

Gender HCI and Microsoft: Highlights from a Longitudinal Study

Margaret Burnett¹, Robin Counts², Ronette Lawrence², Hannah Hanson³

¹Oregon State University
Corvallis, Oregon, USA
burnett@eecs.oregonstate.edu

²Microsoft
Redmond, Washington, USA
{roco,ronettel}@microsoft.com

³Hitachi Consulting
Seattle, Washington, USA
Hannah.Hanson@hitachiconsulting.com

Abstract—Research has emerged over the past decade showing gender biases in software. Although a few methods and prototype systems have emerged to help address this issue, none have been reported to have an impact on the people who actually build software. In this paper, we summarize a few highlights from a year-long field study investigating how Gender HCI methods to address gender biases in software can make impacts on a large software company.

Keywords—GenderMag, developing gender-inclusive software

I. INTRODUCTION

Gender inclusiveness in technology has become a highly visible social issue [3, 4, 19, 39, 42, 44, 45, 52]. One form of gender inclusiveness is *software's* gender inclusiveness. More specifically, because males and females often work differently with software, a software product can be unintentionally biased to the needs of one gender while marginalizing another. Designing software products for gender-inclusiveness does not suggest designing separate products, (e.g., a “pink version” and “blue version”) but rather removing individual inclusiveness barriers to accommodate a wide range of working styles no matter what gender(s) in which they occur.

From this, it follows that helping software professionals see and understand such biases could lead to products that serve more people more fully. Toward that end, we built a Gender HCI method to enable software professionals to concretely find such biases in their own software.

In this paper, we describe a few highlights from an effort to enact a real-world behavior change at Microsoft: to inspire Microsoft software professionals to apply Gender HCI to the software products they create. An appropriate methodology for studying change while attempting to enact it sustainably is known as Action Research. Action Research is an approach “that involves engaging with a community to address some problem... and through this problem solving to develop scholarly knowledge” [25]. As Hayes explains, “the cornerstone ... is that these two cannot be disentangled: the doing and the knowing, the intervention and the learning <about it>” [25].

One way Action Research aims at sustainable change is to be explicitly collaborative with the participants. That is, research is done “with” the participants, not just “to” or “for” or “focused on” them. Thus, some of the participants take on the role of researchers themselves [25] and take part in determining the manipulations of the potential solution being studied throughout the study. For example, Carroll et al. used it in this fashion to research and encourage community engagement in their own information infrastructures [13].

This work supported in part by NSF 1314384 and 1528061, and by Microsoft.

Therefore, our longitudinal study was a fully collaborative endeavor with Microsoft. As per Action Research, we did the research over the course of a year with many of the Microsoft participants serving as researchers, working in the ways they thought best for their views and corner of the company. Our goal was to provide insights into the following question:


What does it take to “land” gender inclusiveness in software development practices at a large technology company?

II. BACKGROUND AND RELATED WORK

We began by introducing into Microsoft a Gender HCI method known as GenderMag (Gender-Inclusiveness Magnifier) [12]. GenderMag is an inspection method to enable software professionals to evaluate software they are building from a gender-inclusiveness perspective. This method eventually spawned numerous derivatives at Microsoft. Thus, we will use the term *GenderMag* only when referring to the GenderMag method alone, and *Gender HCI* when referring to all the variants and derivatives that ultimately came into use.

At GenderMag's core are five facets (shown in Fig. 1) of people's problem-solving approaches that tend to cluster by

Abby (Abigail) Jones¹



Customizable background, age, place of residence, occupation, and interests are here.

- Motivations:** Abby is proficient with the technologies she uses. She learns new technologies when she needs to, but doesn't ... explore technology ... <prefers> methods she is already familiar and comfortable with to achieve her goals.
- Information Processing Style:** Abby leans towards a comprehensive information processing style ... she first gathers information comprehensively to try to form a complete understanding of the problem before trying to solve it. ...
- Computer Self-Efficacy:** Abby has ... low self-confidence in performing computing tasks ... <so> blames herself for problems that she encounters.
- Attitude toward Risk:** Abby is risk averse when she uses computers... When confronted with new software features, Abby worries that she will spend time on them and not get any benefits from doing so...
- Willingness to Explore and Tinker:** Abby doesn't particularly like tinkering with software (ie, just trying out new features or commands to see what they do)... Instead, she prefers following step-by-step tutorials and wizards. (However, when she does tinker, it has positive effects on her understanding of the software.) ...

