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Peter Purgathofer

Assistance: Projektass.(FWF) Dr.techn. Katta Spiel, BSc B.A. MSc
DI Matthias Steinböck

Vienna, 2nd February, 2022

Sabrina Martina Burtscher

Peter Purgathofer



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Erklärung zur Verfassung der Arbeit

Sabrina Martina Burtscher, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 2. Februar 2022

Sabrina Martina Burtscher



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acknowledgements

It takes a village to complete a thesis.

My thanks go to my advisor, Peter Purgathofer and my co-advisors, Katta Spiel and Matthias Steinböck.

Many thanks also to Kristina Weinberger, Matthias Fassel, Alexander Steiner, Lukas Daniel Klausner, Katharina Brunnmayr and everyone else for their help and support through proof reading, providing feedback, listening and just being there whenever I had questions and problems.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

Förderorganisationen verlangen in Calls von Forscher*innen¹ zunehmend, dass sie die Gender-Dimensionen ihrer Arbeit thematisieren sollen – aufgrund des hohen Spezialisierungsgrades, Lücken in der Ausbildung, und einem Mangel an Einsteiger*innen-freundlichem Material stehen Forscher*innen dabei jedoch häufig vor einem Problem. In dieser Arbeit präsentiere ich einen möglichen Lösungsansatz: ein Set an Empfehlungen, wie die Gender-Dimensionen in einem Projekt entdeckt, und das Projekt inklusiver gestaltet werden kann, verfügbar gemacht in Form eines Kartendecks, das Diskussionen innerhalb des Projektteams anregen soll. Um zu diesen Empfehlungen zu kommen, wurde eine vergleichende, tiefgehende Literaturrecherche durchgeführt, die von der Leitfrage “Was können wir hiervon lernen?” geführt wurde. Die Empfehlungen wurden bereits initial mit Interessierten getestet und können auf Basis deren Feedbacks weiterentwickelt werden. Basierend auf den Inhalten von zehn qualitativ hochwertigen Publikationen werden Leser*innen dieser Arbeit und Anwender*innen der Empfehlungen dazu ermächtigt, ihre eigenen bisherigen Arbeitsweisen kritisch hinterfragen und damit ihre (Forschungs-)Arbeit verbessern. Die vorgeschlagenen Empfehlungen können in allen Lebenszyklen eines Projektes angewendet werden, und ermöglichen eine kritische Auseinandersetzung in verschiedenen Gruppensettings, als auch alleine. Im Feld der Human Computer Interaction wird mit dieser Arbeit eine Wissensbasis geschaffen, die in verschiedenen Karrierestufen – vom frühen Studium bis zur etablierten Forschung und Lehre – verwendet werden kann. Die kritisch-fundierte Auseinandersetzung mit existierenden Arbeitsmethoden ermöglicht es, wissenschaftliche Präzision auszubauen.

¹Der hier verwendete Gender-Stern soll darauf hinweisen, dass es mehr als zwei Geschlechter gibt, und diese auch inkludieren. Ausgesprochen wird der Gender-Stern mit einem sogenannten Glottisschlag, wie er z.B. auch im Wort Spiegellei vorkommt.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

Increasingly, funding organisations require researchers to discuss the gender dimensions of their work when handing in proposals. However, because of a high degree of specialisation, gaps in education, and a lack of beginner-friendly materials, such requests pose a problem to many researchers. In this thesis, I present a possible solution: a set of recommendations to discover a project’s gender dimensions, and to make it more inclusive, in the form of a card deck aimed to spark discussions within project teams. To arrive at said recommendations, a contrasting literature review of a small set of publications was carried out, led by the overarching question, “What can we learn from this?”. The recommendations have already been put to test with a small group, and can be improved upon thanks to the group’s feedback. Based on the contents of ten high-quality publications, this work empowers readers and users of the recommendations to critically question their modes of operation and thus improve upon their work. The proposed recommendations can be applied in all phases of a project’s life cycle, and enable a critical reflection in various group settings as well as on one’s own. In the field of Human Computer Interaction, this work provides a knowledge base fit for use in different career stages – from undergraduate studies to established research and teaching. The critical engagement with existing methods of research allows to increase and improve scientific rigor.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Kurzfassung	vii
Abstract	ix
1 Introduction	1
1.1 Research Questions	1
1.2 Motivation	1
1.3 Content Note for Corpus	3
2 Gender, Identity, and HCI	5
2.1 Gender and Identity	5
2.2 Human Computer Interaction and Gender	6
3 Methodology	13
3.1 Corpus Formation	13
3.2 Analysis	14
3.3 Development of Recommendations	15
4 Results	17
4.1 Corpus Findings	17
4.2 Recommendations derived from the Corpus	33
4.3 The Card Deck	41
5 Critical Reflection	45
5.1 Own Pre-Existing Knowledge and Assumptions	45
5.2 Limitations and Constraints	45
5.3 Lessons Learned	46
5.4 On Attribution	46
5.5 On Recursion	47
6 Discussion, Conclusion, and Future Work	49
List of Figures	51
List of Tables	51
Bibliography	53

Introduction

In this chapter, I present the research questions guiding the work on my thesis, the motivation underlying it, and discuss the relevance of the final work for the area of Human Computer Interaction specifically, and Computer Science in general. Finally, I provide a short Content Note regarding the in-depth discussion of the Corpus, as some works deal with trauma-related topics.

The rest of this thesis is structured as follows: Background and Theories in Chapter 2; Methodology in Chapter 3; Description of the Corpus, Recommendations, and Card Deck in Chapter 4. Critical Reflection in Chapter 5. Discussion, Conclusion and Future Work in Chapter 6.

1.1 Research Questions

The effect gender roles and stereotypes can have on projects varies greatly, so there is no “one size fits all” approach to make a project gender-inclusive. This can be seen in projects spanning various areas, ranging from architecture to industrial design (Bardzell, 2010). The goal of this work is to present both researchers and funding organizations with recommendations where to look for gender issues, and how to discuss and handle them. To achieve this, I addressed the following research questions:

- Which gender issues can be found in HCI and how can we reflect on them in our research?
- How can HCI research be made gender inclusive?
- What should a set of recommendations to make HCI research more gender inclusive include in order to help researchers learn about and implement said recommendations?
- In which settings can said set of recommendations spark discussions among team members so that they more easily find ways to make their work more inclusive?

1.2 Motivation

Men are still over-represented in the STEM fields (Science, Technology, Engineering, and Mathematics): While approximately half of the PhD students in Europe are female, only 21% of PhD graduates in computing and 25% of PhD graduates in engineering are women (Ratzer, 2018). The numbers for employees in engineering are similar, with women making up 16.6% of the engineering work force in Europe in 2007 (Erdmann and Schumann, 2010)¹.

¹Such reports currently do not include genders outside the binary (see Section 2.1).

Considering how important technologies are in our societies, it is vital to have adequate participation of all genders in STEM. The technologies developed and deployed by both researchers and engineers in STEM fields have long-term impacts on how people live and work. Thus, discussing gender issues in research and innovation is important to make our societies more just and equal.

My work on this thesis was motivated by a number of factors. Specifically, I had observed that (a) funding bodies increasingly require researchers to describe their projects' gender dimension(s), while at the same time (b) there often is a lack of knowledge regarding gender issues among non-gender studies researchers. Additionally, I wanted to (c) contribute to the development of the HCI community, and (d) make more obvious to readers the politics of gender (research) in HCI. Finally, I had witnessed others' and have my own (e) personal experiences of HCI artifacts that did not account for gender dimensions, resulting in weird and sometimes downright bad user experiences. Each of these I will explore in some depth in the following.

- (a) As stated above, **funding bodies** increasingly require researchers to describe their projects' gender dimension(s), and to position their proposals with regards to equity concerns, often particularly focused on gender.² As academic funding is increasingly driven by third-party money (Wiener et al., 2020), this also relates to research areas not primarily concerned with gender.
- (b) While there is an ever-increasing amount of work published on topics of gender (Stumpf et al., 2020), there is **no well-known low-level knowledge base** providing researchers in specialised fields with concise and actionable guidance in that area. A query for “gender” in paper abstracts within the Guide to Computing Literature of the Digital Library of the Association for Computing Machinery (ACM DL) matches almost 7500 papers.³ Such an amount of potential papers can be in itself overwhelming for researchers unfamiliar with the topic, and cause them struggle to identify *relevant* works, instead relying on “common knowledge”, i.e. stereotypes.
- (c) As demonstrated by the large number of matches in the ACM DL (see Footnote 3), Human-Computer Interaction (HCI) research has developed increased sensitivity towards marginalised populations as seen in e.g., Duarte et al. (2018); Strohmayer et al. (2019); Davis and Farmer (2016); Spiel et al. (2020). This part of the **HCI community** is making great inclusive efforts (see, for example, related discussions in the context of machine learning and artificial intelligence (Leavy, 2018)), particularly when technological artefacts and algorithms operate on scale and permeate infrastructures (Spiel et al., 2019b). As Bellini et al. (2018, p. 8) put it, the – disjoint – communities currently working on gender issues in HCI must “move beyond isolated incidents of feminism towards a community of practice”. This can be done, amongst others, by making visible existing work, spreading their findings, approaches, and methods and helping readers develop sensitivity as to how to explore the relevance of equity within their work. This is where I want to contribute to the community and its efforts.
- (d) Not all items that can be found e.g. in the ACM DL discuss gender critically. Often, “gender” seems to be a characteristic researchers like to draw on in order to find significant differences in their data, even though it is often done as a pure add-on to their actual research (Hines 2004; Maccoby and Jacklin 1974; in Fine et al. (2019)). With more and more digital infrastructures quite literally encode gender e.g. in facial recognition

²For example, the European Commission required H2020 applicants to address gender: <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/support/faq/977>, accessed May 31, 2021; also saved on archive.today

³7 380 matches, queried on April 17, 2020: <https://dl.acm.org/action/doSearch?fillQuickSearch=false&expand=all&field1=Abstract&text1=gender>

technologies (Keyes, 2018), but do not discuss the implications, thus obscuring the **politics of HCI research**.

- (e) Listing **Personal Experiences** could probably fill multiple volumes (see e.g. Wachter-Boettcher (2017) and Criado Perez (2019b)). 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, with watchOS 6⁴. Reiley (2016) reports that Mattel’s “Hello Barbie” at first was 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 (2019a) 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). This is no issue of Google alone: Microsoft and HP software for facial recognition had difficulties recognizing People of Colour, too, and Apple’s FaceID had issues telling Asian women apart. 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; and the worst performance has been reported for Black women (Buolamwini and Gebru, 2018).

As stated above, I hypothesised a lack of knowledge base for researchers who are only starting to look into gender dimensions of their specific fields. While working on Burtscher and Spiel (2020), we could not identify work aiming specifically at researchers to *develop* sensitivity towards the topic, either. Hence, this work offers a starting point allowing researchers to probe their work for relevant matters of (gender) equity. Providing the results from our research not only in the form of academic publications, but also as a publicly available set of discussion starter cards, we aim to make intersectional⁵ sensitivity easier to include in everyday research and development practice. Burtscher and Spiel (2020), my first attempt at an academic publication, seems to have hit a nerve, too: it was accepted into *Mensch und Computer* 2020, and there received a Best Full Paper Honorary Mention Award.

1.3 Content Note for Corpus

Some of the works in the corpus deal with trauma-related topics. In order for all readers to be able to make an informed decision on whether they can or want to deal with the respective contents when looking closer into the case studies, I provide the following content notes:

Clarke et al. (2013) and Ahmed et al. (2014) deal with issues of gendered violence, specifically domestic violence and sexualized harassment in public spaces. Ahmed et al. (2014) especially contains quotes from people describing harassment against women.

Haimson et al. (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.

⁴<https://support.apple.com/en-us/HT210407>, accessed June 14, 2021; also saved on archive.org

⁵i.e., taking into consideration multi-dimensional oppressive factors having specific effects on marginalised experiences such as class, race, sexuality, dis/ability etc. (Crenshaw, 1991; Schlesinger et al., 2017)



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Gender, Identity, and HCI

In this chapter, I give an overview of Gender Theories and discourse around them, present a short history and overview of Human Computer Interaction, and describe how both interact in different kinds of HCI research contribution types.

2.1 Gender and Identity

In order to talk about how gender research is done, and how it could or should be done, it is important to first discuss what “gender” even means, and to introduce some basic vocabulary. This is especially important for the target group of researchers and practitioners who have until now not or only superficially concerned themselves with different conceptualisations of gender and how it permeates societies.

Many researchers operate on notions of *essentialism* or *performance* – both biologically deterministic views assuming that genitalia correspond with sex, and implicate gender. *Essentialism* assumes that there is a strictly binary, immutable, complementary and fixed set of genitalia meaning that there are only two kinds of genitalia, that whichever subset a body has cannot be changed, and that a body can only have one subset of genitalia. The set and its subsets are considered “given”, or natural, and hence can not be changed or challenged. Said genitalia, in the perspective of essentialism, correspond and align with gender, learned behaviour, and societal roles (Bohan, 2016). The concept of *performance* rejects the direct dependence of gender on genitalia, distinguishing between sex (assigned at birth) and gender, where gender is regarded as a reiterated performance that is mostly aligned with the assigned sex. As the sexed body (the body which has been assigned a sex) still is the foundation for the performed gender, any performance that does not align with the assigned sex is deemed deviant (Nicholson, 1994).

In The Wiley Handbook of Human Computer Interaction, Breslin and Wadhwa (2017, p. 74) present a distinction adhering to the performance concept, describing sex as “[a] person’s biological classification as male or female, including physical appearance, chromosomes, hormones, reproductive organs, and secondary sex characteristics.” In contrast, gender in their usage refers to “[a] person’s self-identification and expression in relation to what is considered socially masculine, feminine, or other gendered category, within a given context.”

More recently, Gender Studies theorists have started to understand both gender and sex as social constructs not at all confined to a binary (Fausto-Sterling, 2000). Additionally, it is not possible to clearly distinguish between a presumed ‘biological’ sex and the ‘performed’ gender, as both are dependent on continuous social practices and customs, e.g. ascribing sex to certain body parts (referred to by Fausto-Sterling (2000) as “sexing” of the body), or ascribing gender to certain societal roles. When a person, instead of being externally assigned such – in many societies pervasive and fundamental – categories of personality, consciously assumes a

gendered *identity*, this represents an exercise of self-determination (Hunt, 2018). In this thesis, and the connected works such as Burtcher and Spiel (2020, 2021), my view is informed by an understanding of gender as self-determined and independent of genital makeup.

