Understanding Gender Equity in Author Order Assignment

KIRSTIN EARLY*, Yahoo Research, USA
JESSICA HAMMER, Carnegie Mellon University, USA
MEGAN HOFMANN, Carnegie Mellon University, USA
JENNIFER A. RODE, University College London, United Kingdom
ANNA WONG, Carnegie Mellon University, USA
JENNIFER MANKOFF, University of Washington, USA

Academic success and promotion are heavily influenced by publication record. In many fields, including computer science, multi-author papers are the norm. Evidence from other fields shows that norms for ordering author names can influence the assignment of credit. We interviewed 38 students and faculty in human-computer interaction (HCI) and machine learning (ML) at two institutions to determine factors related to assignment of author order in collaborative publication in the field of computer science. Interview outcomes informed metrics for our bibliometric analysis of gender and collaboration in papers published between 1996 and 2016 in three top HCI and ML conferences. Based on our findings, we make recommendations for assignment of credit in multi-author papers and interpretation of author order, particularly with respect to how these factors affect women.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**;

Keywords: Gender; Publishing; Collaboration.

ACM Reference Format:
Kirstin Early, Jessica Hammer, Megan Hofmann, Jennifer A. Rode, Anna Wong, and Jennifer Mankoff. 2018. Understanding Gender Equity in Author Order Assignment. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 2, CSCW, Article 46 (November 2018). ACM, New York, NY. 21 pages. https://doi.org/10.1145/3274315

1 INTRODUCTION

Collaboration is an important aspect of academic performance, particularly in multi-disciplinary fields such as human-computer interaction (HCI). Yet, evidence suggests that credit attribution for collaborative publication can introduce bias into the evaluation process for academic success [66]. Specifically, some collaborators may not receive appropriate credit for their work [31, 32], because

*Work done while the author was a student at Carnegie Mellon University.

Authors' addresses: Kirstin Early, Yahoo Research, 701 First Avenue, Sunnyvale, CA, 94089, USA, kearly@oath.com; Jessica Hammer, Carnegie Mellon University, Human-Computer Interaction Institute and Entertainment Technology Center, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA, hammerj@cs.cmu.edu; Megan Hofmann, Carnegie Mellon University, Human-Computer Interaction Institute, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA, meganh@cs.cmu.edu; Jennifer A. Rode, University College London, Knowledge Lab, Gower Street, London, United Kingdom, jen@acm.org; Anna Wong, Carnegie Mellon University, Department of Statistics, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA, graceanw@andrew.cmu.edu; Jennifer Mankoff, University of Washington, Paul G. Allen School of Computer Science & Engineering, 185 E Stevens Way NE, Seattle, WA, 98195, USA, jmankoff@cs.washington.edu.

2573-0142/2018/11-ART46 $15.00
https://doi.org/10.1145/3274315

authorship, and authorship order, are ambiguous reflections of effort [65]. Individual fields have authorship norms that address issues of credit, but often they are not documented. In addition, these norms lack finesse in reflecting degree of effort. When norms are not codified, they are influenced by a potentially problematic negotiation process about author order at submission time and are open to biased interpretations by readers. For example, a study of the impact of publication on tenure in economics found that men benefited more from multi-gender collaborations than women because author order was alphabetical, presumably leaving the assignation of credit to reader interpretation [66].

Clearly, author order, gender, credit, and academic success are linked in complex ways. This paper's contributions are to: (1) document challenges in negotiating author order in two subfields of computer science (CS) with varying presence of women, HCI and machine learning (ML), and (2) explore population-level differences in author order presumed to reflect trends at negotiation time. For the first goal, we interviewed 38 academic researchers about the process they used to determine author order. For the second, we performed a large-scale bibliometric analysis of 7376 authors, in which we quantified gender differences in co-authorship and author order outcomes. To our knowledge, ours is the first research to study the relationship between gender and author order in computer science publications.