¹ Footnote is here linking to the persona's data foundations.

Fig. 1. The Abby persona as it was in the first version used at Microsoft. (Portions elided, others enlarged for readability.)

gender [2, 5, 9, 10, 14, 15, 20, 24, 30, 38, 47]. GenderMag makes the facets concrete with a set of four faceted personas—“Abby” (Fig. 1), “Patricia,” “Patrick” and “Tim”. Tim’s facet values are frequently seen in males, Abby’s are those frequently seen in females that are the most different from Tim’s, and the two Pats’ (identical) facet values add coverage of a large fraction of females and males different from Abby and Tim.

GenderMag intertwines these faceted personas with a specialized Cognitive Walkthrough (CW) [48, 51]. In a GenderMag CW, evaluators answer two CW questions for each step of a detailed use case with respect to the five facets, from the perspective of one of the above personas:

- Will <persona> know what to do at this step? (Yes/no/maybe, why)
- If <persona> does the right thing, will s/he know s/he did the right thing & is making progress toward their goal? (Yes/no/maybe, why)

Uses of GenderMag have begun to emerge in the literature [11, 18, 27, 28, 35], but none of these works investigates the use of practices or methods for *gender-inclusive software* over a period *longer than a couple of hours*. That is the gap this paper helps to fill.

III. METHODOLOGY: ACTION RESEARCH

Action Research (Fig. 2) is an approach to long-term field research that involves engaging with a community for a dual purpose. As a type of field study, it aims to develop scholarly knowledge about a potential solution to a problem. However, it is unlike most studies in that it does not have controls and is “hands on”—it aims to *address* the problem in that particular community *during* (throughout) the field study [25, 50].

Our initial input into Microsoft was to introduce the GenderMag method, but along the way this “treatment” evolved into a variety of Gender HCI methods. We iteratively collected data on what happened next while also iterating on the method(s) to improve them, looping around to collect more data, and so on, as per the Action Research loop.

Despite the fact that the “treatment” (here, Gender HCI methods) changes continuously, Action Research shares with other empirical methodologies an emphasis on rigor. In Action Research, rigor focuses especially on credibility and validity, which are attained primarily through member checking (verifying interpretations of events by the participants themselves) and through triangulation (ascertaining whether multiple sources of evidence produce the same conclusion).

As per Action Research, our participants were of two sorts. Participant-researchers were Microsoft employees who stepped forward to try out GenderMag or to otherwise contribute to seeing where it would lead within Microsoft. As of April 2016, we counted 20 participant-researchers, but there may be more who did not identify themselves to us. In addition to participant-researchers were “regular” participants: those who also participated in the changing processes. These participants did not do research per se, but they participated in and experienced variants of the method, and had opportunities to act upon what they experienced. We do not know exactly how many Microsoft employees ultimately fell under this category, but we know that there were at least 655 of them.

We collected a wide variety of data sources about usage

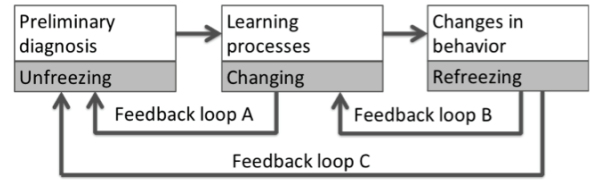


Fig. 2. Action Research’s three stages.

across the company: detailed accounts of using the method and its derivatives, download counts of internal videos about the method and its derivatives, attendee counts at Gender HCI events, anonymous workshop feedback forms, emails and memos about Gender HCI, internal artifacts, surveys collecting Gender HCI data, and user experience reports on product outcomes. These multiple sources of evidence allowed extensive triangulation.

IV. THE “UNFREEZE” STAGE

Our investigation began with the Action Research “Unfreeze” stage (Fig. 2, left node). The Unfreeze stage is recognizing a problem and generating an impetus for change.