The fact that sex and gender may not be confined to a binary is also recognized in various countries and societies, as well as their respective legal systems. Some have adapted recently, some have been aware for a long time already (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. In Austria, a total of six options are available: female, male, inter, divers, open, and the possibility to have no entry at all. However, these are all explicitly limited to people who can provide an expertise stating a bodily “variant of genital development” (Humer, 2020).

However, people can hardly be accurately described and defined by only one trait or part of their identity, and inequality does not operate exclusively based on gender. This is where the concept of *intersectionality* comes in: different modes of discrimination might overlap and even amplify each other, meaning that a *white* woman and a Black woman will experience different kinds of oppression, as described by Crenshaw (1991):

Feminist efforts to politicize experiences of women and antiracist efforts to politicize experiences of people of color have frequently proceeded as though the issues and experiences they each detail occur on mutually exclusive terrains.

If gender and other factors are not made explicit, and their possible influence is not considered, researchers risk operating within *unmarked norms* of, e.g., masculinity, whiteness, ableness and (often upper) middle-class (Brekhus, 1998). As Schlesinger et al. (2017) put it, the impact of identity on context and design in HCI must be a focus of intersectional analysis.

2.2 Human Computer Interaction and Gender

Human Computer Interaction studies the interfaces and interactions 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 which factors to evaluate, and different methods of evaluation), as well as the study of end users as both giving input, and receiving output. For the second part of the Wiley Handbook of Human Computer Interaction, Kirakowski and Norman pose the question of the effects technologies have on “us humans”, studying the interface itself, and the interactions taking place (Kirakowski and Norman, 2017). Shackel (1997) also lists issues of accessibility and “special needs”, social aspects of users and communities as important parts of HCI – even 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, and non-reflected way. 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).

2.2.1 Design Pitfalls

Users’ needs and requirements are rather specific depending on the context, and their adapting to and of systems are transient and hard to foresee. When trying to account for these unknown or fuzzy needs and requirements, designers and researchers routinely get snared by pitfalls described by Breslin and Wadhwa (2017): I-methodology, “One Size Fits All” approaches, gender

stereotyping, and gender difference. Falling for these generalisations happens because of the pervasive nature of the (gender) norms incorporated in our societies and technologies.

In the following, I describe the pitfalls and illustrate them with examples of design issues. The explanation of these pitfalls tends to use examples with only binary genders, which is due to the fact that if gender is given consideration, this mostly happens in a binary way (as discussed also in Schlesinger et al. (2017)).

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), or derive ideas for products from their own experience. A common adjacent is the “hallway testing method”, where designers/developers ask people available by walking down a hallway (e.g. to an office next door, or maybe one on another floor) to test an artefact.

While these approaches are not per se bad, they introduce bias into the pools of ideas and test subjects which are often not discussed or accounted for. According to data quoted in Breslin and Wadhwa (2017, p. 72), 80–90% of position holders in tech are men, and 92–94% are Whites and Asians at, for example, Google, Facebook and Twitter. This bias is not an issue only seen in Silicon Valley or the USA, but can be found in the field of computer sciences in general. At Vienna’s technical university TU Wien, my alma mater and Austria’s largest technical university, the numbers are similar: about 85% of students in computer science are men¹ who are likely to share a certain set of experiences, for example having studied at a HTL² (Grabher et al., 2014).

These sets of experiences often 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 be used to track one’s intake of many nutrients, but did not offer the functionality of tracking menstrual cycles³. Similar bias in training data sets can impact the basic functionality of software up to complete dysfunction (DeVries et al., 2019). And while low-pitched voices can operate voice controlled software without any issue, people with higher pitched voices (often women) have been reported struggling to do so in various settings (Criado Perez, 2019a; Tatman, 2016).

Drawing from their own needs, experiences and requirements, many ideas and products developed by the well-situated white men in IT center and serve privileged lives. Ride-share services, food delivery and cooking services, laundry services with pick-up and delivery – many successful companies of the share economy could be dubbed “technologies replacing [the developers’] mums”⁴.

One Size Fits All: “Everyone is the same”

From the above-mentioned fact that the work forces in IT and HCI are very homogeneous derives the method of “one size fits all”, or “works for me” thinking. 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. If you are supposed to interact with something, it should recognize you - but often, cameras

¹While since the beginning of 2019 it is possible in Austria to use gender markers other than “male” or “female” in one’s documents, they are as of October 11, 2020 not displayed either in TU Wien’s numbers (publicly available at https://tiss.tuwien.ac.at/statistik/public_lehre?locale=en), or Statistics Austria’s data (available at https://statistik.at/web_en/statistics/PeopleSociety/education/universities/index.html)

²College for Higher Vocational Education with a focus on technical subjects.

³<https://www.cnet.com/how-to/apple-watchos-6-cycle-tracking/>, accessed October 11, 2020; also saved on archive.org

⁴<https://www.businessinsider.com/san-francisco-tech-startups-replacing-mom-2015-5>, accessed October 11, 2020; also saved on archive.org

and optical sensors cannot recognize People of Color, resulting in soap dispensers not giving out soap to Black people, or proctoring software more likely to flag students of color as cheating on their tests⁵. Even protective gear, no matter the area of application, is often modeled to accommodate men’s bodies – thereby leaving women unprotected, occasionally even hindering them (Criado Perez, 2019a).

In software, 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, big 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, 2019a).

Gender Stereotyping: “All women are alike”

The pitfalls mentioned above are cases of implicit gendering, where people just did not think about whether, or how, gender has impacts on their work. Gender stereotyping, however, is explicit gendering – here, people did realize that there might be differences in how women and men might use whatever is being designed or marketed. However, reflection about why women buy and use certain things, compared to which products are bought by men, did not necessarily take place. Making explicit assumptions that informed decisions and also the processes of how decisions were reached offers a way to improve design and research. However, explicit “design for women” can easily result in “Shrink it and pink it” (SNP), or other ways of negative 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 hands, 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.

Design approaches like SNP cater to stereotypes and reproduce them instead of making technology more inclusive. The underlying stereotypes also make it hard for men to use “girly” designs. Furthermore, they disregard the fact of inter-group diversity, like 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 works the other way around, making products “more manly” in order to make “unmanly” products more appealing to men. Examples for this include, but are not limited to, chocolate⁶, and hot sauce⁷. Hot sauce in Central Europe and the USA often has a masculine connotation, with labels depicting symbols of death and destruction (using skulls, fire, and the likes) 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 “manly BBQ” in the 1950s⁸.

A less obvious, often-used example of gender stereotyping in computer science is prompting someone to “explain something so your grandmother can understand it”. This prompt is often used in job interviews, or oral exams, to see if the interviewee really understood a concept. This

⁵<https://www.technologyreview.com/2020/08/07/1006132/software-algorithms-proctoring-online-tests-ai-ethics/>, accessed October 11, 2020; also saved on archive.org

⁶“Because normal chocolate is not manly enough.”, showing a picture of a box of various chocolate snacks in black wrappers branded “Männersache” (“men’s business”), <https://twitter.com/freakingmuse/status/731057719403270145>, accessed October 11, 2020; also saved on archive.org

⁷<https://www.youtube.com/watch?v=jitjTtWTHYQ>, accessed October 11, 2020

⁸<https://www.vox.com/the-goods/2019/6/28/18760073/barbecue-grilling-men-stereotype>, accessed October 11, 2020; also saved on archive.org

completely ignores the possibility of the testee’s grandmother being an avid user of the concept in question⁹.

To avoid stereotyping, many institutions use personas, which are abstract representations of users. Using personas is supposed to help developers/researchers identify and understand their target audience (Pruitt and Grudin, 2003). A persona 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 various user groups, when applied properly. For example, good personas would show that the usage of a washing machine differs depending on household size, or if the washing machine was placed in a students’ dorm. However, creating good personas needs time, information, and insight (Marsden et al., 2017). If not enough time is put into creating the personas, one runs the risk of relying too heavily on stereotypes, thus introducing into the project what was tried to be avoided in the first place.

Gender difference: “Women are inherently different”

There are some things where men and women may be 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 looks (Breslin and Wadhwa, 2017, p. 73, citing Ceci et al., 2009). For example, while pink and blue are nowadays rather strictly assigned to girls and boys respectively, the assignment was less strictly handled until the first half of the 20th century (Stimpson, 1930; Paoletti, 1987).

Issues when looking at gender difference are: binary division, mashing up sex and gender, and the alleged immutability of gender (Keyes, 2018; Schlesinger et al., 2017; Breslin and Wadhwa, 2017). *Binary division* has been discussed in section 2.1. *Mashing up sex and gender* is easily explained: one speaks about “gender differences”, but actually discusses a characteristic like body hair distribution, and uses gender categories as proxy variables. *Immutability of gender* describes the assumption that gender is something fixed – once assigned, this assignment cannot be changed. Most of the time, gender assignment happens without consent, and from without (think of doctors proclaiming a baby’s gender based on their visible genitals right after birth). Immutability of gender disregards the fact that a person’s gender fluctuate, that sometimes it may challenge gender norms, and sometimes adhere to them, depending on context, place, or time. Also, a person may very well feel content within their assigned gender for some time, but later in life feel the need 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. One does not make her less of a woman, the other does not make her less of a hacker.

2.2.2 Gender in current HCI research

Looking at the increase of publications addressing the influence of gendered dimensions on HCI, in Burtscher and Spiel (2020) we analyse some of the activity regarding gender recommendations in the forms of *analytical reviews*, *software analysis* and *guidelines*. For each of these groups we looked at some recent works and how these conceptualise issues of gender and equity, describing how Burtscher and Spiel (2020) and this thesis fit in with the bigger picture.

Analytical Reviews

Roig-Maimó and Mas-Sansó show how researchers need to be wary of generalisations based on user studies, if they focus solely on mathematical parameters of significance, but do not account

⁹Examples in this 2019 twitter thread include a mother with a PhD in Materials Science, mothers who indeed *do* use the concept in question every day, and so on: <https://twitter.com/danrkports/status/1108083064104587264>, accessed October 11, 2020; also saved on archive.org

for the distribution of gender among their participants (Roig-Maimó and Mas-Sansó, 2019). The authors suggest researchers should discuss the limitations of their research due to their participant samples. Using a binary view on gender, this work fits in with most traditional HCI research.

Stumpf et al. discuss gender-inclusive HCI research focusing specifically on cognitive and behavioural works and the dangers of stereotyping in technology design Stumpf et al. (2020). Even though the authors explicitly attend to issues surrounding a dominant binary notion of gender, their approach risks essentialising gender (cf. Keyes et al. (2020)) and does not operate from a notion of self-identification.

In Burtscher and Spiel (2020), we take a prior literature review by Schlesinger et al. as a starting point due to its intersectional approach Schlesinger et al. (2017). The authors determined how much of existing research is narrowly focused on single markers of identity without attending to overlapping and amplified modes of oppression. They investigated how HCI research identifies and/or classifies target populations and research participants within a corpus of 140 papers published between 1982-2016 at the ACM CHI Conference on Human Factors in Computing Systems (CHI). Particularly, they focused on diversity and intersecting categories of identification along gender, ethnicity, race, class, and sexuality (though leaving out, e.g., age, localisation and disability).

Software Analysis

Burnett et al. identified gender as a relevant factor in software (Burnett et al., 2016a), even if the software itself was not explicitly concerned with gender (González-González et al., 2018). The authors and further collaborators created an application that systematically steps through source code to identify instances in which gender bias occurs (Vorvoreanu et al., 2019). The method relies on gender-inclusive personas Marsden et al. (2017) in order to avoid stereotyping (Hill et al., 2017). More recently, some of the authors were also involved in an attempt at generalising the method to other issues of equity (Mendez et al., 2019). This approach requires not only that researchers are already aware of and have a basic understanding of gender (or other) identity factors appropriately. Furthermore, it assumes that these issues can be feasibly addressed by a cognitive walk-through. While this method comprise a first step in aiming at (gender) equity in software, it also presents a formulaic approach to an issue that requires situated sensitivity.

Burtscher and Spiel (2020) and this thesis present a potential starting point for researchers to develop knowledge relevant to apply such methods critically and effectively.

Existing Guidelines

In the context of qualitative settings, Rode calls for more reflexivity (also on gender) in HCI research (Rode, 2011), and Brulé and Spiel developed considerations on how to negotiate gender (and disability) between researchers and participants in participatory design (Brulé and Spiel, 2019).

Jaroszewski et al. investigated how in two rather different contexts, namely tumblr users and fantasy football players, the same form to inquire into participants' gender would be met (Jaroszewski et al., 2018). The free form field was used by the first group to provide a range of unique genders and to praise the approach, whereas the second group was hostile towards the freedom of expression. Based on these insights, the authors derived a generalised approach for how to inquire into gender, which has been further refined, including a call for situated nuance (Spiel et al., 2019a). The *HCI Gender Guidelines* (Scheurman et al., 2020) stem from a grassroots initiative of HCI researchers and provide hands-on guidance for gender equity in writing and research as well as the organisation of academic events.

Burtscher and Spiel (2020) and this thesis supplement this approach based on lived experience and activist knowledge by analysing existing works and drawing higher-level recommendations from peer-reviewed publications as different modes of assessing these issues lead to stronger recommendations.

Additionally, who we are, and how we are seen, also defines which resources we have access to, resulting in different points of view on the world we live in (Haraway, 1988; Harding, 2015).

Operating from a biologically deterministic perspective might seem like an acceptable approach when researching body differences. However, it is important to make explicit which differences are researched, instead of using blanket terms like “male” and “female”. Defaulting to these categories, and assuming shared patterns e.g. of fat distribution, actually makes nuanced discussions of outcomes difficult. Making explicit what was included in the research/design, and how, actually enables critical engagement, feedback and improvements on one’s work.

In HCI/Computer Science, bodily differences can relate to, for example, work on fitness or 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 more active metabolisms, or 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, societal norms code certain tasks as “women’s work”, which impacts the number of men completing them, as well as the tasks’ overall image and importance. Looking at gender differences means to scrutinize 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 snow plowing pattern. One town in Sweden changed their snow plowing 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 which 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 as well as on the city’s health cost, 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 with the choice of study field or hobbies, and continuance in CS careers, respectively. They look at what image women and men have of computers in general, or of CS studies, where these images come from, and how it impacts who enters a field, and who stays. Other work focuses on how software can implement different features to be better usable for women, for example to help them steer clear of harassment on social networks.

Initially, gender is assigned to people at birth, based on their visible genitals; throughout their life, this assignment is repeatedly reinforced (by clothing a child according to norms; by offering a child certain toys, but not others; by reprimanding it for not adhering to its ascribed role). Time and again, this external attribution does not match a person’s 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.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Methodology

In this chapter I describe and explain the methods I used to collect, analyse and interpret the data used in this thesis. At points, I am going to use “we” rather than “I”, as part of this work has been conducted together with Katta Spiel while preparing our joint publication (Burtscher and Spiel, 2020). For a detailed discussion of this aspect, see Section 5.4.