Our qualitative interview results show that *author order does not necessarily reflect what authors think fairly represents their own contributions to a paper.* From a quantitative perspective, though publishing rates are similar for women and men in our dataset, the distributions of author compositions and positions are not. For example, significantly more multi-author UIST papers have no female authors (63%) than chance would predict (55%). Our predictive analysis shows that *being female can reverse the impact of multiple factors on likelihood of first, last, and middle author position compared to the population as a whole.* For example, when an author has many lower-ranked co-authors, that author is more likely to be last author. In CS publications, this prestigious last-author slot indicates intellectual leadership of a project. However, if a *female* author has many co-authors of lower rank, she is less likely to be last author. When interpreting findings from our two studies, we frequently use theory derived from prior published studies to explain what we observe. We summarize these prior studies in the Related Work section that follows.

We recommend strategies for researchers to make more inclusive decisions about author order. We also suggest ways we can, as a field, support a more effective author assignment process.

2 RELATED WORK

This section first briefly describes how gender affects interpersonal relationships and power dynamics. We then examine the impact of gender in academia. *Gender is a culturally defined, fluid concept* that affects the actions and expectations of individuals and communities [19] and is embedded in structural effects and organizational processes, such as power relationships [48]. Additionally, *gender is a process* that is performed daily, reaffirmed in conjunction with gender norms that are being socially reconstructed in response to changing social practices [84]. Thus, studies of gender look not at biological differences but at the impact of gender across personal, interpersonal, structural, and cultural contexts. Individuals may identify outside the gender binary [60], but their actions are performed "at the risk of gender assessment," typically either female or male [84].

Gender is not the only axis along which discrimination and bias occur: people of any marginalized group—whether of gender, race, sexual orientation, disability, socioeconomic class, *etc.*—are disadvantaged due to power differences, social codes, and implicit biases. In contrast, members of the dominant group do not face this systemic disadvantage and are treated as a "default" in society. This position of benefit for the dominant group is referred to as *privilege* [52]. People with multiple marginalized identities experience bias in compounded ways, meaning that multiple social power

structures, not only gender, may contribute to a woman's disadvantage [21, 55]. This paper focuses on gender alone, but we recognize the importance of future work on these overlapping spheres of difference and how they impact women's lived experiences in the academic realm [90].

2.1 Gender and power differences

Gender roles specify modes of interacting, and children are socialized into these roles [57]. These roles privilege men (*e.g.*, [42, 57]). For example, prized leadership traits are seen as "male" qualities (*e.g.*, management ability) [62]. When women lead in traditionally "feminine" ways (*e.g.*, emphasizing collective work), they are not seen as effective. However, when they lead in "masculine" ways (*e.g.*, emphasizing independence), they are also penalized [62]. These dynamics may affect authorship decisions. For example, a "feminine" leadership style, emphasizing collaborative work, may lead co-authors to devalue a woman's contribution and place her in a less prestigious author position [31, 32].

While overt sexism is declining, subtle biases persist, disadvantaging women more insidiously because well-intentioned men may not realize the hierarchy they perpetuate (*e.g.*, [76]). Implicit bias is often shown through *microaggressions*, "brief and commonplace daily verbal, behavioral, and environmental indignities, whether intentional or unintentional, that communicate hostile, derogatory, or negative...slights and insults" [76]. These seemingly small acts have real world consequences—for example, general statements about stereotypes can hamper individual performance on exams [70, 72].

Gendered power dynamics affect all social interactions, including academic collaboration and authorship negotiation. This paper focuses on gendered outcomes of author order.

2.2 Gender bias in academia

As in many institutions, evidence suggests that gender bias exists in academia [37, 47, 83]. For example, multiple studies have demonstrated bias in assessment of resumes based on presumed gender. In one recent (2012) double-blind study, 127 faculty in biology and physics each reviewed a resume with a randomly assigned female or male name. Men were rated as being more competent, hireable, and worthy of mentoring and higher pay than women [53]. Faculty gender, seniority, field, and age had no effect on outcome. An earlier (1999) psychology study of 238 faculty who received four gender-randomized resumes showed similar outcomes [73].