The organization’s “unfreeze” from the perspective of Gender HCI began with a changing of the guard. In February 2014, Microsoft’s Board of Directors appointed Satya Nadella as Chief Executive Officer and member of the Board of Directors, replacing S. Ballmer, who had been CEO for 14 years. In June 2015, Nadella announced a new mission statement for Microsoft that rested heavily upon Microsoft’s ability to build inclusive products [49]:

CEO: Our mission is to empower every person and every organization on the planet to achieve more.

The CEO’s internal company email [7] accompanying the new mission statement emphasized the need for growth in three areas: an “obsession” with customers, collaboration, and—especially pertinent to this paper—diversity and inclusion:

CEO: Diverse and inclusive: The world is diverse. We will better serve everyone on the planet by ... be<ing> open to learning our own biases and changing our behaviors... We don’t just value differences, we seek them out... as a result, our ideas are better, our products are better and our customers are better served...

To allow the new mission to take hold, the CEO then needed to propagate throughout the company the “unfreezing” with respect to that mission. Toward that end, the CEO initiated ways to hold employees accountable for inclusiveness in general—including gender-inclusiveness. As P17, a participant a few levels below the CEO, put it:

P17:<Accountability for inclusiveness is> precisely what happened. CorporateVP2 is making this a top priority for all of the <Division>. This is coming down to all the managers.

Lesson 1 (A precondition: *inclusiveness accountability*): We believe that the top-down propagation of accountability for making progress on software inclusiveness was key to our participants’ willingness to consider incorporating Gender HCI methods into their practices.

V. THE “CHANGING” STAGE

From a Gender HCI perspective, the Action Research “Changing” stage (Fig. 2, middle node and Feedback loop A)

began in July 2015.

In July 2015, Researcher1 gave a 20-minute presentation on GenderMag to a group of faculty researchers and Microsoft engineers at the 2015 Microsoft Faculty Summit. In attendance were about 30 Software Engineering faculty members from around the world and about 30 Microsoft employees (mostly engineers). Researcher1's talk attracted significant interest. Several follow-up meetings from the GenderMag talk ensued, and on July 29, 2015, Researcher1 and TeamGroupA employees P1 and P3 decided to try out the GenderMag process on some of TeamGroupA's products.

Lesson 2 (Political perils and optionality of gender inclusiveness): A number of teams that at first expressed interest did not follow through. This can occur with any new approach, but in the case of Gender HCI it was sometimes tied to political perils such as the gender politics of talking explicitly about gender, or not wanting to argue with a vocal naysayer about its importance. Such issues were sometimes magnified by competing priorities winning out over gender inclusiveness when gender inclusiveness was viewed as being too controversial or "more optional" than other priorities.

TeamGroupA was a group of 5 development teams working for P3 (TeamGroupA's Engineering Lead); they were the first 5 teams at Microsoft to try the method. As the first 5 teams, they had the greatest lifespan in our yearlong investigation, and we have the most data about their experiences.

TeamGroupA had inherited a multi-platform product targeting IT professionals, and wanted to broaden the customer base to small business owners. Toward that end, their UX team had recently prepared personas that reminded P3, TeamGroupA's Engineering Lead, of the GenderMag personas Tim and Abby. Further, since half of all small business owners are female, P3 saw GenderMag as being particularly pertinent to their goals:

P3: If I can't make the product work for Abby, I will have failed.

The first GenderMag session at Microsoft was attended by five males and one female from a TeamGroupA team. Two were software engineers or engineering managers, one was a program manager, and three were user experience (UX) people. They began to walk through a fairly basic product and immediately started finding issues. At the time of their GenderMag evaluation, their product was partially implemented and partially in design mockup form.

By the end of that session, their GenderMag CW forms revealed that 54% of the features they evaluated had usability issues, half of which (27% of all the features evaluated) were gender-inclusiveness issues as per the definition in [11].

The team was enthusiastic about the results. Five out of the six said it had revealed issues they hadn't realized before. The sixth, P1, added credibility: s/he said that, in prior user research s/he had conducted, P1 had actually seen all of the issues the GenderMag session revealed. The team agreed on a list of 14 changes to be made in their software product.

As the week progressed, the TeamGroupA teams brought more advanced functionalities into the evaluation. Since those functionalities involve even more problem-solving by the user, usability issues in general went up to an average of 77%, and gender-inclusiveness issues relating to problem-solving styles

also increased, averaging around 52% in the last three sessions, as per the teams' recordings on their GenderMag forms.