3.1 Corpus Formation

For the GEECCO literature review, I started out with some research on the history of HCI to get a feeling for the field as a whole, and to find out at what point in time gender topics came into view in the field. As soon as I started looking for specific works on HCI and gender issues, I realised that compiling a comprehensive overview would by far go beyond the scope of a Master’s thesis. In light of this, I decided to look for more specific research, and arrived at Schlesinger et al. (2017), an analysis of how researchers in ACM Special Interest Group on Computer-Human Interaction’s (CHI) conference publications construct “the users” they are researching/working with.

For their report, Schlesinger et al. looked into 140 papers presented at CHI, mapping how in those papers “the user” is constructed. They conclude 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 a binary construction barely are discussed at all.

From Schlesinger et al.’s corpus, I selected seven papers covering a range of topics, and aiming for geographical diversity as it relates to authors’ institutions and research environments. Together, these papers show different ways of studying, constructing, and considering gender in research. While working on Burtscher and Spiel (2020), we added three papers published since 2017 to account for more recent developments in the field. To locate and summarise the knowledge and information we were looking for, we conducted a close reading of the selected papers that in our opinion could function as an introduction into how the field thinks about gender. An overview of the corpus (Burtscher and Spiel, 2020) is shown in Table 3.1 on Page 14.

Paper	population	locale	context	treatment
Blackwell et al. (2016)	LGBT parents	USA	social media	self-id
Karuei et al. (2011)	general	CA/DE	wearables	binary/biology
Otterbacher (2015)	general	CY	crowdsourcing	binary/external
Clarke et al. (2013)	DA survivors	UK	photo-sharing	binary
Ahmed et al. (2014)	general	BD	online platform	binary light
Haimson et al. (2014)	MSM	USA	dating ads	ambiguous
James DiSalvo et al. (2011)	black men	USA	game development	ambiguous
Keyes (2018)*	trans people	USA	gender recognition	self-id
Metaxa-Kakavouli et al. (2018)*	general	USA	web interfaces	self-id
Fernandez and Birnholtz (2019)*	trans people	USA	dating platforms	self-id

Table 3.1: Characteristics of papers part of our close reading corpus. Papers indicated with * have been added outside of Schlesinger et al.’s corpus. *DA* stands for “domestic abuse”, *MSM* stands for “men who have sex with men”, and *self-id* stands for “self-identification”. Table adapted from Burtscher and Spiel (2020).

3.2 Analysis

After the initial research phase for the literature review and some deliberation with Dr. Ratzler, I decided to not create a systematic, cursory review covering a great number of papers. Instead, I would read, compare, and discuss a smaller set of papers (see 3.1), which would give me the opportunity to do so in much greater detail. This approach was also kept up during work on Burtscher and Spiel (2020). Drawing on Martin (2005), *close reading* was applied as a method for both the initial literature review and the subsequent paper. This enabled us to analyse the corpus in detail, looking at different approaches and underlying motivations. We analysed the papers according to the implications they could have for HCI research and knowledge production. In short, we conducted a *contrasting review* of a deliberately non-exhaustive list of papers in order to find difference, diversity and variety.

Each work within the corpus was analysed with regard to implicit and explicit gender aspects in research questions, methodologies, and language used. This analysis was driven by the question, “What can we learn from this?”, so we could later derive a set of recommendations. This guiding question was chosen to ensure the derived recommendations would remain relevant also for researchers and practitioners who are not yet familiar with the topic. You will find said question as a recurring sub-section in Chapter 4. In line with current funding requirements, and due to the focus of GEECCO, I initially focused on gender, but soon expanded my lens to include intersectional characteristics (Crenshaw (1991); Schlesinger et al. (2017)). This broadening of scope was also kept up for Burtscher and Spiel (2020), and we hope to spark HCI researchers’ curiosity on what else – aside from gender – they might have to consider to adequately address equity and justice in their research.

I do not believe in objectivity Harding (2015), and we are very well aware that this work is *interpretive*. However, we argue that rigour can be established within a notion of *partial perspective* (Haraway, 1988). This means that we do not claim generalisability of our findings, but rather present an informed analysis on which we base a set of situated recommendations. Our analysis is influenced by our marginalised positions as a woman and a non-binary person within gendered power dynamics (Collins, 1997). At the same time, being white and culturally tied to a majority perspective within central Europe, our analysis remains outside of a lived experience of, e.g., race (Ogbonnaya-Ogburu et al., 2020) or forced migration (Talhok et al., 2018). During the time I was working on the literature review, the paper, and the card deck, I came to terms

with being neurodivergent¹. Making these positional markers visible allows readers to reflect on the potential limitations of our analysis and contrast their position. Katta Spiel and I also want to invite and encourage further work from perspectives different to ours (Hartsock, 2017).

3.3 Development of Recommendations

As described in Section 1.2, I hypothesised from the beginning that researchers would benefit most not from a simple summary on gender research in HCI, or a checklist, but would rather need some sort of assistance – like a set of recommendations to follow. While working on the GEECCO literature review, during the analysis of the initial corpus, I already picked up on some recurring themes. These I developed into rather general recommendations, which also influenced, in a manner of speaking “on the fly”, how I proceeded with the literature review (and everything else I have been doing since). I detail this recursive aspect of my work in Chapter 6.

During work on Burtscher and Spiel (2020), we found some more patterns that could be refined and made more concrete (and thus, more practical). It was Katta Spiel’s idea to fit the recommendations into the temporal structure of a research project (Burtscher and Spiel, 2020). These temporal “boxes” are only one way to see connections, similarities, and possible synergy effects in between different recommendations. While preparing a workshop where I would use the resulting (digital) card deck, I derived a set of categories that turned out to work well for my participants (see Section 4.3.2 for details). The most important aspect while working with the recommendations is to provide the space necessary for thorough discussion and reflection – the categories, as well as the recommendations, mostly function as catalysts.

In the next chapter, I describe the corpus papers and detail the findings of the close readings. In Chapter 6, I then present the recommendations we derived from the close readings, and the development of the card deck.

¹Meaning that some of my brain functions do not work like the majority’s, i.e. I have a type of brain functioning that is not neurotypical (see also <https://www.merriam-webster.com/dictionary/neurodiversity>, accessed June 18, 2021; also saved on archive.org)



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Results

This chapter presents the corpus papers, as well as some relevant examples from outside of academia, and for each answers the question “What can we learn from this?” (Section 4.1). This is followed by the description of the recommendations, grouped by our temporal classification (Section 4.2, page 33ff). Finally, I present the process of developing the card deck, from the initial decision on the format to the go-live of the online version (Section 4.3, page 41 ff.).

4.1 Corpus Findings

In this section, I discuss each of the corpus papers in detail to illustrate the close readings that took place during work on the GEECCO literature review and on Burtscher and Spiel (2020). I decided to fully list all authors of each paper when I first introduce it. This way, I want to acknowledge all authors’ contributions, and hope to improve visibility of otherwise marginalised authors. Additionally, I aim 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. For each paper, I present a short description of the research and, if applicable, the associated project(s) and their connection to HCI, followed by the gender aspects included in the research. The *best practices* we found are put in italics. A summary of these best practices and specific recommendations targeted at researchers and practitioners is presented in section 4.2.

4.1.1 Complexities of Disclosure in Blackwell et al., 2016

For this paper, Lindsay Blackwell, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck looked into how LGBT¹ parents use social media. LGBT parents here references people who are parents and LGBT, as opposed to being parents of LGBT children. The Human-Computer Interaction connection lies in the focus on usage of social media: which platforms 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)? While the term “gender” is not in the papers’ keywords, it is visible already in the title that gender is an important aspect of the research.

Blackwell et al. group their findings into three primary themes:

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.” (p. 614)

¹The acronym stands for Lesbian, Gay, Bisexual and Trans. For the purpose of describing the paper, we use this term to stay close to the text, even though more inclusive abbreviations are available, e.g. LGBTQ2IA* (Lesbian, Gay, Bisexual, Trans, Queer, Two-Spirit, Intersex and Asexual with * indicating incompleteness of these identities). For further information, see, e.g., the glossary in Brant (2016).

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 (and, e.g., partnership status) 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." (p. 615) Some participants feel that their mere existence and visibility are acts of advocacy and resistance in times of anti-LGBT politics (here: 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." (p. 616) 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, it may have 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. provide detailed information on who their study participants are (Blackwell et al., 2016, p. 613). In addition to sexual orientation and gender identity, the authors also discuss location (rural vs. urban), and socioeconomic status as *dimensions of intersectionality* (it may be easier to be out & proud in big cities as opposed to rural areas), further enriching their descriptions. They suggest that some of their work may also be applied to LGBT persons who are not parents, but that most is specific to LGBT parents. They acknowledge the in-group differences within LGBT communities, for example pointing 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. try to show what kind of influence which parts of their participants' identity 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' life and/on social media.

The authors *pay special attention to the vocabulary* used throughout the paper. When talking about gender, sexual orientation, cis/trans, or chosen names vs. "legal names", they explain what special terms mean. By explaining the term cis instead of trans, they avoid presenting cis as the norm. Additionally, Blackwell et al. not only use sensitive language, but also stress the importance of respecting participants' choice of *self-identification*. Interviewers reported paying specific attention to the language used by each individual participant regarding their identity, so the interviewer could adapt their own language for any further communication. Rather than using pre-established vocabulary and categories, participants' self-descriptions and preferred expressions were used (p. 614). The authors recruited their participants *through peer groups* on various social media platforms (e.g., Twitter, Facebook). Even though the focus was on social media use, the researchers did not seek out their participants' online profiles out of respect to their individual *privacy* and to dignify the participant's desired presentation towards the researchers.

Apart from their methodological modulations, we can take from this work that researchers should communicate a *non-judgemental safe environment* to participants, and strive to create a space in which it is comfortable to share personal experiences as well as to *refrain from probing* deeper into areas and topics participants do not readily disclose.

4.1.2 Diversity of Embodiments in Karuei et al., 2011

Idin Karuei, Karon E. MacLean, Zoltan Foley-Fisher, Russell MacKenzie, Sebastian Koch, and Mohamed El-Zohairy examined body locations for tactile displays. Testing for sensitivity toward vibrations and impairments by movement, impact of visual workload, expectation of location, and gender, the authors tested performance of tactile displays. They also investigated if users had subjective preferences to any of these conditions. The HCI connection lies in 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.

What can we learn from this?

The team balanced gender (male/female) in their experiments, as they consider it an indicator for, e.g. differences in body fat distribution (p. 3270). This is an example for a biologically essentialist understanding of gender. And while the authors indeed report differences in the success of signal detection and response time when analysing by participant's gender, they point out themselves that these results are neither consistent nor significant.

Overall, Karuei et al. did try to make their research gender aware, but stopped just shy of discussing some fundamentally 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, or skirts) designed for women often do not have pockets big enough for smartphones, if they have any at all (Burman, 2002). Another point that could have been included in their work, is the fact that presumably, all participants were able-bodied. This is something we have to assume, as there is no mention of disability in their description of norms and baselines: “Feet are the baseline for sites, male for gender, sitting for movement, no workload for workload, and first trial for trial number.” (p. 3271).

With regards to methodology, the authors made sure their participants wore the *same type of clothes* (sportswear) to reduce effects of different clothing styles. They do not disclose how they inquired into gender, and only report it for eight participants (male). From their results, we can take that gender does not transfer well to a specific concept of embodiment. Instead, *body characteristics should be explicitly and directly analysed* rather than inferred from identity characteristics.

4.1.3 Development of Gender Bias in Otterbacher, 2015

Jahna Otterbacher looked into the results of *ESP*², a Game With A Purpose (GWAP). ESP was created by Von Ahn and Dabbish in order to crowdsource textual metadata for images. To this end, users are randomly paired and challenged to agree on as many words to describe an image (labels) as possible within 2.5 minutes. For every word the players agree on, they earn points. After successfully completing 15 images, the players receive a bonus. In order to create specific labels instead of general ones, the game will show a list of taboo words, which are selected based on their popularity within the broader player community. 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”. Thus, players contribute to the accumulation of big database structures. At the same time, they might introduce personal and collective biases into these systems and the algorithms learning from them.

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 decision making systems. Otterbacher states that there is already a noteworthy body of research on how tasks may be gamified, and on the incentives 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 between 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*.³
- The proportion of subjective adjective labels that have *positive/negative prior polarity*.
- The most frequent strongly subjective adjectives.

What can we learn from this?

Otterbacher’s results provide evidence that players rather describe *how* they perceive the women to be (labeling them with adjectives) while describing *what* the men are (using nouns, e.g. occupations). Looking at the strongly subjective adjectives most frequently used to describe either men or women, Otterbacher finds that all reference “physical appearance (for example, sexy, ugly, cute) or disposition or character (for example, angry, happy, fun)” (p. 1960). In detail, Otterbacher found impressive gender differences: of the top 10 subjective adjectives used to describe men, only two specify their appearance. In contrast, from the top 10 subjective adjectives used to describe women, 5 relate to their appearance. Overall, 2,425 images of women (16.6%) were labeled using the most frequent, subjective label of “sexy”. The same label was used to describe only 20 pictures (0.1%) 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 to be 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).

²The name refers to the abbreviation for ‘extrasensory perception’ as an indicator for the distributed collaborative nature of the game (Von Ahn and Dabbish, 2005).

³Subjective adjectives, as the name suggests, are adjectives used to express one’s subjective belief or opinion about something.

Otterbacher also includes some critical thoughts about the images included in the game and the analysis. Considering that they were collected from online resources and selected randomly, the images probably represent a wide variety of depictions of people. Yet, 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 than “normal” women available, while for men this ratio might be more even. In order to find out if the gender biases discovered in their first analysis hold true for less biased sets of images, Otterbacher also analyzed images in contexts where labels related to occupations had been assigned (p. 1961). One result of this in-depth analysis is that players have distinct biases regarding doctors and gender: half of the images displaying women doctors had been labeled “nurse”, while the same label never occurred on depictions of men. This indicates a stereotype that women are most likely assumed to be nurses, while men are more likely deemed to be doctors (p. 1962).

The first thing, and most easy to do, we can learn from this work concerns the pronouns Otterbacher uses when talking about people of unknown gender: the author 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 [...]”.

In addition, Otterbacher explicitly discusses that there are more than two genders, and more sexual orientations than hetero- and homosexuality. The work also presents us with explicit 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. The author designs a second analysis in order to test the results of the first.