The impact of small differences in evaluation of female candidates of 1-5% (a purported degree to which gender bias affects performance ratings [4]) can be used to simulate how discrimination will impact gender distributions over time in an organization [50] as well as productivity [20]. In simulation, such differences accumulate, leading to approximately a 2:1 difference in measures of success at the most senior levels [20, 50]. Disadvantages due to systemic bias accumulate and affect assessment of merit, grant proposal acceptance, and assessment of a publication's authority [88] as well as promotion, pay, publication rate [9, 51, 63, 69, 78], and longevity, particularly going from Ph.D. to faculty [67]. Racial and ethnic minorities face more issues, and these are compounded with gender discrimination for women [86].

2.2.1 Gender bias extends to authorship. Among differences that affect academic success, publication and credit assigned for publishing are salient given this paper's focus and the fact that authorship is a significant factor in promotability.

Much research has focused on differences in the number of publications and citations, with women having fewer of both, on average, than their male counterparts (*e.g.*, [18, 28, 41, 43, 68, 71, 85, 89]). Hengel [34] identified one factor that can deter women's publishing. The review cycle for papers by women takes, on average, six months longer to complete than for men *Econometrica* despite the papers being more readable—measured using five separate scales, such as the Flesch Reading

Ease scale—than those authored by men. The simplest explanation is that reviewers hold women's papers to a higher standard than men's [34] in venues without double-blind review, such as many computer science journals.

There are also differences in author position for women and men. For example, West *et al.* [85] determined gender for 73% of authors in JSTOR. They examined overall gender frequencies and gender frequencies by author position (first, second,..., last). Results varied by field, but first authorship roughly matched the fraction of women in a field (increasing over time), while last authorship lagged behind: female authors are typically in less prestigious positions than male authors [85]. Similarly, Pierson's study of the gender gap in arXiv [59] found that "women are more likely to be first authors in fields in which they are better represented...even controlling for how many authors on the paper are women."

These studies focus solely on women's publishing outcomes. Considering the assignment and interpretation of credit to women and men in jointly authored papers reveals further biases women must overcome (*e.g.*, [34, 66]).

2.2.2 Benefits and challenges of multi-author papers. Modern science in general, and CS in particular, are collaborative fields, and papers often have multiple authors. Collaboration has been shown to improve productivity (*e.g.*, [23, 29, 40, 41]) and work quality [74]. For example, collaborative authors are more productive even when adjusting for number of authors to account for reduced individual contribution [23].

A group's gender composition can mediate the benefits of collaboration for women. For example, in a study of undergraduate engineers, women's verbal participation was highest in majority-female groups, and women's confidence and career aspiration were lowest in majority-male groups [22]. While these effects have not been studied in groups of authors, teams of male faculty with male graduate students produce more papers than other gender compositions [29]. Also, non-collaborating women produce less than non-collaborating men [41]. Though collaboration benefits men and women, women who collaborate still produce less than men who collaborate [29, 41].

2.2.3 Credit and authorship. In academia, authorship translates collaborative work into professional credit. Yet, in collaborations, the relation between effort, authorship, and credit remains opaque.

Despite the fact that research is usually a group effort, credit for a paper is often assigned to one person. This can then affect hiring, promotions, other researchers' esteem, *etc.*, greatly affecting individuals' careers. Holden *et al.* [35] propose crediting a researcher with only a fraction of a multi-author paper when making promotion decisions. They also highlight the impact of *how credit is assigned in collaborative efforts* on an author's career.