One issue was the amount of pre-work (setup time) required. To lighten the process, Researcher1 and P1 collaboratively experimented over the week. By iterating on the method over the week's sessions, they were able to make it much lighter, ultimately eliminating most of the pre-work.

By the end of that first week of working with TeamGroupA, we learned two more lessons:

Lesson 3 (More than inclusiveness evaluation): GenderMag had to serve a dual purpose: it had to not only effectively help the team find improvements their product needed, but also to *educate* people about diversity in software products. At times, this duality added enough weight and difficulties to threaten the entire effort.

Lesson 4 (Talking about gender by not talking about gender): The five facets were a non-gendered vocabulary that gave software teams explicit and actionable ways of talking about gender biases in their software. This aspect turned out to be key to some people's willingness to participate in gender inclusiveness conversations.

VI. TOWARD REFREEZING: INITIAL IMPACTS

Due in part to interest from leadership and in part to TeamGroupA's successful experiences with GenderMag, a variety of other teamgroups started experimenting with GenderMag and derivative Gender HCI methods that they began to derive to fit their teamgroups' particular needs and processes.

Using data from all of these teams, after about 10 months, we began to measure initial impacts across the company.

A. Impacts on Software Products

We know of twelve teams who changed the software products they produce as follow-ups to using some variant of a Gender HCI method. For example:

P20: (I) rewrote the compiler error messages in ProductX to make use of Researcher1's research.

P14: ... looking at ProductD-xFeature ... uncovered many general usability and gender related issues... <The TeamGroupD team working on xFeature> along with two other key teams are addressing <the issues> in the next version of ProductD.

By November 2016, nine other teams from TeamGroupD had also made changes like those P14 describes.

The facets as a vocabulary turned out to be a particularly accessible tool for the teams. P13 explicitly referred to using them in expanding upon P14's description of the ProductD changes based on the facets:

P13: This sprint <ProductD-xFeature improvement> was a first step towards addressing gender inclusion...and we know there are other factors that contribute to gender bias... We will continue to explore these factors in the coming months.

As for TeamGroupA, we have the most data about their product, because their product changes have been in effect long enough to collect customer ratings data. Old vs. new ratings were a little challenging to compare because they measured slightly different things, but one of TeamGroupA's feature sets' customer ratings improved by over 40% (i.e., about 1.5 times the previous score); other feature sets had more modest but still

positive gains. In qualitative user studies of TeamGroupA's product redesign, results were also positive, with comments generally preferring the redesigns that followed from using Gender HCI methods. For example:

UserStudyParticipant1: This is really easy to do. <Compared to older product>, this is all in one place.

UX research report: ...Very positive responses from both novice and experienced <customers>.

Ultimately, TeamGroupA was so encouraged by the new customer ratings, they posted about the product improvements on their "Brag Wall."

B. Impacts on Work Processes

As to impacts on people's work processes, these varied greatly. The event that facilitated work process impact the most was probably the deconstruction of GenderMag into multiple intellectual tools, because the deconstruction afforded teams' ability to evolve the bits and pieces they wanted into their own work processes.

P2: I have Abby's persona pinned to my corkboard with all the good bits (facets) highlighted. ... It also allowed people who were emotionally involved in the product creation the freedom or ability to pull themselves out of the equation and talk about Abby.

P3: After the first session, my team members came up with a simpler process – that eliminates any pre-work ...

P14: We're not actually implementing <GenderMag> as-is, ...much too difficult for any team without a facilitator. <But> we are looking at <facets> ...

P16: We use GenderMag kinds of questions in <our> heuristic evaluation walkthroughs.

The part of GenderMag that seems to have impacted people's work practices the most are the facets and the vocabulary they bring about ranges of problem-solving approaches.

P14: The most useful pieces for us have been thinking in spectra rather than binary ...<the spectra of facet values> have proven directly applicable to decision making in SW design.

Building upon this vocabulary and more, TeamD2 is creating intellectual tools and activities to inspire other teams throughout the company to embrace Gender HCI practices in their own work:

P13: We'll focus on activating more teams with simple starting points for gender inclusion, led by learning styles and self-efficacy.