Considering methodology, researchers can take from this work to be cautious about *pre-existing bias* in online pictures of people, as these are shaped by gendered societal expectations. Otterbacher also reflects on the *limits of assigning gender* to pictures (and people), especially if gender is reduced to binary labels. When testing outcomes for women against results for men, and finding differences, this work reminds us that *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. Finally, from Otterbacher’s results, researchers can take note on how to ensure equivalence in describing people regardless of gender, in order to *consistently articulate the same information* when writing about participants, case studies, personas or example stories by using the same type of words (adjectives, nouns etc.) and classifiers (e.g., jobs, colours etc.).

4.1.4 Agency around Trauma in Clarke et al., 2013

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 such a life disruption 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 evading their abuser(s). The technologies used in Clarke et al.'s work were not developed from scratch - rather, the focus was on new ways of using photographs and storytelling as a way for the survivors to (re-)gain agency over their narrative. The project took place over the span of several months and design iterations. The gendered aspect of this research was domestic violence, which is often framed as a problem within a binary concept, where women are seen as survivors and men as perpetrators (Bellini et al., 2019). Additionally, Clarke et al. talk about the digital gender gap, meaning that many women are less proficient and confident using various kinds of technologies.

While other facets of identity and their influence are briefly touched upon, Clarke et al. exclude them as research foci, reducing domestic violence to an issue between men and women. However, researchers call for an intersectional approach toward the topic (Ristock and Timbang, 2005). For example, class (Anderberg et al., 2016) and culture (Clarke et al., 2012) have specific impacts on the power dynamics at play. Disabled people are reported to be at a higher risk of being abused (Cramer and Plummer, 2010) and LGBTQ2IA* people often have to face additional hurdles in order to receive support (Calton et al., 2015).

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, . . .) as well as photography itself incorporate acts of storytelling (by choosing what to show, and how) which in turn helps constructing identity and building relationships. The researchers aimed to foster an open and long-term engagement to work on (individual) collections and displays of images in various (non-)physical states: printed, displayed on screens, incorporated into videos.

Clarke et al. designed what is usually called a “cultural probe” 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 to assist the participants in photographing or recording the aspects of their lives they wanted to share and/or retain. 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 build up confidence, experience a sense of achievement and affirm their agency.

The team conducted ten sessions of two hours per week between November and February, with breaks for Christmas and school holidays. Initially, they planned to videotape the sessions, but the participants were uncomfortable with that idea. Instead, the process was documented using anonymized field notes for each session, and interviews with the women's center staff, an outreach worker and the center coordinator, and a recorded group discussion (p. 2521).

What can we learn from this?

The first, and probably most obvious lesson to learn from this project is: Remember that *there are no “standard users”*. Clarke et al. use gender as an analytical focus to restrict the population of their research, and acknowledge that women who have experienced and survived domestic violence have specific needs which not all women have; on the other hand, people of other genders with similar experiences around domestic violence might share these needs. Life disruptions, and a, hopefully good, life thereafter, exist (Clarke et al., 2013; Wachter-Boettcher,

2017). Designs and research touching on such life disruptions have to be handled carefully. Assumptions about for example stable family or personal life must be thoroughly examined (p. 2525).

In order to handle life disruptions with the appropriate sensitivity, Clarke et al. *worked with topical experts*, in this case survivors of domestic violence and staff at the women's center. They examined the vocabulary they would use (p. 2519f), as well as the scope (p. 2520), and ways of documenting the project (p. 2521). One of the authors volunteered at the women's center for several months before the project's sessions started in order to get in touch with future participants, and to inform the research questions (p. 2519). This way, Clarke et al. found a way of *building careful relationships* with their participants. Not holding workshops during times schools were closed was a way of respecting the women's life outside of the project and *attending to participants' needs and availability*, as many had care duties for young dependants.

As seen in other case studies, and as recommended by Schlesinger et al., Clarke et al. provide readers with rich details on their participants, and how they got involved with the project: “[...] a self-selected group of 6 South Asian women who were aged between 25-38 and had been separated or divorced for between 1 - 6 years [...]” (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 et al., 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 in between the researchers and the research participants, who are the actual experts on their own situation. Clarke et al. *disclosed their perspective* and motivations on the research topic not only in their documentation, but also during their workshop sessions in order to create a trustworthy setting for their participants.

Finally, Clarke et al. also *reflect on the power imbalances and ethical challenges* coming with 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)

4.1.5 Attending to Everyday Violence in Ahmed et al., 2014

The Protibadi (a Bangla word meaning “one who protests”) project 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 project’s focus thus was on the gendered issue of sexualized harassment and violence and its impacts.

The authors describe how they applied user-centered design in order to create 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 stories around sexual harassment in public places launched in August 2013. Three months later, the website had 110 registered users (20 self-identified as men, the rest as women). The users had shared 24 reports of sexual harassment from different parts of Dhaka city⁴. Three months after launch, the team conducted user studies and analysed public responses to Protibadi in order to evaluate the strengths and limitations of the system.

Ahmed et al. reference several similar projects spanning the world, as well as localized projects, for example from Egypt, whose insights and design approaches they could build on. The participants of their focus group discussion reported strong feelings of shame, sadness, and regret as well as defiance, anger, and a strong desire for change. In semi-structured interviews the women talked about their understanding and experiences of sexual harassment as well as their requirements and needs for design targeted at this issue. 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).

Participants who had shared their experiences with others before (for example with family members and sometimes friends), reported 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 following the launch was positive in general. Several interviewees noted the strong need for a system like it, 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 in other papers discussed in this thesis (e.g. Blackwell et al. (2016); James DiSalvo et al. (2011); Fernandez and Birnholtz (2019)), Ahmed et al. give *detailed insight into who their participants were, and why they chose to select them*. Their decision to focus on university students was based on three factors: first, their access to technology (mobile phones and internet) would enable them to actually use the system; second, the authors assumed they would be more attuned to issues of gender discrimination and sexual harassment, due to their education and socioeconomic position; third, they argue that 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). More *privileged participants* also might feel safer sharing experiences that are associated with feelings of shame.

⁴A 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 was last active in 2016.

Discussing these *socio-economic factors*, Ahmed et al. critically reflect on *how their focus creates limitations* to the generalizability of their work. The authors discuss further dimensions of intersectionality e.g. in the discussion of future steps. They mention an NGO planning to adapt Protibadi for rural areas with less mobile coverage, as well as a group wanting to extend it towards harassment in the workplace. Comparing Protibadi to other similar projects around the world, Ahmed et al. point out the need to *consider localized differences*, like 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.

In their group discussion and interviews, Ahmed et al. used various ways to try and *make their participants feel comfortable* when talking about sexual harassment. The authors for example provided the participants the option of writing things down, and having the written statements read aloud by the woman conducting the discussion. 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 the participants agreed. Furthermore, the participants were reassured they could stop at any point of the process, request their data be destroyed, and walk out. It seems that the research team was successful in communicating will to put the participants at ease, as they report one participant requesting that her data should not be shared with male members of the research team, and another interviewee bringing a companion who helped her share her story.

Finally, Ahmed et al. *acknowledge the high prevalence of sexual harassment* and violence and their gendered impact. Understanding what damaging effects repeated violations can have on some participants might help contextualise, e.g., the work of Karuei et al. (2011) discussed above.

4.1.6 Language on Sexual Health in a Haimson et al., 2014

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 understand how users (re)present themselves, and how they communicate. Haimson et al. 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, the authors studied the temporal changes in sexual health related (SHR) language, 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 could 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 language 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 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?

Haimson et al. *use exact and concise language* to describe their target group (“men who have sex with men”, or MSM for short), with the intent of including e.g. bisexual men, who would be excluded or made invisible when using either “homosexual men” or “gay men”. Also, they present an instance of *self-disclosure and discuss the team’s positionality* when describing the coding team: “[f]our coders, including two gay-identified men” (p. 1618), similar to Blackwell et al. (2016).

However, Haimson et al. do not explicitly discuss what model of gender they base their work on. As they only talk about ‘men’ without further specification, they implicitly rely on the dominant binary model.⁵ In order to avoid confusion, researchers should briefly *reflect on their inclusion and exclusion criteria* even when their focus lies on one gender only, for a better understanding of which populations they refer to.

The authors *discuss how and why groups were constructed* regarding the age statements in the ads, taking into account the specific history of HIV/AIDS (p. 1620) rather than using conventional age clusters. Categories are defined for example based on the user’s phase of life (and thus the fact whether or not they probably were sexually active at that time) during the 1980s, when most people first learned of HIV/AIDS. Another group consists of those too young to remember the HIV/AIDS epidemic, who may display more risky behaviors than older men (p. 1620).

⁵Given other work involving Haimson (e.g., Haimson (2018); Starks et al. (2019); Saha et al. (2019); Haimson et al. (2019)), this might well be an instance of inadvertently adhering to a binary model due to ambiguity.

4.1.7 Intersectional Identities in James DiSalvo et al., 2011

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. The authors report that in general, although African American⁶ 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” or “nerd”: 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 and learn about computing. Participants received an apprenticeship as game testers working full-time during the summer, and part-time during the school year. Their tasks encompassed quality assurance for pre-release games of various game companies.

To measure effects of the program, pre- and post-test surveys were conducted with 21 Glitch participants. Students joining the program in 2009 were categorized in the study as “oldtimers”, those joining in 2010 as “newcomers”. The survey measured participants’ perceptions of technical competency for both their peers and themselves, and the access to technical resources in their social groups. In addition, the team conducted interviews and focus groups, and researchers spent over 800 hours finding, observing, and engaging with participants.

In the survey, participants were asked to list close friends and 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 resources) they were asked to report relationship, closeness, and technical expertise of each connection. James 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.

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, e.g. that of oldtimers over newcomers, may have positive impacts on technology adoption and identification.

What can we learn from this?

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

The authors also state that when recruiting participants, they did not explicitly include gender and race in their selection criteria. Rather, they had been looking for students interested in video games who stemmed from low-income households (measured by Free and Reduced Lunch eligibility). All applicants who met these criteria happened to be young African American men.

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 (dis-identification).

⁶As before, stay with the term used by the authors, although more recent works seem to prefer the labels ‘Black/African Descent’ Ogbonnaya-Ogburu et al. (2020).

They explicitly take into account how gendered stereotypes masculinity are influenced, and even amplified by culture and racialisation (Crenshaw, 1991).

Combining this *knowledge of identity construction and intersectionality* of their participants' lives, the authors found an appropriate way to make computing more interesting for them. Often, projects to pave the way into computing for newcomers try to make it “fun” and “easy”, using elements of gamification. With Glitch Game Testers, participants had the opportunity to see computing as a future occupation, and as something they are able learn and do.

Together with more recent work, e.g., Ogbonnaya-Ogburu et al. (2020), this research challenges readers to reflect on the *unmarked norm of Whiteness* in computing, and shows how many professions in technology and especially computing inaccessible not just for (white) women. Thus, it shows that when thinking about equity and gender, researchers need to *critically interrogate their viewpoints* and standpoints Hartsock (2017) in order to be able to appropriately contextualise both the knowledge they produce and its limits.

4.1.8 Automating Stereotypes in Keyes, 2018

In their extensive meta analysis, Os Keyes explored how in the fields of automated gender recognition (AGR) and HCI the term “gender” is understood. They argue that the “automatic computational identification of a person’s gender from photographs or videos” (Keyes, 2018, p. 88:4) is inherently trans-exclusive, as it presumes that a person’s gender could be recognized, and thus assigned, externally.

From the top seven pattern recognition publication venues, Keyes analysed 58 papers published between 1995 and 2017. On the side of HCI, they selected 12 papers from the International AAAI Conference on Web and Social Media (ICWSM), the Conference on Hypertext and Social Media and SIGCHI conferences and journals that used an implementation of AGR as part of their methodology. In both categories, Keyes looked for implicit or explicit references to the traditional view of gender (binary, immutable, physiological). HCI papers were additionally probed for discussion of the AGR technology’s model of gender, or limitations thereof, and if any “efforts to actively test whether the model matched reality” were undertaken (or rather, documented).

Keyes “found a remarkably consistent operationalisation of gender” in the papers they sampled (Keyes, 2018, p. 88:7). In 94.8% of papers gender was either implicitly or explicitly treated as binary, in 72.4% as immutable, and in 60.3% as physiological. In the group of papers with a focus on gender, references to a binary and/or immutable view were just a little less numerous, but references to a physiological model were twice as frequent than in the papers without a focus on gender. The small set of works that did not refer to traditional views on gender were “largely those which did not discuss their model of gender, or essentialised [...] elements of external appearance other than physiology.” (Keyes, 2018, p. 88:7) Keyes argues that such findings likely hold for racialised contexts as well (Keyes, 2018), rendering approaches of “recognizing” individual’s traits or characteristics as inherently risky of amplifying pre-existing inequalities.

What can we learn from this?

In a very straightforward way, Keyes recommends to generally *avoid automated identity assignment*. Their suggestion that researchers should not (try to) externally define a person’s self-representation and identity, or aspects thereof, can also be expanded to non-automated assignment, i.e. assuming participants’ gender based on their looks and behaviour. Keyes also suggests to *frame gender explicitly and trans-inclusively* and calls for *appropriate resources for researchers*. Since the publication of their paper, some such resources have been created, e.g. the HCI Gender Guidelines (Scheuerman et al., 2020) or Spiel et al.’s recommendations for survey design (Spiel et al., 2019a). The thesis at hand is an attempt to add to this corpus of resources.

4.1.9 Interfacing in Metaxa-Kakavouli et al., 2018

The layout of a physical space, or its decoration, has been shown to have an impact on visitors' feelings of ambient belonging (Cheryan et al., 2009). Metaxa-Kakavouli et al. hypothesised that the visual design of a webpage may have a similar impact on users' sense of belonging. To test this hypothesis, they created two webpages representing an introductory course in Computer Science. The pages showed the same content (from a Stanford University course), but differed in design and layout. One layout featured a Star Trek reference as background image, and green-on-black monospace design (referencing a stereotypical hacker's command line interface), while the other showed an image of the sun shining through a leafy tree, and a white-on-green sans-serif text layout. The former is dubbed the '(stereotypically) masculine' layout throughout the paper, while the latter is referred to as 'gender-neutral' (Metaxa-Kakavouli et al., 2018). In a pre-test, the authors recruited Amazon MTurk workers similar in age to their target population and asked them to rate the designs on scales of 1 to 7 representing (1) 'not at all' to 'very stereotypical of a computer scientist' (2) 'very feminine' to 'very masculine' (Metaxa-Kakavouli et al., 2018).