The process of translating authorship information to professional credit is subject to bias. For example, in a study of the impact of papers on the careers of economists, mixed-gender, co-authored papers yielded tenure gains for men but not women [66]. One possible explanation is that author order in economics is alphabetical, so the reader decides who deserves credit for published work. This highlights two important issues. First, *perceptions by readers of the contribution of authors in multi-author papers may be subject to bias*, even if author order is not alphabetical. Second, when collaborators decide author order, *cultural assumptions may significantly affect the assignment of credit in a process that may be subject to gender bias.*

2.2.4 Determining author order. While we have established the reader's potential for bias in assigning credit, we have yet to address the degree of author bias in doing so, or, more broadly, how author order is established in an academic context.

Author order and citation style can both impact credit assignment: the former often reflects who did the work, and the latter affects whose name is linked to a result. Common citation approaches are numerical ([n]), abbreviation (part of first author's name and year), and APA (one or several names and year) [58]. Recognizing that citation style affects recognizability, the Conference on Human Factors in Computing Systems (CHI) includes first names in references [39].

Overall, author order is decided in subfield-specific ways that reflect a combination of the culture of the publishing community and the training and personal preferences of co-authors. In some fields, including most of CS, author order typically represents contribution. However, the meaning of author positions is not consistent across publications; *e.g.*, the second author may be a major contributor or, at other times, more peripheral. Therefore, readers cannot always deduce the contributions of each author [11]. Some researchers have proposed making author contributions explicit, such as through interactive interfaces [6] or contribution disclosures [30, 80].

Regardless of field norms, author order is best understood as *a negotiated process* among authors. Past work has shown that if criteria are not defined upfront, women are disadvantaged [82]. Women may be unable to advocate for themselves because of social costs for speaking up. For example, they may be penalized for initiating negotiations, especially by men [13]. Furthermore, contribution types are *gendered*. For example, recommendation letters emphasize men's research abilities and agency but women's education and contribution to community [47, 79]. However, research ability is more valued, especially in male-dominated fields [45, 47]. Thus, women's contributions may be presented as and deemed to be less worthy than men's, which could result in women receiving less prestigious author slots.

2.3 Summary

Gender influences authorship in multiple ways. These differences may be partly driven by specific features of the authorship negotiation process. For example, if co-authors do not discuss author order, structural effects about the assumptions of gendered contributions can come into play. If co-authors do discuss author order, gendered power dynamics can affect negotiation. However, the meaning of author order and the process by which it is established are still opaque, which we work to clarify with qualitative interviews. In addition, the impact of women's representation in papers has not been documented in HCI, a gap addressed in our bibliometric analysis.

Discussion of sex and gender is complex. Most prior work on authorship is cisnormative—*i.e.*, it assumes gender identity matches physical characteristics and ignores trans and genderqueer identities). We asked interviewees to provide their gender free-form, and all identified as female or male. For our bibliometric study, we asked Mechanical Turk workers to identify an author's self-identified gender (female/male/other), not sex. We instructed workers to look for gender pronouns in articles about the author, if possible. This paper uses the terms "female" and "male," as opposed to "feminine" or "masculine." We do so because, as queer scholars and scientists argue, both sex and gender are socially constructed [3, 16]. Thus, neither set of terminology is inherently more "accurate," so we use the terminology found in prior work for ease of comparison.

3 INTERVIEWING ACADEMICS ABOUT AUTHOR ORDER

Computer science is an increasingly collaborative discipline, and most work is done in teams [25]. This paper seeks to characterize the nature of collaborations in producing papers.

To interpret the meaning of positions in author order, we first interviewed 38 researchers (16 women, 22 men) at two U.S. research universities. These researchers were at various stages of their careers, from Ph.D. students to senior professors. About two-thirds of the interviewees were from the field of HCI, while the remainder were from ML. Two authors of this paper participated as interviewees: there is value in authors' engaging in reflexive science [15] and bringing personal