C. Impacts on Awareness and Mindset

The greatest impact overall has been a mindset shift in the way Microsoft employees talk and think about their users.

P11: ... making our software more inclusive without resorting to cliché.... There is immense value in ...understanding that gender biased software arises in subtle ways....

Measuring changes in awareness and mindset is not easy, but we can provide at least some evidence from interest in an internal "Gender-Inclusiveness Summit," held in late June 2016. On the day of the event, 268 people attended: 85 people in person, and 183 via Skype. The presentations were recorded and uploaded to the company intranet, and within the first month it had been viewed 524 times, making it the second-most popular video on the company-wide research server that month.

Another indicator lies in the feedback forms from attendees at some of the Gender HCI training events over the course of the year. Among them are comments employees made saying that their minds had changed in some way. For example:

P19: We learned that when you build for the average case, you could very easily leave out one gender or the other. That was a big Aha. Applying <the Gender HCI tools> to everything we do is the kind of thing I'll take back to my day job.

P25: I never thought about the way women think vs. the way I think and want things done.

P23: ...I ... learned something that changed the way I think.

P21: I realize that the people who are using <product> look, feel, act and identify in very different ways from each other.

P22: ...Now I have a better way to think about how to address issues that will benefit everyone, but especially women.

VII. CONCLUDING REMARKS

The time may be right for Gender HCI methods to have a chance of taking hold in real-world software companies. The results of our year-long study suggest that integrating Gender HCI methods into software development can pay off—for both the business case and the social justice case. Among the lessons we learned are:

- Impacts: Gender HCI practices have begun to impact not only processes but also mindsets ("...changed the way I think") and software products. Improvements have been noticeable and encouraging: one team's use of Gender HCI practices improved their customer ratings by over 40%.
- From top-level interest to widespread ownership of inclusiveness: Top-level interest in gender-inclusiveness was key to Gender HCI methods having a chance initially, but grassroots ownership brings staying power.
- Gender HCI minefields: Gender HCI is fraught with political peril. Among the key factors that enabled the Gender HCI work to survive these dangers were its initial foundations of extensive empirical data; a vocabulary that was about facets instead of gender; and bringing people *toward* understanding that one gender or another can be disproportionately affected by design designs, but *away* from binary bucketing like "women are pink, men are blue".

Our journey to understand what it takes to bring Gender HCI methods to a large technology organization is still ongoing. One thing we have learned is the fragility of such an endeavor. Our work was beset with competing priorities, erratic funding, differing philosophies on gender inclusiveness, and the political sensitivity of talking about gender.

Despite these difficulties, Gender HCI is gaining traction at Microsoft. We hope that other researchers and practitioners will brave the pitfalls, armed in part by the lessons reported here, to embark on Gender HCI journeys in their own settings.

REFERENCES

- [1] T. Adlin and J. Pruitt. 2010. The Essential Persona Lifecycle: Your Guide to Building and Using Personas. Elsevier, San Francisco, CA.
- [2] M. Arcand, J. Nantel. 2012. Uncovering the nature of information processing of men and women online: The comparison of two models using the think-aloud method. J. Theor. Appl. Elec. Comm. Res. 7, 2: 106-120.