In their experimental setup, participants were presented with either the 'masculine' or the 'gender-neutral' webpage. After looking at the stimulus webpages, participants answered questions on "their interest in the course, sense of belonging, anticipated success in the course, self-confidence in their computing skills, interest in studying computer science in the long-term, and anxiety about how others in the course would perceive their gender." (Metaxa-Kakavouli et al., 2018, p. 614:2). The participants also rated the webpages on the same scales as the MTurk workers (see above). Metaxa-Kakavouli et al. report that men did not distinguish much between the designs, as opposed to women (Metaxa-Kakavouli et al., 2018). These findings hold both with regard to perceived masculinity of the design and intention to enroll in the course. For both variables, the 'neutral' design was assessed similarly by men and women, but enrollment intention dropped significantly for women looking at the stereotypically masculine webpage, and they rated it as significantly more masculine, even 'juvenile' (Metaxa-Kakavouli et al., 2018, p. 614:4). Metaxa-Kakavouli et al. state that women's answers indicated lower sense of belonging in the context of stereotypically masculine designs. The authors explicitly mention that no participants reported to be nonbinary, and list this as well as the sample's lack of racial diversity as limitations of their study.

What can we learn from this?

Metaxa-Kakavouli et al. work with a gender model of *self-identification*, with cited related work predominantly using a binary model of gender. In the limitations section, the authors reflect on their sample's lack of diversity both with regard to gender and race. Another limitation they touch upon, but could have discussed just a little more is the sample's large reliance on current students. With 67% of participants currently studying, information on their fields of study is actually important, especially when thinking about answers (and answer options) to questions regarding enrollment in the course, or future choice of study field.

The authors *point out different conceptualisations of gender* across cultures may be connected to gendered differences found in previous work. From their results, researchers can take pointers on how to reflect on their designs of prototypes and software, and how these might be perceived as stereotypical within their target group to *appropriately control for gendered impressions*.

4.1.10 Identity Disclosure in Fernandez and Birnholtz, 2019

In their paper, Fernandez and Birnholtz focus on dating platforms as a topic of social computing research. They conducted one-on-one interviews with 20 trans dating platform users in the United States of America to discuss forms of and reasons for disclosure of ‘trans status’ (Fernandez and Birnholtz, 2019, p. 226:1).

While everyone has to think about which information to disclose where and how, trans people, and particularly trans women of colour (Fernandez and Birnholtz, 2019, p. 226:6) can rarely be certain whether it is safe to disclose their gender in online contexts. Disclosing their gender to the wrong person may lead to slurs, harassment, hate speech and/or violence by other (cis) persons, which is why trans people have developed strategies that allow them to assess the safety of a given situation. In their interviews, Fernandez and Birnholtz identified two forms of disclosure: direct (clear and open) and indirect (with the option of plausible deniability) (Fernandez and Birnholtz, 2019). Fernandez and Birnholtz report that participants would often combine various forms of disclosure, and that different modes of disclosure would pursue different goals. They found that safety and certainty were key motivators for disclosure. According to Fernandez and Birnholtz, concerns for one’s physical safety lead to users disclosing their gender within the platform rather than face-to-face, as they considered receiving a negative reaction on the app less risky than in person (Fernandez and Birnholtz, 2019, p. 226:9). Certainty, as a motivation, refers to users’ desire to avoid rejection, and to make sure (to a certain degree) that a developing relationship would have a positive outlook.

What can we learn from this?

Fernandez and Birnholtz recruited using several *community-based platforms* (e.g., an LGBTQI2A* community centre, trans-specific mailing lists) and targeted Facebook advertisements. While they are not sure about the reasons, they do discuss that this strategy led to a sample with a bias towards White trans people. *Participants’ preferences* were taken into account for the *interview setting* (video chat or in person). Additionally, during each interview Fernandez and Birnholtz explicitly *asked about the participant’s pronouns* and made clear these would be used in publications. The authors also reflect on their status as cis researchers and *identify themselves as allies*.

Fernandez and Birnholtz allocate time and space in their paper to discuss their model of gender, and why trans users face different struggles when using online dating platforms. In their analysis, they take into account that rural and urban communities may have different effects on trans individuals’ strategies of disclosure. Finally, they also stress that the decision on how to, if at all, disclose one’s gender is deeply individual.

4.1.11 Outside of Academia

As Human Computer Interaction is not only an academic field, a lot 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 products to fail. For this insight into the industry, I have found examples in the work of Sara Wachter-Boettcher (2017), Caroline Criado Perez (2019a) 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 2019⁷. Reiley (2016) reports that Mattel's "Hello Barbie" at first was not able to understand its prime target group (little girls). Similar issues of voice recognition failing to recognize voices with a relatively high pitch are reported by Criado Perez (2019a) 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 (cis) men's voices. Another case that can be traced back to slanted training and test data was when Google Pictures' algorithm failed to recognise Black people on pictures, tagging them as gorillas instead. This is a symptom of the test and training data sets 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. This is no issue of Google alone: Microsoft and HP software for facial recognition had difficulties recognizing People of Colour, too⁸, and Apple's FaceID was reported to have issues telling Asian women apart.

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 (cis) men; the worst performance has been reported for Black women (Buolamwini and Gebu, 2018). As discussed by Keyes, people outside traditional gender models are also frequently mis-recognized.

Not only software is affected by disregarding gender dimensions, 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 different body shapes are smartphones and other hand-held devices that are too large for small hands (Criado Perez, 2019a), or GoPro chest mounts⁹.

What can we learn from this?

The most important lesson to be learned from the incidents described above is to be critical of the things one uses, be they data sets, parts of software or full applications. Researchers should spend some time trying to find out what the people behind their resources had in mind, and where they might have overlooked something. In order to find stress cases (Wachter-Boettcher, 2017) and handling them well, it is important to think outside the box (and also that of your colleagues'), asking whether things have to be done the way they have been done.

Further, the cases reported here make clear that it is important to consider the situatedness of knowledge (Haraway, 1988). For data sets, whether they are statistical data to mine information from, or survey data from user studies, there should be discussion and reflection on how they came to be, as done .e.g in Otterbacher (2015).

⁷<https://www.cnet.com/how-to/apple-watchos-6-cycle-tracking/>, accessed October 11, 2020; also saved on archive.org

⁸https://www.pcworld.com/article/209708/Is_Microsoft_Kinect_Racist.html, accessed October 10, 2020; also saved on archive.org

⁹<https://twitter.com/HackyScientress/status/1361295141798051845>, accessed March 1, 2021; also saved on archive.org

4.2 Recommendations derived from the Corpus

As described earlier, I developed an initial set of recommendations after completing the literature review for GEECCO. This initial set of recommendations, which has similar content, but is worded in more general terms, can be found in the literature review on Human Computer Interaction and Gender at the GEECCO website¹⁰. The recommendations were improved upon during work on Burtscher and Spiel (2020), resulting in a set of high level recommendations on accounting for gender while *designing research, acquiring funding, conducting* and *presenting* HCI research. After completing work on Burtscher and Spiel (2020), the recommendations were used to create a deck of cards (to be discussed in Section 4.3), which I used during a workshop at ditact summer studies 2020. At said workshop, I collected some feedback on both the recommendations and the cards as a tool.

Overall, the goal of these recommendations is to assist project team members in becoming aware of the gaps in their ideas, data, and plans, and to show how these gaps might be attended to and even mended. Other approaches that might help, e.g. the GenderMag method (Burnett et al., 2016b) or the “gender patches” described in Spiel et al. (2019b) have been discussed in Section 2.2.2.

While a checklist might have been easier to create, and could be the faster and easier way to go for the readers, these recommendations only come in a rather verbose form. This is due to the fact that we wanted to create sustainable change aside from just checking items off a list. Thus, I want to point out here that it does take time to attune ourselves to experiences that are different to our own. As we put it in Burtscher and Spiel (2020, p. 437), we want to “encourage our readers to dig further, read some of the texts presented here [...] in depth and continuously reflect on their practices. Knowledge production attending to power imbalances, oppression and marginalisation comprises a process of ongoing atonement and negotiation.”

In the following descriptions, I use “we” as a denominator pointing to “us”, the community of researchers, designers, and developers of HCI-related content. As discussed in Section 3.3, we grouped the recommendations into temporal “boxes” as one way to show connections, similarities, and possible synergy effects. An overview of said boxes is presented in Table 4.1 on Page 34.

¹⁰http://www.geecco-project.eu/resources_results/geecco_material/, last access March 13, 2021; also saved on archive.org

Project Life Cycle Phase	Designing Research Protocols	Acquiring Funding	Conducting Research	Presenting Research
<p>Phase description: What happens in this phase?</p> <p>Recommendations: Short name/title of the recommendations associated with this phase</p>	<p>defining what you are going to research, how, and including/excluding whom.</p> <ul style="list-style-type: none"> • Articulate Gender Explicitly • Plan for Individual and Collective Reflection on Positionality • Probe Methods for Accessibility and Sensitivity • Deliberately Assess Exclusions 	<p>finding a source of funding for your project, writing your proposal(s).</p> <ul style="list-style-type: none"> • Check Expectations and See How They May Be Extended • Understand Proposal Writing as a Teaching Opportunity 	<p>interacting with your research partners, making sure you did not overlook or ignore important points.</p> <ul style="list-style-type: none"> • Follow Participants' Choices in Identifying Them • Attend to Different Needs and Preferences • Seek Critical Feedback • Actively Look for What's Missing 	<p>documenting your process, your findings, strengths and limitations of your work.</p> <ul style="list-style-type: none"> • Provide Appropriate Context Information • Reflect on Representation and Prior Work • Choose Mindful Language

Table 4.1: Overview of project life cycle phases and how the recommendations are grouped into these phases.

4.2.1 Designing Research Protocols

Research happens not only in academic settings, but also in other project contexts. When developing an artefact, designers and developers also have need to conduct research in order to discover their target user group(s), their needs, preferences, and aversions.

As such, even before funding for a project can be sought, and before work on it can start in earnest, project members have to plan and design their research protocols.

If we reflect on gender sensitivity early on and become aware of the structures and dynamics of power that affect potential participants, target populations or informants, we can create a strong basis for future work.

Articulate Gender Explicitly

In the background section (Section 2.1), we described three different approaches to understanding gender: *essentialist*, *performative* and *identity based* models. As Keyes (2018) suggests, we need to articulate our understanding of gender in advance. This way, we can later on communicate more clearly the implications of the model we rely on. Further, making factors explicit enables us - and our readers, and participants - to reflect on the inferences of gender we might otherwise have made implicitly.

Plan for Individual and Collective Reflection on Positionality

Positionality describes the fact that due to our individual position within groups and societies, we can only see, learn and understand things from a specific point of view. This often leads to presumptions and stepping stones that may influence a given project and its associates. In order to identify these influences, we should plan for individual and collective reflection on positionality and personal perspectives, starting early on. The corresponding plan can be a living document that we adjust to emerging and changing needs and practices. Ongoing reflection can be in the form of shared or private diaries for every individual involved, regular meetings with documentation, or informal check-ins with an accountability structure (Rode, 2011). For different people, different modes might be more suitable. Just as with making our view on gender explicit, thinking about potential ways to capture the project and its process from a personal perspective from the start makes it easier to disclose and discuss positionality in the presentation of research later on (compare Clarke et al. (2012); Fernandez and Birnholtz (2019); Blackwell et al. (2016)).

Probe Methods for Accessibility and Sensitivity

We often build our work on prior concepts and research, without giving a lot of thought about the history of these predecessor technologies. They usually were developed in and for certain contexts, so sometimes adaptations become necessary. This relates to gender insofar as body related work often is tested only on a homogeneous population, most often able-bodied cis men. Reflecting on how (in)accessible the methods we choose are in a specific context allows us to also understand more about the limitations of our work. Within the corpus, Clarke et al. adapted their chosen method after discussing it with experts in order for it to make sense for their participants.

Of course, mishaps can still happen - however, they should be met with an honest interest to do better. Deliberations on gender can be a starting point for discussing identity factors more broadly so we can further develop and sharpen our sensitivity for marginalising tendencies in research and development.

Deliberately Assess Exclusions

Any project will exclude some people from participating and include others. Deliberately assessing in advance who might or will be excluded, and why, can mitigate situations where exclusions might happen unintentionally or even in opposition to a given research question or project goal.

As an example from the corpus, Ahmed et al. acknowledge that for their participants to feel safe and secure enough to talk about topics associated with stigma, they had to focus on a fairly small and privileged part of the population (Ahmed et al., 2014). Another way to look at this topic might be by considering unmarked norms as James DiSalvo et al. did this for Whiteness in computing (James DiSalvo et al., 2011).

4.2.2 Acquiring Funding

In order to fund research activities, researchers need to increasingly engage with funding bodies who award finances competitively. Especially those funding bodies tied to political or governmental entities, e.g., the European Union or national science funds such as the Deutsche Forschungsgesellschaft (DFG), the Austrian Science Fund (FWF) or the Vienna Science and Technology Fund (WWTF) distribute available funds according to various specific criteria. Being themselves funded largely through taxpayer money, these funding bodies also need to attend to the usefulness of research to advance societies to become more just and equal. Hence, they often require researchers to describe their (planned) research work's impacts on gender norms, rules, or stereotypes, if relevant. We argue that this gives researchers the opportunity to not just align with funding bodies' requirements, but to even push them into a more progressive position (while simultaneously presenting their projects as progressive as well) and to disseminate knowledge about more substantial issues regarding equity to reviewers and boards (Burtscher and Spiel, 2020).

Check Expectations and See How They May Be Extended

As touched upon above, funding bodies often publish a request to attend to gender issues when calling for submissions. These requests are often based on stereotypical binary ideas, and/or tied to an essentialist view on gender (see Section 2.1). Researchers could inspect the call and review its limitations in order to create a proposal such that it attends to issues of intersectionality and equity more generally, while still falling under the umbrella of what the funding body expects. Such an understanding and extension allows, e.g., the work on Black men in computing detailed above (James DiSalvo et al., 2011).

Understand Proposal Writing as a Teaching Opportunity

While being domain experts, proposal reviewers might not be entirely familiar with concepts such as intersectionality or gender as a self-determined identity. Thus, such approaches need to be communicated accessibly across disciplines, and proposal writing becomes an opportunity for disseminating knowledge on these topics. In detailing one's assumptions on gender and equity, clarifying specifically why and where this matters within the proposal, researchers may create a best practice. In that regard, being precise in language and acknowledging potential biases in data sets (whether pre-existing or created; cf. Otterbacher (2015)) helps to guide the reviewers' understanding of the subject matter.