PID	Rank	Gender	Field	Univ.	PID	Rank	Gender	Field	Univ.
StudF1	Ph.D. student, year 4	F	ML	U1	StudM20	Ph.D. student, year 2	M	HCI	U2
StudF2	Ph.D. student, year 2	F	ML	U1	StudM21	Ph.D. student, year 3	M	HCI	U1
StudM3	Ph.D. student, year 2	M	ML	U1	StudM22	Ph.D. student, year 4	M	HCI	U2
StudM4	Ph.D. student, year 4	M	ML	U1	StudF23	Ph.D. student, year 3	F	ML	U2
StudM5	Ph.D. student, year 5+	M	ML	U1	PostF1	Postdoc	F	ML	U1
StudM6	Ph.D. student, year 4	M	ML	U1	PostM2	Postdoc	M	ML	U1
StudM7	Ph.D. student, year 4	M	ML	U1	PostM3	Postdoc	M	HCI	U1
StudF8	Ph.D. student, year 3	F	HCI	U1	PostF4	Postdoc	F	HCI	U1
StudF9	Ph.D. student, year 3	F	HCI	U1	JFacF1	Assistant professor	F	HCI	U1
StudM10	Ph.D. student, year 5+	M	HCI	U1	JFacF2	Assistant professor	F	HCI	U1
StudF11	Ph.D. student, year 3	F	HCI	U1	JFacM3	Assistant professor	M	ML	U1
StudF12	Ph.D. student, year 2	F	HCI	U1	JFacM4	Assistant professor	M	HCI	U1
StudM13	Ph.D. student, year 2	M	HCI	U1	JFacF5	Assistant professor	F	ML	U2
StudM14	Ph.D. student, year 2	M	HCI	U2	JFacM6	Asst. teaching prof.	M	ML	U2
StudF15	Ph.D. student, year 5+	F	HCI	U2	SFacF1	Associate professor	F	HCI	U1
StudM16	Ph.D. student, year 5+	M	HCI	U2	SFacM2	Full professor	M	ML	U1
StudM17	Ph.D. student, year 3	M	HCI	U1	SFacM3	Full professor	M	HCI	U1
StudF18	Ph.D. student, year 4	F	HCI	U2	SFacF4	Associate professor	F	HCI	U2
StudM19	Ph.D. student, year 3	M	HCI	U1	SFacM5	Associate professor	M	ML	U2

Table 1. IDs indicate rank ("Stud" for Ph.D. student, "Post" for postdoc, and "J/S Fac" for junior/senior faculty) and gender. Seven of the 27 students and postdocs had advisors who were interviewed. To protect anonymity, we do not disclose these relationships. Students published less than half of their papers with these advisors.

experiences to research [61]. The authors were interviewed in the same way as other participants, which enabled standardized incorporation of their experiences. Two authors interviewed participants. Table 1 lists interviewees' rank, gender, field and university affiliation. To protect anonymity, all participants are not identified further.

Interviews lasted thirty minutes and included questions like *"What factors do you consider when deciding author order?"* and *"How do you discuss author order with your collaborators?"* The full list of interview questions is in Appendix A. Some prompts asked about specific papers: *e.g.*, *"Identify a paper where you were not happy about the author order and discuss it."* Participants were given lists of their publications from the bibliography database DBLP [46] but were not restricted to that list. We did not directly ask about gender because we wanted to hear lived experiences, not beliefs about the influence of gender. We paid close attention to gender-related issues and the relationship between participant gender and topics.

Two authors conducted interviews and anonymized all individuals. Participants were identified by the authors, recruited over email, and not compensated. We analyzed the data by coding emerging themes. We reached consensus on the themes through discussion among the authors. The resulting themes focused on norms for author order, interpretation of author position, strategies for determining order, experiences discussing order, dissatisfaction with author order, managing co-authorship and order in long-term collaborations, and how participants learned to determine author order. After presenting these themes, we discuss differences across genders and institutions.

3.1 Interview results on authorship

Our 38 participants revealed that author order can be determined by rules, negotiation, and randomness. Five specifically called out the process as being "really/very complicated" (StudM4, StudM5), "thorny" (StudM19), "not always easy to determine" (SFacM2), and "really *ad hoc*" (SFacM3). Despite this complexity, some common themes emerged.