- [3] S. Bardzell. 2010. Feminist HCI: Taking stock and outlining an agenda for design. *ACM CHI*, 1301-1310.
- [4] S. Bardzell, E. Churchill. 2011. Feminism and HCI: New perspectives. *Special Issue Eds. Introduction. Interacting with Computers* 23, 5: iii-xi.
- [5] L. Beckwith, C. Kissinger, M. Burnett, S. Wiedenbeck, J. Lawrance, A. Blackwell, C. Cook. 2006. Tinkering and gender in end-user programmers' debugging. *ACM CHI*, 231-240.
- [6] A. Bhargava. 2002. Gender bias in computer software programs: A checklist for teachers. *Information Tech. Childhood Educ.* 2002, 1: 205-218.
- [7] T. Bishop. 2015. Satya Nadella reveals Microsoft's new mission statement, sees 'tough choices' ahead. *Geek Wire*. June 25, 2015.
- [8] M. Borkin, C. Yeh, M. Boyd, P. Macko, K. Gajos, M. Seltzer, H. Pfister. 2013. Evaluation of filesystem provenance visualization tools. *IEEE Trans. Visualization and Computer Graphics* 19, 12: 2476-2485.
- [9] M. Burnett, L. Beckwith, S. Wiedenbeck, S. Fleming, J. Cao, T. Park, V. Grigoreanu, K. Rector. 2011. Gender pluralism in problem-solving software. *Interacting with Computers* 23, 5: 432-460.
- [10] M. Burnett, S. Fleming, S. Iqbal, G. Venolia, V. Rajaram, U. Farooq, V. Grigoreanu, M. Czerwinski. 2010. Gender differences and programming environments: Across programming populations. *ACM/IEEE ESEM*.
- [11] M. Burnett, A. Peters, C. Hill, N. Elarie. 2016. Finding gender inclusiveness software issues with GenderMag: A field investigation. *ACM CHI*.
- [12] M. Burnett, S. Stumpf, J. Macbeth, S. Makri, L. Beckwith, I. Kwan, A. Peters, W. Jernigan. 2016. GenderMag: A method for evaluating software's gender inclusiveness. *Interacting with Computers*, in press.
- [13] J. Carroll, M. Horning, B. Hoffman, C. Ganoe, H. Robinson, M. Rosson. 2011. Visions, participation and engagement in new community information infrastructures. *J. Community Informations* 7, 3: 783-812.
- [14] S. Chang, V. Kumar, E. Gilbert, L. Terveen. 2014. Specialization, homophily, and gender in a social curation site: Findings from Pinterest. *ACM CSCW*, 674-686.
- [15] G. Charness, U. Gneezy. 2012. Strong evidence for gender differences in risk taking. *J. Economic Behavior & Organization* 83, 1: 50-58.
- [16] D. Compeau and C. Higgins. 1995. Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly* 19, 2: 189-211.
- [17] J. Contrera. 2016. The end of 'shrink it and pink it': A history of advertisers missing the mark with women. *Washington Post*, June 8, 2016.
- [18] S. Cunningham, A. Hinze, D. Nichols. 2016. Supporting gender-neutral digital library design: A case study using the GenderMag toolkit. *Int. Conf. Asia-Pacific Digital Libraries*, 45-50.
- [19] M. Day. 2016. Curious LinkedIn search results highlight tech's diversity problems. *Seattle Times*. Aug 31, 2016.
- [20] T. Dohmen, A. Falk, D. Huffman, U. Sunde, J. Schupp, G. Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *J. European Economic Association* 9, 3: 522-550.
- [21] A. Forte, J. Antin, S. Bardzell, L. Honeywall, J. Riedl and S. Stierch. 2012. Some of all human knowledge: Gender and participation in peer production. *ACM CSCW*, 33-36.
- [22] E. Friess. 2012. Personas and decision making in the design process: an ethnographic case study. *ACM CHI*, 1209-1218.
- [23] V. Grigoreanu, M. Mohanna. 2013. Informal cognitive walkthroughs (ICW): Paring down and pairing up for an agile world. *ACM CHI*.
- [24] J. Hallström, H. Elvstrand, K. Hellberg. 2015. Gender and technology in free play in Swedish early childhood education. *Int. J. Tech. Design Educ.* 25, 2: 137-149.
- [25] G. Hayes. 2014. Knowing by doing: Action Research as an approach to HCI. In *Ways of Knowing in HCI*, J. Olson and W. Kellogg (eds.). Springer, NY, 49-67.
- [26] I. Heemskerk, G. ten Dam, M. Volman, and W. Admiraal. 2009. Gender inclusiveness in educational technology and learning experiences of girls and boys. *J. Research Tech. Educ.* 41, 3.