4.2.3 Conducting Research

In order to make our research practices more inclusive takes time, requires earnest work, and comes in many different shapes and sizes. However, when we conduct research in an accessible way and adapted to the needs of otherwise potentially marginalised participants, we can better work **with** them (or at least, their input), instead of designing and developing **for** them, based on our assumptions **about** them. The knowledge we uncover and produce in such a way has the potential to be more thorough and more reflected along its limitations.

Follow Participants' Choices In Identifying Them

As illustrated by Fernandez and Birnholtz (2019) and Haimson et al. (2014), it is a basic sign of respect towards our participants to explicitly ask for and also use their pronouns (Fernandez and Birnholtz, 2019), even if we do not directly interacting with them (Haimson et al., 2014). When we value their choice on how they want to self-identify without judgment (Blackwell et al., 2016), we also stay close to the data and information provided to us, and stay honest towards our participants. In order to receive this sensitive and personal information, researchers need to provide participants with the space and freedom to express themselves as they want to. In interviews, this can be managed during dialogue. Smaller scale studies can provide free form fields for disclosing gender. In larger scale studies, it becomes relevant to know the dominant approach to gender within the target population (Jaroszewski et al., 2018) in order to appropriately design questions into gender identity (Spiel et al., 2019a). We also urge researchers to explicitly offer participants an option not to disclose their gender (Spiel et al., 2019a).

Attend to Different Needs and Preferences

While it is understandable that researchers are focused on their research questions and tasks, it helps to be mindful about participants' lives outside of a given project context.

As early as during recruiting, researchers can familiarise themselves with the cultural environment a target population might have. This can be done by, or help in recruiting through peer groups (Blackwell et al., 2016) or community-based platforms (Fernandez and Birnholtz, 2019).

Attending to participants' needs can mean to actively approach and include marginalised groups, and may e.g. take the form of offering lunch during workshop days and tweaking the project schedule to accommodate for care responsibilities (Clarke et al., 2013), as well as making counselling available for participants (Ahmed et al., 2014; Clarke et al., 2013). Additionally, flexibility in the choice of research setting creates an inclusive research environment, as done e.g. by Fernandez and Birnholtz (2019) in conducting interviews based on participant technology preference. To further create a safe and non-judgmental environment, we suggest to refrain from probing into topics where participants do not readily disclose aspects of their lives.

Seek Critical Feedback

Critical feedback may help to find limitations that were previously overlooked, and challenge us to further improve the work done so far. It can be sought in three ways, as demonstrated by Clarke et al. (2013): by building long lasting engagements and careful relationships with participants, so they feel safe in challenging our ideas; by collaborating with topical professionals, so they can bring in their expertise; and by actively disclosing our own perspectives and motivations to participants (Clarke et al., 2013). However, it takes time to appropriately account for where we can do better (Stengers, 2018).

Actively Look for What's Missing

One of the hardest things to do is to be aware of the gaps in our ideas, the data we have collected, and the plans we have developed and deployed, and to then set out and try to mend them. Finding those gaps can be especially difficult if we discuss our work only within our own domain(s). In lieu of, or in addition to obtaining critical feedback from without, researchers can also attune themselves to actively attend to what is missing in their research, akin to hunting software bugs in engineering practices). Sometimes, this serious examination of our work may result in the realisation that, e.g., not using gender as a variable might be more appropriate and a different, more nuanced view on people (Keyes, 2018) or their bodies (Karuei et al., 2011) could be more meaningful to a given research question or context.

4.2.4 Presenting Research

We document our research in different ways, i.e. in writing or talking, and at different points in time, i.e. while in progress, or as a final presentation at the end of the project's cycle. In any case, we need to reflect on and be thoughtful with regard to the implications of our research as it pertains to gender and other identity factors.

Provide Appropriate Context Information

In order for readers and fellow researchers to build on existing work, detailing and explaining our choices is paramount, as it provides necessary and appropriate context. This applies e.g. for describing and explaining our motivation to include and operationalise gender within a given research context (Keyes, 2018). We also suggest to discuss the origin of pre-collected data, if present, and analyse it for incompleteness or skew (Otterbacher, 2015) – for example, workplace environments likely change across large enough time spans, possibly making surveys and their findings regarding the topic volatile. Here, we further recommend to acknowledge the specific locale for knowledge, understanding the limits inherent in this and extending solidarity to diversify knowledge constructed in different geographical contexts (see also, Kumar et al. (2019)). To further expand on the example, workplaces within one geographic region may differ in between industries; workplaces within one industry may differ over regions. Hence, surveys on the topic might not be easily translatable from either one to another.

Disclosing positionality can also be relevant contextual information as done e.g. in Haimson et al. (2014); Fernandez and Birnholtz (2019); Blackwell et al. (2016). As we pointed out above in referencing prior work (Haraway, 1988; Brulé and Spiel, 2019), disclosing identity allows readers to contextualise the viewpoint of given works, which – contrary to dominant practices – remains relevant for post-positivist as well as constructivist approaches. Additionally, if researchers are themselves part of the groups they analyse, this comprises another point of explicit representation (see Section 4.2.4) and provides an additional level of lived expertise with a given subject matter (Keyes, 2018).

Reflect on Representation and Prior Work

Representation can refer to examples, pictures and imagery used. Here, researchers can reflect on unmarked norms embedded in their visualisations and textual examples and diversify where possible. As most prior work operates from a binary, essentialist understanding of gender (Stumpf et al., 2020), denaturalising prior work (i.e. making explicit that the categories or vocabulary used in it is a product of its environment, not “neutral”) can also mean broadening representation and attending to cultural differences where appropriate (see Metaxa-Kakavouli et al. (2018)).

Choose Mindful Language

The use of inclusive and precise language can help to avoid perpetuating exclusionary stereotypes. For example, we suggest to use singular they when referring to “abstract” individuals (as done e.g., by Keyes (2018)) and to state all options and choices when presenting statistics (e.g., “45 % identified as women, 44 % as men, 7 % as nonbinary, 4 % chose not to disclose their gender; participants could select multiple options”). Further we recommend asking for and using participants' self-descriptions and pronouns (Blackwell et al., 2016; Haimson et al., 2014; Fernandez and Birnholtz, 2019), as well as gathering feedback regarding language used by representatives of the target population.

Again, as not all readers and fellow researchers may be familiar with concepts of intersectionality or gender as a self-determined identity, some additional information should be provided (cf. Section 4.2.2).

4.3 The Card Deck

In this section, I describe the process of developing the card deck, including the reasons for choosing the format, making the deck available online, and changes made from the recommendations detailed in Section 4.2.

The idea to create not only a set of recommendations, but a physical set of cards came up while writing the thesis proposal. It is a bit difficult to reconstruct how I arrived at the idea, as many different people were involved in the discussions on the way. For example, it was made clear early on by the Dean of Studies, Hilda Tellioglu, that the literature review for GEECCO would not be enough for a thesis. GEECCO's Brigitte Ratzler and I had already decided to derive a set of recommendations from the literature. In addition, when talking to, amongst others, Brigitte Ratzler, the point was made that a simple list of recommendations might be seen by many as just another checklist to work off. In contrast, we wanted the recommendations to have mid- to long-term effects on the ways researchers work. As a way to achieve this, my primary supervisor, Peter Purgathofer, suggested to create a card deck. One of the benefits we considered a card deck would provide was an additional physical/haptic component that a simple checklist would not offer.

I intended the cards to be used by project teams, or subgroups thereof, to encourage discussion and reflection on the topic of (gender) sensitivity, and make dissemination of corresponding knowledge easier. We theorized that cards would offer high flexibility with regard to group size: a single person could use the deck on their own, and groups could use the deck while sitting together. Group size would only be restricted if the group decided that one unique card should be used per member, which would imply the number of cards in the deck as an upper bound for group size.

The card deck offers users various different points of view to look at their project from, no matter if they use it alone or in a group setting. When used in a group setting, distributing one card per participant, each participant will *have to assume* a different position from which to discuss their project from. Each card is a hand-sized text snippet posing questions, offering information and proposing actions to take, in language close to one's own (non-academic German and English, respectively).

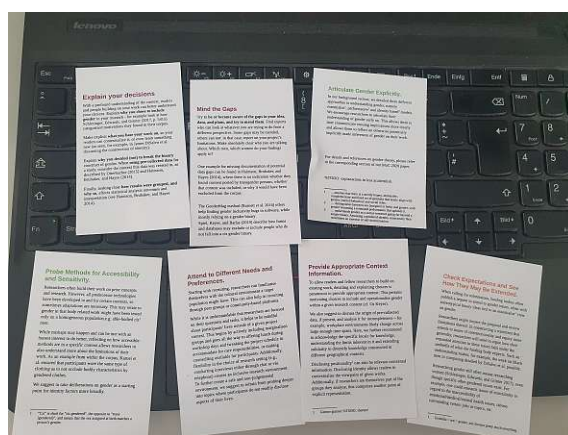
4.3.1 Development of Text and Online Version

The recommendations with their verbose and brief descriptions were used as the text and headers for the first draft of cards. Development of the cards happened using Libre Office Writer, using page format A5 and a split-page setup to come close to the intended card size. The text then was (a) translated to a more vernacular German and (b) cut down so each recommendation would fit on one card (1/2 page in A5 format). I decided to translate to a less academical German because I was already planning a workshop that would be taking place at a German-only venue (ditact.ac.at), and because people at the workshop should be able to use the cards no matter their level of education or their focus of research. This is also an aspect of accessibility and bringing the topic closer to those who until now had (near to) no contact with it. The workshop was planned to be an opportunity to receive feedback from people who before had not heard of the project (as compared to most of my surroundings). Another difference between the original text as presented above, and the card deck is the amount of footnotes. Many cards feature one or several footnotes which mostly cover recurring concepts (cis, trans, intersectionality) in order to establish a baseline for the users. My motivation for this was that not everyone using the cards would know these concepts (even if they heard the terms), and I wanted this crucial information to be available *on each card*. This way, if someone in a group using the cards does not know a term or phrase, they would not have to "out" themselves as not knowing to the group or the person holding onto the description.

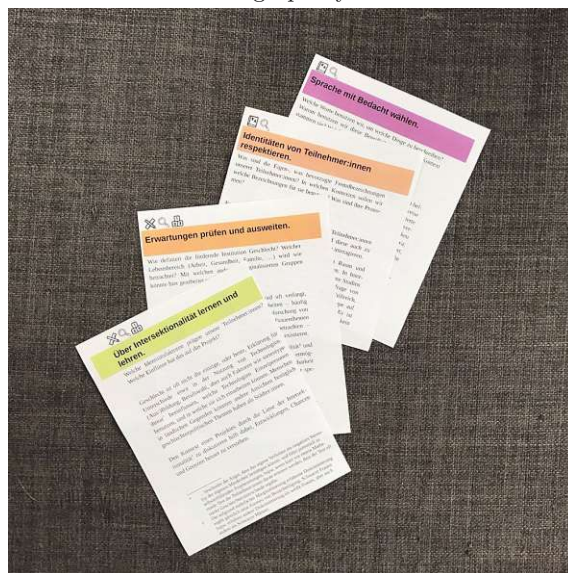
Finally, in addition to the categories we had already established (see Section 4.2), I constructed further categorizations by looking at thematic rather than temporal connections. Basically, I asked the question “How, apart from over the life cycle of a project, could these cards be connected?”. The temporal categories in the initial design were color-coded into the title of each card (see Figure 4.1a), and later in the background of the title (see Figure 4.1b), to improve legibility. The thematic categories are displayed using icons on the top left of each card (see Figure 4.1b).

On the technical side, the base document for the cards is a Libre Office Writer document. I decided to use Writer because I figured it would be easy to edit. The WYSIWYG character of the program immediately made changes like the switch from single-page A6 to double-page A5 visible, and made it easier to compare the two versions.

While working on the card texts, I also printed the work in progress cards to see if the format I chose would be okay for physical use. I also planned to make the PDF file available online from the start, so people could share and use the card deck widely.



(a) Work-in-progress print-outs.
Photograph by me.



(b) Print-outs of the final layout.
Photograph by Katta Spiel.

Figure 4.1: Development of card layout.

While I had planned on having a printable version of the card deck available online, I initially focused on the physical version, in order to test it with groups in Vienna, and at the ditact workshop. When it became clear that the latter would be held remotely, and that (local)

contact restrictions due to CoVid-19 would stay in place, I had to prioritise development of a fully-fledged online version including a “draw a card” feature. The initial setup was basically a proof-of-concept, consisting of screenshots of the PDF generated by Writer. A PHP script would return a random image drawn from the folder containing said screenshots. Many thanks to Matthias Steinböck for the code and helping setting up the web version. In this form, the card deck was not at all accessible for people who use screen readers, and also did not scale on different screen sizes. This was pointed out by Alex Steiner, a friend who works on web and accessibility, who insisted it had to be fixed before the workshop. He helped by creating a script that would convert markdown files to HTML, and incorporating this into Matthias’ PHP setup. This means that the final version makes the cards screen reader accessible, and mobile-friendly. Additionally, each card on the online version is a separate HTML file with a designated URL, so people can actually share links to specific recommendations. The online version can be accessed at <https://div.uber.space>, including some explanations on the project, information on how to contact me or contribute to the card deck, and the printable PDF.

4.3.2 Testing and User Feedback

Initially, I planned to test my card deck with focus groups of research and teaching staff at TU Wien, but after talking about my work in general and the master’s thesis at a FLINT¹¹ in tech meet-up at the 2019 Chaos Communication Congress¹² in Leipzig, a member of the ditact organisational team encouraged me to submit a corresponding proposal. I opted for the workshop format rather than a talk, because I saw the potential of trying out the cards in a setting similar to the one I had in mind, and gathering feedback on it. This would also give me the option to test my *hypothesis* that an interactive format would be a good way to use the cards, and to encourage discussion and reflection on the topic.

As the target group for courses at ditact are both students and people working in IT-related sectors, I decided to keep both the title and the course description practical and pragmatist. The title, “Projekte inklusiv gestalten – Diversität Abseits von Personas” roughly translates to “Inclusive project design – diversity aside from personas”. The (German only) description is available on the event website¹³.

The workshop attracted 5 participants from various backgrounds. For introductions, I asked each participant what their current connection to diversity and inclusion are, and whether they had any current projects where they would like to apply the card deck. We discussed what their experiences with diversity and inclusion measures were in the past, and followed that with an input where I presented and explained the development and content of the card deck, and offered space and time for questions. After that, I asked each participant to draw a card for themselves and (on their own) try to find ways of applying the respective recommendation to their current project (or their job in general). Most participants used the first card the tool presented them with. One person reported looking at multiple cards before starting on the task, using the card they were “handed” initially. Finally, we discussed each card they had drawn, and what their ideas on applying it were.