3.1.1 Norms for ordering authors in multi-author publications. All participants were currently in academia, and a few had prior industry experience. Academic collaborations typically involve at least one student and their advisor(s). Typically, the student is first author, and the advisor the last.

Contribution	Mentioned by
Work	*Stud*: F2, M4, M5, M6, F8, F9, F11, F12, M13, M14, F15, M16, M17, M19, M20, M21, M22, F23; *Post*: M2, F4; *JFac*: F1, M4, F5; *SFac*: F1, M2, F4, F5 "My fundamental philosophy: authorship should reward work." -JFacF1
Ideas	*Stud*: F1, F2, M3, M4, M5, M6, F9, F11, F12, F15, M17, M20, M22, M16; *Post*: F1, M3; *JFac*: F2, M4; *SFac*: F1, M2, M3, F4, M5 "First is idea." -StudF2 "Work is more important than idea generation." -StudF1
Study design	*Stud*: F2, M7, F8, F9, F15, F18, M21, M22; *Post*: M2, M3; *JFac*: F1, F2
Implementation	*Stud*: F1, M3, M7, F11, F12, M13, M22, M16; *SFac*: M3
Data collection	*Stud*: F1, F2, M7, F8, M16, M21; *Post*: M2; *JFac*: F2, M3; *SFac*: M2 "Anyone who collects data gets their name on the paper because there's so much work involved." -PostM2 "I found it a little weird that I've never met or talked to this [data] person, yet I have to add him." -StudM7
Analyses	*Stud*: F1, F2, M6, F9, F11, F12, F15, M16, F18, M21; *Post*: M3; *JFac*: F2; *SFac*: M2
Writing	*Stud*: F1, F2, M3, M5, M6, M7, F8, F9, F11, F12, F15, M16, M17, F18, M21; *Post*: M2; *JFac*: F1, F2; *SFac*: F1, M2, M3, F4, M5 "I think the key thing is writing the paper." -SFacM3
Organization	*Post*: F1; *JFac*: F1, M3 "Managing process, which is work that often goes uncounted." -JFacF1
Funding	*Stud*: M7, F11; *JFac*: F2; *SFac*: F1, M3, F4, M5 "I wouldn't weigh that very highly—if somebody's putting funding in and not doing [work], there's a problem." -JFacF2

Table 2. Aspects of contribution mentioned by interviewees.

Publications can also arise from group projects in classes, where author order is often difficult to determine since all students did similar amounts of work. Participants who mentioned papers from course projects were usually unhappy about author order for this reason (StudF2, StudM3, StudM5).

Academia-industry collaborations often involve a student intern, an industry advisor, and occasionally a faculty advisor. SFacF1 stated that it is common for the student to be first and the industry advisor to be last. Two people (StudM3, StudM5) previously published in industry. They reported a similar advisee-advisor relationship as in academia, with advisees being first authors.

3.1.2 Meaning of author positions. Table 2 summarizes aspects of contribution mentioned by participants. Most thought first author should have done the most work. There were more options for the meaning of last author, but 27 people said last position is for the lead advisor or principal investigator (PI): "last author...provided overall intellectual and practical resources" (JFacF1). Others considered last position random or unimportant (two faculty); the least-contributor, regardless of seniority (six students); or the most senior, regardless of contribution (six students). Overall, 31 of the 38 interviewees considered last author to be for the advisor or most senior author (some participants gave multiple answers for the meaning of last author). Ten said middle positions are unimportant: "If you're not first and you're not last...it doesn't matter that much" (JFacF1).

3.1.3 Determining author order. A very common pattern used for author order is for a Ph.D. student to be first author, and their advisor last author. SFacM3 said: "To be first author, you have to write most of the paper. But we encourage the most appropriate person to do the writing." Here, his comment about the most appropriate person is referencing his belief that sometimes the appropriate person (by which he meant the Ph.D. student) needs to be nudged and guided to do the things that merit first authorship, meaning this is not simply a negotiation.