- [27] C. Hill, S. Ernst, A. Oleson, A. Horvath, M. Burnett. 2016. GenderMag experiences in the field: The whole, the parts, and the workload. *IEEE VLHCC*.
- [28] C. Hill, M. Haag, A. Oleson, C. Mendez, N. Marsden, A. Sarma, M. Burnett. 2017. Gender-inclusiveness personas vs. stereotyping: Can we have it both ways? *ACM CHI*.
- [29] R. Ibrahim, G. Wills, and L. Gilbert. 2010. deGendering Games: Towards a gender-inclusive framework for games. *IADIS Int. Conf. Games and Entertainment Tech. (MCCSIS)*, 127-130.
- [30] C. Kelleher. 2009. Barriers to programming engagement. *Advances in Gender and Education* 1, 1: 5-10.
- [31] K. Lewin. 1958. Group decision and social change. *Readings in Social Psychology*, T. Newcomb, E. Hartley (eds.) NY, Holt, Rinehart, Winston.
- [32] C. Lewis, P. Polson, C. Wharton, J. Rieman. 1990. Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. *ACM CHI*, 235-242.
- [33] T. Mahatody, M. Sagar, C. Kolski. 2010. State of the art on the cognitive walkthrough method, its variants and evolutions. *Int. J. Human-Computer Interaction* 26, 8: 741-85.
- [34] N. Marsden, M. Haag. 2016. Stereotypes and politics: Reflections on personas. *ACM CHI*, 4017-4031.
- [35] N. Marsden, M. Haag. 2016. Evaluation of GenderMag personas based on persona attributes and persona gender. *HCI Intl.*, Part I, 122-127.
- [36] A. Massanari. 2010. Designing for imaginary friends: Information architecture, personas, and the politics of user-centered design. *New Media & Society* 12, 3: 401-416.
- [37] T. Matthews, T. Judge, S. Whittaker. 2012. How do designers and user experience professionals actually perceive and use personas? *ACM CHI*.
- [38] J. Meyers-Levy and B. Loken. 2015. Revisiting gender differences: What we know and what lies ahead. *J. Consumer Psych.* 25, 1: 129-149.
- [39] MIT Media Lab. 2016. Media Lab conference addresses gender bias, diversity, and inclusion in STEM. Retrieved Sept. 17, 2016 from <http://news.mit.edu/2016/media-lab-conference-addresses-gender-bias-diversity-inclusion-in-stem-0914>
- [40] L. Nielsen and K. Hansen. 2014. Personas is applicable: A study on the use of personas in Denmark. *ACM CHI*, 1665-1674.
- [41] L. Palen. 2010. Better ODDS than "Snowballs in Hell?" -or- What might HCI do with action research? *Human-Computer Interaction Consortium*.
- [42] D. Paredes. 2016. Gender equity: Why it matters. Retrieved Sept. 17, 2016 <http://www.cio.co.nz/article/606066/gender-equity-why-it-matters/>
- [43] J. Pruitt, J. Grudin. 2003. Personas practice and theory. *ACM DUX*, 1-15.
- [44] J. Rode. 2011. A theoretical agenda for feminist HCI. *Interacting with Computers* 23, 5: 393-400.
- [45] L. Rosenauer. 2016. Closing the STEM gender gap. *Corpus Christi Caller Times*. Retrieved Sept. 17, 2016 <http://www.caller.com/opinion/forums/closing-the-stem-gender-gap-3c8e68ce-2934-264e-e053-0100007f4d05-393756441.html>
- [46] A. Sears. 1997. Heuristic walkthroughs: Finding the problems without the noise. *Int. J. Human-Computer Interaction* 9, 3: 213-234.
- [47] A. Singh, V. Bhaduria, A. Jain, A. Gurung. 2013. Role of gender, self-efficacy, anxiety and testing formats in learning spreadsheets. *Computers in Human Behavior* 29, 3: 739-746.
- [48] R. Spencer. 2000. The streamlined cognitive walkthrough method, working around social constraints encountered in a software development company. *ACM CHI*, 353-359.
- [49] N. Statt. 2015. Microsoft CEO Nadella wants to help the world 'to achieve more'. *CNET*. Retrieved Aug. 24, 2016 <http://www.cnet.com/news/microsoft-ceo-satya-nadella-outlines-new-mission-statement/>
- [50] E. Stringer. 2007. *Action Research*. Sage, Newbury Park, CA.
- [51] C. Wharton, J. Rieman, C. Lewis, P. Polson. 1994. The cognitive walkthrough method: A practitioner's guide. *Usability Inspection Methods*. Wiley, NY, 105-140.
- [52] G. Williams. 2014. Are you sure your software is gender-neutral? *Interactions* 21, 1: 36-39.