At the end of the workshop, I asked the participants explicitly for verbal feedback, and announced that I would email them the link to an online feedback form, which in the end was filled in by one person.

¹¹abbreviation for German “Frauen, Lesben, inter, trans”, translating to “women, lesbians, inter, trans”

¹²details on 36C3 are available at <https://events.ccc.de/2019/10/04/36c3-in-leipzig/>, last accessed June 18, 2021; also saved on archive.org

¹³<https://ditact.ac.at/course/1956-kurs-online-projekte-inklusive-gestalten>, last accessed June 18, 2021; also saved on archive.org.

Verbal Feedback

For the verbal feedback, I asked for the participants' opinions on both the workshop and the card deck. I made my own local notes which I hoped would give me pointers for aspects to improve as well as things to keep.

The group setting was well received by the participants: One participant noted that the mixed group (most probably regarding the participants' various fields of work/expertise) was interesting, another pointed out that the group size worked well for the virtual setting. Yet another participant stated that they liked the interactive character of the setting ("guter Austausch"), and that they approved of my sticking to our half-day schedule. One participant remarked that they had hoped for a broader focus of the workshop overall - even though I tried to bring in other dimensions of diversity, we mostly discussed issues of gender.

Regarding the cards, several participants remarked that the content clearly leaned towards research and HCI issues - however, they also pointed out that they could see how to adapt the deck for different fields of work. One participant said that they found the questions on the cards to be well developed. They described the temporal grouping of the cards as "smart and fascinating" (translated from "sehr schlau & spannend"). Two participants explicitly stated that they really liked the idea of using cards.

During the verbal feedback, one participant asked for the complete printable PDF file to be put up on the website so they could print it and use the card deck at work¹⁴. I consider this request to be an instance of positive feedback, as it implies that the participant considered the card deck to be useful, and wanted to share it with their environment.

What really impressed me was that even the participant who during the introduction round gave me seemed very sceptical of the method, and even of the necessity of improving diversity and inclusion, said that they had learned some new things. More specifically, they stated in the feedback round that they would in the future try to scrutinize implicit assumptions more often.

Written Feedback

After the workshop, I emailed the participants some additional information, and the links to the workshop documentation. As announced, I also included the link to a feedback form asking them for more detailed feedback. While only one person filled in the form, they provided me with a lot of detailed feedback to guide the further work on the card deck.

Most importantly, they had very specific feedback for two cards. For #11 "Actively look for what's missing"¹⁵, they recommended to add more examples, and to include other dimensions of diversity than gender. They also noted that it might be useful to include information on the concept of positionality on this card. Their feedback for #1 "Talk About Gender Explicitly"¹⁶ took a similar direction: the card itself "is great, but I wondered why other dimensions don't exist as individual cards"¹⁷. They continued to list some examples like background (German: "Herkunft") and race/racialization (German: "Hautfarbe"), or to change the card to "Explizit über Identität sprechen" ("Talk About Identity Explicitly").

In general, they said they really liked the idea, and would love to see more cards. For the online drawing tool, they mused about an option to draw more than one card.

Asked for specific use cases, the respondent noted planning events (e.g. conferences), as well as scientific projects. Prompted for whether they would advise the card deck to their colleagues, or what would have to be improved on in order for them to do so, they responded they already had told others about the deck.

¹⁴As mentioned before, I had planned on doing so, but did not get to it before the workshop.

¹⁵translated from German: "Aktiv suchen, was fehlt"

¹⁶translated from German: "Explizit über Geschlecht sprechen"

¹⁷translated from German: "passt super, aber ich habe mich nur gefragt, warum andere beispiele [sic] nicht explizit als karten [sic] existieren".

Critical Reflection

In this chapter, I want to reflect what working on this thesis has sparked within me, how it has influenced me and my work – both regarding the thesis, and other topics. I start with a short look into my own pre-existing knowledge and assumptions, describe limitations and constraints, lessons learned, and discuss the topic that was the major cause for the delay in writing up my thesis: attribution. Finally, I want to briefly touch upon the topic of recursion.

5.1 Own Pre-Existing Knowledge and Assumptions

Regarding pre-existing knowledge, I have completed several courses at TU Wien that dealt – either explicitly or implicitly – with gender issues in engineering and computer science. Additionally, I participated in a course on scientific Theories and Methods in Gender Studies, taught by Katta Spiel at University of Vienna. In my voluntary engagements at the student representative body, I worked towards the betterment of the situation of women both in university, on the fringes, and beyond. Thus, I considered myself to already be well-versed in topics of gender equality and social justice.

My most important, and most basic, assumption was that there is a need to be able to understand the gender dimensions of a project. Based upon discussions with researchers from various fields, and in different contexts, I learned that this motivation could be based either intrinsically (“I want to do better”), or extrinsically (e.g. “I have to do this, because the funding organisation requires it”). Finally, from my experience as a student of computer science, I know that for future researchers, there is not enough space to discover and explore gender dimensions within the curriculum (and often, not even outside of it).

5.2 Limitations and Constraints

The papers used for the corpus were selected so that they would cover the wide range of topics included in HCI (see Table 3.1). Not all authors have made explicit where their research took place. Still, a strong bias towards works from the US (6) 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). All papers were published by SIGCHI or in ACM journals on HCI, focusing our analysis of on what may be considered “mainstream” regarding gendered issues in HCI. However, we might therefor have missed developments that are already (partially or fully) addressing our critique.

Related to the topic of incompleteness, I acknowledge that the recommendations we derived from the corpus may seem incomplete to some. This is visible e.g. in the feedback I received on

the workshop at ditact, documented in Section 4.3.2. However, rather than providing a complete “how to”, my goal for this work was to create a starting point, a trigger for fruitful discussions.

How we see marginalisation, and how it is discussed in this work, is closely tied to me as a main author, and my contributors and supporters, and can never be complete. All of us are white members of – various levels of – academia. We are culturally tied to majority perspectives within central Europe, and thus can hardly talk about marginalisation due to, e.g., race or forced migration. However, our own personal experiences of marginalisation inform our work, and thus have implicitly influenced our analysis of existing works. These aspects should be taken into account by readers and fellow researchers, when they themselves ask “What can we learn from this?”.

This list of limitations and constraints is in itself most likely incomplete. I am very much looking forward to feedback on the work I have done, on new collaborations to improve upon it, and to discussions that hopefully broaden all participants’ horizons.

5.3 Lessons Learned

Gender issues have existed in HCI pretty much from the start. However, in widely cited historical articles such as Shackel (1997), these topics are not at all discussed. Especially in the last years, lots of research has been done (as documented in Footnote 3), and as both HCI and gender studies are steadily developing, it is hard for newbies to get started. Orientation is hard to develop, and becoming overwhelmed by too much information is a problem.

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, as I realised that context was a thing that should be included. Working on this thesis, some of the works and people I got in touch with challenged my own beliefs and views, for which I am thankful.

Then, a rather pragmatic lesson I learned about myself: I am not very good at working alone. I discovered accountability buddies to be a method that works very well for me¹, as are well-maintained and well-structured to do lists. Sometimes, an action item appears to be huge when seen from afar, but when you break it down and approach it bit by bit, it suddenly becomes very manageable.

Finally, I had to realize, remember, and repeatedly be reminded of the fact that the standards I applied (and probably still apply) to this master’s thesis of mine are way higher than those some people may apply to their dissertation².

5.4 On Attribution

As I have pointed out earlier in this work, attribution and crediting have been bugging me. Pronouns and descriptors used in this work are a bit complicated, as the journey from finding the topic to finally handing in the thesis was not always straightforward.

Often in scientific works, a sense of neutrality, objectivity and scientific-ness is established by having an unspecified “we” or a de-personified “author” act, decide, and report. However, when I use “we”, I want to indicate and make explicit that some things arose and formed in collaboration and exchange with my friends and colleagues.

¹I could not find a comprehensive definition for the concept or method of accountability buddies. For me, it means to have friends check in on my progress relating to specific work packages or the project in general, and sometimes hang out with me while writing or researching, so I would not have to get started and keep working on my own. One fitting description can be found at ADDitude Mag: <https://www.additudemag.com/getting-things-done-with-adhd/>, last accessed June 28, 2021; also saved on archive.org.

²see e.g. the case of the former Austrian minister of labour, as documented by The Guardian: <https://www.theguardian.com/world/2021/jan/09/austrian-minister-resigns-amid-plagiarism-scandal>, last accessed January 1, 2022; also saved on archive.org

Here is how this thesis came to be: I started out writing a literature review for the GEECCO project, on which this thesis should be built. Brigitte Ratzler advised me on some important points that should be covered within the literature review. I myself decided which papers to use as sources, how to analyse them, and what recommendations to take from them.

Then, Katta Spiel took an interest to this work, and proposed we write a paper together for MuC. They also read & analysed the papers, and after comparing notes, we came up with a modified set of recommendations. The paper was submitted, accepted, and even honored with a Best Paper Honorary Mention Award.

I translated the recommendations to less-academic German, found additional categories to group them, and created the initial card deck. Kristina Weinberger provided feedback on these texts. The online version of the card deck was built with help from Matthias Steinböck (who provided the base to draw random cards) and Alexander Steiner (who made the tool screenreader friendly).

I took great care to attribute work correctly, defaulting to a – sometimes explicitly specified – collective “we” as no work can ever be completely done by one person alone. Only where I could say without a doubt that a task, idea, or artefact originated mostly or only from my own work, attribution is made using “I”.

5.5 On Recursion

A rather (un)popular computer science joke is that, “in order to understand recursion, you first must understand recursion”. I started work on the literature review in late 2018. Since then, the topic has barely left my head. Even when I did not actively work on the thesis, or the literature review, I was still busy thinking about it. In my head, there exists the picture of a gigantic omnidirectional graph of things that have influenced each other during these three years: working on the thesis has influenced how I saw and did things in my spare time, or my respective wage job(s). Experiences in my spare time and wage job(s), as well as volunteer work, have influenced how I looked at what I was doing with my thesis. And working on the thesis itself has influenced how I worked on it further – even while working on the corpus, I already started seeing my own limitations, and trying to amend them e.g. by applying the recommendations that we derived from the corpus already into my presentations, discussions, and, of course, this document.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Discussion, Conclusion, and Future Work

For this thesis and the accompanying publications (Burtscher and Spiel (2020, 2021), the card deck, and the drawing tool available online), I conducted a close reading of ten selected HCI papers together with Katta Spiel. We engaged thoroughly with the corpus and their various foci on gender, asking “What can we learn from this?” in order to create a knowledge base for HCI researchers and practitioners who are interested in the gender dimensions of their work, but hitherto had not received any guidance with regard to (intersectional) gender studies basics. The recommendations we deduced from the corpus cover the complete life cycle of HCI research, from designing research protocols to acquiring funding and conducting research to presenting work and outcomes. In Burtscher and Spiel (2020), we additionally present how the recommendations resulting from our work relate to different contribution types as described by Wobbrock and Kientz (2016).

The research questions defined in Section 1.1 are as follows:

1. Which gender issues can be found in HCI and how can we reflect on them in our research?
2. How can HCI research be made gender inclusive?
3. What should a set of recommendations to make HCI research more gender inclusive include in order to help researchers learn about and implement said recommendations?
4. In which settings can said set of recommendations spark discussions among team members so that they more easily find ways to make their work more inclusive?

Answering all these questions in full turned out to go beyond the scope of a Master’s thesis. However, I found partial answers to each of the questions, and discovered some pointers that could be followed up in future research.

Regarding Question 1, “Which gender issues can be found in HCI and how can we reflect on them in our research?”, I described some gender issues that could be found in the literature in Section 2.2. Ideas of how they may be addressed are detailed in the recommendations derived from the corpus in Section 4.2.

To answer Question 2, “How can HCI research be made gender inclusive?”, I again point to Section 4.2. The recommendations therein aim to enable HCI researchers and practitioners to question their current work practices, in order to (better) address gender dimensions as well as other dimensions of marginalisation and discrimination.

6. DISCUSSION, CONCLUSION, AND FUTURE WORK

Within the scope of this thesis, I was able to develop a set of recommendations, created a card deck to function as discussion starters, and conducted an initial user test with a small group of people. However, the test users were no HCI experts. Thus, in order to answer to Question 3, “What should a set of recommendations to make HCI research more gender inclusive include in order to help researchers learn about and implement said recommendations?”, further work together with HCI experts will be necessary.

Finally, in order to answer Question 4, “In which settings can said set of recommendations spark discussions among team members so that they more easily find ways to make their work more inclusive?”, I conducted a test with a small group of 5 participants, with a small individual assignment and a plenary discussion. This setting was reported to work well (see Section 4.3.2). Another workshop with a larger group of participants, using sub-group assignments and a plenary discussion has been conducted, and the card deck has been used as part of an assignment in a lecture at TU Wien’s faculty for Informatics. Findings from both will be presented in future work.

List of Figures

4.1	Development of card layout.	42
-----	-------------------------------------	----

List of Tables

3.1	Characteristics of papers part of our close reading corpus. Papers indicated with * have been added outside of Schlesinger et al.'s corpus. <i>DA</i> stands for “domestic abuse”, <i>MSM</i> stands for “men who have sex with men”, and <i>self-id</i> stands for “self-identification”. Table adapted from Burtscher and Spiel (2020).	14
4.1	Overview of project life cycle phases and how the recommendations are grouped into these phases.	34



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Bibliography

- Ahmed, Syed Ishtiaque, et al. Protibadi: A Platform for Fighting Sexual Harassment in Urban Bangladesh. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14. ACM, New York, NY, USA (2014) pp. 2695–2704.
- Anderberg, Dan, et al. Unemployment and Domestic Violence: Theory and Evidence. In: *The Economic Journal* (2016). volume 126(597):1947–1979.
- Bardzell, Shaowen. Feminist HCI: Taking Stock and Outlining an Agenda for Design. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10. ACM, New York, NY, USA (2010) pp. 1301–1310.
- Bellini, Rosanna, et al. Feminist HCI: Taking Stock, Moving Forward, and Engaging Community. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18. ACM, New York, NY, USA (2018) pp. SIG02:1–SIG02:4.
- Bellini, Rosanna, et al. Mapping the Margins: Navigating the Ecologies of Domestic Violence Service Provision. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19. Association for Computing Machinery, New York, NY, USA (2019) pp. 1–13.
- Blackwell, Lindsay, et al. 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*, CHI '16. Association for Computing Machinery, New York, NY, USA (2016) pp. 610–622.
- Bohan, Janis S. Regarding Gender: Essentialism, Constructionism, and Feminist Psychology:. In: *Psychology of Women Quarterly* (2016).
- Brant, Cathy AR. Teaching our teachers: Trans* and gender education in teacher preparation and professional development. In: *Teaching, Affirming, and Recognizing Trans and Gender Creative Youth*, pp. 47–61. Springer (2016).
- Brekhus, Wayne. A Sociology of the Unmarked: Redirecting Our Focus*:. In: *Sociological Theory* (1998).
- Breslin, Samantha and Wadhwa, Bimlesh. Gender and Human-Computer Interaction. In: *The Wiley Handbook of Human Computer Interaction*, pp. 71–87. Wiley-Blackwell (2017).
- Brulé, Emeline and Spiel, Katta. Negotiating Gender and Disability Identities in Participatory Design. In: *Proceedings of the 9th International Conference on Communities & Technologies - Transforming Communities*, C&T '19. ACM, New York, NY, USA (2019) pp. 218–227.
- Buolamwini, Joy and Gebru, Timnit. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*. In: *Proceedings of Machine Learning Research 81* (2018). pp. 1–15.

- Burman, Barbara. Pocketing the difference: Gender and pockets in Nineteenth-Century Britain. In: *Gender & History* (2002). volume 14(3):447–469.
- Burnett, Margaret, et al. Finding Gender-Inclusiveness Software Issues with GenderMag: A Field Investigation. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016a) pp. 2586–2598.
- Burnett, Margaret, et al. GenderMag: A Method for Evaluating Software's Gender Inclusiveness. In: *Interacting with Computers* (2016b). volume 28(6):760–787.
- Burtscher, Sabrina and Spiel, Katta. "But where would I even start?": Developing (gender) sensitivity in HCI research and practice. In: *Proceedings of the Conference on Mensch Und Computer, MuC '20*. Association for Computing Machinery, New York, NY, USA (2020) pp. 431–441.
- Burtscher, Sabrina and Spiel, Katta. "Let's Talk about Gender" – Development of a Card Deck on (Gender) Sensitivity in HCI Research and Practice Based on a Contrasting Literature Review. In: *i-com* (2021). volume 20(1):85–103.
- Calton, Jenna M., et al. Barriers to Help Seeking for Lesbian, Gay, Bisexual, Transgender, and Queer Survivors of Intimate Partner Violence. In: *Trauma, Violence, & Abuse* (2015).
- Cheryan, Sapna, et al. Ambient belonging: How stereotypical cues impact gender participation in computer science. In: *Journal of Personality and Social Psychology* (2009). volume 97(6):1045–1060.
- Clarke, Rachel, et al. Sharing narrative and experience: Digital stories and portraits at a women's centre. In: *CHI '12 Extended Abstracts on Human Factors in Computing Systems, CHI EA '12*. Association for Computing Machinery, New York, NY, USA (2012) pp. 1505–1510.
- Clarke, Rachel, et al. Digital Portraits: Photo-sharing After Domestic Violence. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13*. ACM, New York, NY, USA (2013) pp. 2517–2526.
- Collins, Patricia Hill. Comment on Hekman's "Truth and method: Feminist standpoint theory revisited": Where's the power? In: *Signs: Journal of Women in Culture and Society* (1997). volume 22(2):375–381.
- Cramer, Elizabeth P and Plummer, Sara-Beth. Social work practice with abused persons with disabilities. In: *Domestic violence: Intersectionality and culturally competent practice* (2010). pp. 128–154.
- Crenshaw, Kimberlé. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. In: *Stanford Law Review* (1991). volume 43(6):1241–1299.
- Criado Perez, Caroline. The deadly truth about a world built for men – from stab vests to car crashes. In: *The Guardian* (2019a).
- Criado Perez, Caroline. *Invisible Women: Exposing Data Bias in a World Designed for Men*. Penguin Random House (2019b).
- Davis, Hilary and Farmer, Jane. Digital participation: Engaging diverse and marginalised communities. In: *Proceedings of the 28th Australian Conference on Computer-Human Interaction, OzCHI '16*. Association for Computing Machinery, New York, NY, USA (2016) pp. 672–675.

- Dervin, Fred. Discourses of Othering. In: *The International Encyclopedia of Language and Social Interaction*, pp. 1–9. American Cancer Society (2015).
- DeVries, Terrance, et al. Does Object Recognition Work for Everyone? In: *arXiv:1906.02659 [cs]* (2019).
- Duarte, Ana Maria Bustamante, et al. Participatory Design and Participatory Research: An HCI Case Study with Young Forced Migrants. In: *ACM Transactions on Computer-Human Interaction* (2018). volume 25(1):3:1–3:39.
- Erdmann, Vera and Schumann, Tanja. EU engineering report. Technical report, Institut der deutschen Wirtschaft Köln (2010).
- Fausto-Sterling, Anne. *Sexing the Body: Gender Politics and the Construction of Sexuality*. Basic Books, New York, NY (2000).
- Fernandez, Julia R. and Birnholtz, Jeremy. "I Don't Want Them to Not Know": Investigating Decisions to Disclose Transgender Identity on Dating Platforms. In: *Proceedings of the ACM on Human-Computer Interaction* (2019). volume 3(CSCW):226:1–226:21.
- Fine, Cordelia, et al. Eight Things You Need to Know About Sex, Gender, Brains, and Behavior: A Guide for Academics, Journalists, Parents, Gender Diversity Advocates, Social Justice Warriors, Tweeters, Facebookers, and Everyone Else. In: *S&F Online* (2019). (15.2).
- González-González, Carina S., et al. Gender Design Methods for Engineering, Responsible Innovation and Interaction. In: *Proceedings of the XIX International Conference on Human Computer Interaction*, Interacción 2018. Association for Computing Machinery, New York, NY, USA (2018) pp. 1–2.
- Grabher, Angelika, et al. Sonderauswertung der Studierenden-Sozialerhebung 2011 für die Technische Universität Wien. Technical report, Institut für Höhere Studien (2014).
- Haimson, Oliver. Social Media as Social Transition Machinery. In: *Proceedings of the ACM on Human-Computer Interaction* (2018). volume 2(CSCW):63:1–63:21.
- Haimson, Oliver L., et al. 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*, CHI '14. ACM, New York, NY, USA (2014) pp. 1615–1624.
- Haimson, Oliver L., et al. Life Transitions and Social Technologies: Research and Design for Times of Life Change. In: *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, CSCW '19. Association for Computing Machinery, New York, NY, USA (2019) pp. 480–486.
- Haraway, Donna. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. In: *Feminist Studies* (1988). volume 14(3):575–599.
- Harding, Sandra. *Objectivity and Diversity: Another Logic of Scientific Research*. University of Chicago Press, Chicago (2015).
- Hartsock, Nancy CM. The feminist standpoint: Developing the ground for a specifically feminist historical materialism. In: *Karl Marx*, pp. 565–592. Routledge (2017).
- Hill, Charles G., et al. Gender-Inclusiveness Personas vs. Stereotyping: Can We Have it Both Ways? In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17. Association for Computing Machinery, New York, NY, USA (2017) pp. 6658–6671.

- Humer, Tobias. Neuer Erlass zum dritten Geschlechtseintrag - Ein Teilerfolg. https://www.ots.at/presseaussendung/OTS_20200910_OTS0011/neuer-erlass-zum-dritten-geschlechtseintrag-ein-teilerfolg (2020).
- Hunt, Sarah. Embodying self-determination: Beyond the gender binary. In: *Determinants of Indigenous Peoples' Health*, pp. 22–39. Canadian Scholars' Press, Toronto (2018).
- James DiSalvo, Betsy, et al. African American Men Constructing Computing Identity. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11. ACM, New York, NY, USA (2011) pp. 2967–2970.
- Jaroszewski, Samantha, et al. "Genderfluid" or "Attack Helicopter": Responsible HCI Research Practice with Non-binary Gender Variation in Online Communities. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18. ACM, New York, NY, USA (2018) pp. 307:1–307:15.
- Karuei, Idin, et al. Detecting Vibrations Across the Body in Mobile Contexts. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11. ACM, New York, NY, USA (2011) pp. 3267–3276.
- Keyes, Os. The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. In: *Proc. ACM Hum.-Comput. Interact.* (2018). volume 2(CSCW):88:1–88:22.
- Keyes, Os, et al. Reimagining (Women's) Health: HCI, Gender and Essentialised Embodiment. In: *ACM Transactions on Computer-Human Interaction* (2020). volume 27(4):25:1–25:42.
- Kirakowski, Jurek and Norman, Kent L. Introduction: Human-Computer Interaction Yesterday, Today, and Tomorrow. In: *The Wiley Handbook of Human Computer Interaction*, pp. 1–6. Wiley-Blackwell (2017).
- Kumar, Neha, et al. Engaging Feminist Solidarity for Comparative Research, Design, and Practice. In: *Proceedings of the ACM on Human-Computer Interaction* (2019). volume 3(CSCW):167:1–167:24.
- Leavy, Susan. Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. In: *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering*, GE '18. Association for Computing Machinery, New York, NY, USA (2018) pp. 14–16.
- Marsden, Nicola, et al. Developing personas, considering gender: A case study. In: *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, OZCHI '17. Association for Computing Machinery, New York, NY, USA (2017) pp. 392–396.
- Martin, JR. Close reading: Functional linguistics as a tool for critical discourse analysis. In: *Researching language in schools and communities: Functional linguistic perspectives* (2005). p. 275.
- Mendez, Christopher, et al. From GenderMag to InclusiveMag: An Inclusive Design Meta-Method. In: *2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)* (2019) pp. 97–106.
- Metaxa-Kakavouli, Danaë, et al. Gender-Inclusive Design: Sense of Belonging and Bias in Web Interfaces. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18. Association for Computing Machinery, New York, NY, USA (2018) pp. 1–6.

- Nicholson, Linda. Interpreting Gender. In: *Signs* (1994). volume 20(1):79–105.
- Ogbonnaya-Ogburu, Ihudiya Finda, et al. Critical Race Theory for HCI. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20. Association for Computing Machinery, New York, NY, USA (2020) pp. 1–16.
- Otterbacher, Jahna. Crowdsourcing Stereotypes: Linguistic Bias in Metadata Generated via GWAP. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15. ACM, New York, NY, USA (2015) pp. 1955–1964.
- Paoletti, Jo B. Clothing and Gender in America: Children's Fashions, 1890-1920. In: *Signs: Journal of Women in Culture and Society* (1987). volume 13(1):136–143.
- Pruitt, John and Grudin, Jonathan. Personas: Practice and Theory. In: *Proceedings of the 2003 Conference on Designing for User Experiences*, DUX '03. ACM, New York, NY, USA (2003) pp. 1–15.
- Ratzer, Brigitte. GEECCO Project. <http://www.geecco-project.eu/home/> (2018).
- Reiley, Carol. When bias in product design means life or death (2016).
- Ristock, Janice and Timbang, Norma. Relationship violence in lesbian/gay/bisexual/transgender/queer [LGBTQ] communities. In: *Violence Against Women Online Resources* (2005).
- Rode, Jennifer A. Reflexivity in digital anthropology. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11. Association for Computing Machinery, New York, NY, USA (2011) pp. 123–132.
- Roig-Maimó, Maria Francesca and Mas-Sansó, Ramon. The Female Effect: The Case of Gender Parity on User Studies. In: *Proceedings of the XX International Conference on Human Computer Interaction*, Interacción '19. Association for Computing Machinery, New York, NY, USA (2019) pp. 1–6.
- Saha, Koustuv, et al. The Language of LGBTQ+ Minority Stress Experiences on Social Media. In: *Proceedings of the ACM on Human-Computer Interaction* (2019). volume 3(CSCW):89:1–89:22.
- Scheuerman, Morgan Klaus, et al. HCI Guidelines for Gender Equity and Inclusivity. <https://www.morgan-klaus.com/gender-guidelines.html> (2020).
- Schlesinger, Ari, et al. Intersectional HCI: Engaging Identity Through Gender, Race, and Class. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17. ACM, New York, NY, USA (2017) pp. 5412–5427.
- Shackel, Brian. Human-Computer Interaction—Whence and Whither? In: *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE* (1997). volume 48(11):970–986.
- Spiel, Katta, et al. How to do better with gender on surveys: A guide for HCI researchers. In: *Interactions* (2019a). volume 26(4):62–65.
- Spiel, Katta, et al. Patching Gender: Non-binary Utopias in HCI. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19. ACM, New York, NY, USA (2019b) pp. alt05:1–alt05:11.
- Spiel, Katta, et al. In the details: The micro-ethics of negotiations and in-situ judgements in participatory design with marginalised children. In: *CoDesign* (2020). volume 16(1):45–65.

- Starks, Denny L., et al. Designing Technology to Support Safety for Transgender Women & Non-Binary People of Color. In: *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion, DIS '19 Companion*. Association for Computing Machinery, New York, NY, USA (2019) pp. 289–294.
- Stengers, Isabelle. *Another Science Is Possible: A Manifesto for Slow Science*. John Wiley & Sons (2018).
- Stimpson, George William. *Popular Questions Answered*. G. Sully, Incorporated (1930).
- Strohmayr, Angelika, et al. CHIInclusion: Working Toward a More Inclusive HCI Community. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, CHI EA '19*. Association for Computing Machinery, New York, NY, USA (2019) pp. 1–10.
- Stumpf, Simone, et al. Gender-Inclusive HCI Research and Design: A Conceptual Review. In: *Foundations and Trends® in Human-Computer Interaction* (2020). volume 13(1):1–69.
- Talhok, Reem, et al. HCI and refugees: Experiences and reflections. In: *Interactions* (2018). volume 25(4):46–51.
- Tatman, Rachael. Google's speech recognition has a gender bias. <https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/> (2016).
- Turkle, Sherry. Computational reticence: Why women fear the intimate machine. In: *Technology and Women's Voices*, pp. 41–61. Cheris Kramarae (ed.), New York, NY, USA, pergamon press edition (1986).
- Von Ahn, Luis and Dabbish, Laura. ESP: Labeling images with a computer game. In: *AAAI Spring Symposium: Knowledge Collection from Volunteer Contributors*, volume 2 (2005) p. 1.
- Vorvoreanu, Mihaela, et al. From Gender Biases to Gender-Inclusive Design: An Empirical Investigation. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19*. Association for Computing Machinery, New York, NY, USA (2019) pp. 1–14.
- Wachter-Boettcher, Sara. *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*. W. W. Norton & Company, New York, NY, 1 edition edition (2017).
- Wiener, Melanie, et al. The shift towards entrepreneurial universities and the relevance of third-party funding of business and economics units in Austria: A research note. In: *Review of Managerial Science* (2020). volume 14(2):345–363.
- Wobbrock, Jacob O. and Kientz, Julie A. Research contributions in human-computer interaction. In: *Interactions* (2016). volume 23(3):38–44.