A further complicating factor is the increasingly common case where papers have multiple student authors. In these cases, everyone agreed that "contribution" should play a large role in determining author order. Aspects of contribution mentioned include: amount of work, ideas, study design and methods, implementation, data collection, analyses, writing, organization, and funding. While participants could easily identify types of contribution, it was difficult to assign values to each since collaborators often make multiple contributions to a project.

A few participants believed that contributions were mandated by tradition and should not determine authorship. Several students questioned why advisors who fund work but do not contribute

intellectually should be authors ("If someone doesn't help, even if they funded the research, they shouldn't be on the paper...but that's not realistic. The PI is always going to be on it." (StudF1)).

Expectations for what merits authorship vary by field. For example, some ML interviewees who work mainly with hard-to-acquire data did not question giving authorship to someone whose sole contribution was data collection. Others thought data providers should be recognized in the acknowledgements, rather than as authors. Authors who work with such datasets (usually biological data) explained that when someone runs a lab whose primary goal is to collect data, they do not have time to write papers with those data. Thus, they receive author slots on papers using their data so they can get recognition and continue receiving grants to fund their data collection work.

Most faculty (eight) said they consider the impact on students' careers when determining author order since professors are already established: "Whenever any of my Ph.D. students are primarily involved in the project, they always go first, independent of contribution" (SFacM3). Several students said their advisors told them students always come before faculty in author lists. However, five faculty and three students had stories of how, when they were students or postdocs, an advisor or senior person took the first author slot when the student did the vast majority of the work. As students, they agreed to this because they "didn't have a say, and he was more senior" (JFacM3), "figured if it wasn't that way, it wouldn't get done" (SFacM2), "did not have the power to say otherwise" (JFacF1), or felt their advisor "had power; I didn't want to upset her" (StudM16).

3.1.4 Discussing author order with co-authors. Seven faculty said they initiate discussions of author order with students to correct inappropriate orderings, to teach them about the process, and to lay out expectations. Many students and postdocs reported minimal discussion of author order with all their collaborators "because it's usually so obvious" (StudM10) or because it is uncomfortable. When students do negotiate order, it is often with other students for first author position.

Faculty who initiate authorship conversations often do so early, before writing starts: "It's best to sort this stuff out in advance unless you don't care" (SFacM2). Other researchers discuss order while writing once they have an idea of relative contributions: "Sometimes it's not clear a paper's going to come out...it seems awkward to [discuss order] before anyone's done any work" (JFacF2). Some wait until submission: "After the work's done...you've seen everyone's contributions" (StudF2).

Interviewees reported considering their collaborators' feelings when proposing an author order (SFacM3) and accepting an authorship they might not want (PostM2) to avoid uncomfortable conversations. Discussion can break down when a collaborator makes demands that others find unreasonable but often accept to avoid or end an argument: "There wasn't much of a discussion because it got unpleasant very quickly so we just stopped" (JFacF2). People said they expect collaborators to be reasonable: "I've always been working with very reasonable people" (StudM4); "It's based on reasonable behavior" (StudM5). In some cases, they considered behavior in choosing collaborators: "No one's going to fight over fourth author. I tend not to work with those people" (PostM2); "I only want to work with people where I can really feel like we're on the same side...it's not either adversarial or exploitative" (JFacF1).

3.1.5 Dissatisfaction in author order. Most participants (26 of 38) could cite at least one paper where they were unhappy with the author order. Three categories of dissatisfaction characterized these cases. The most prevalent (14 instances from 14 authors) was due to the culture of authorship and lack of discussion—certain people were expected to be authors, even if they had not contributed much in interviewees' eyes or someone made the decision without consulting the interviewee. The next most prevalent (10 instances from 9 authors) occurred when a powerful co-author insisted on a prestigious slot. Finally, other cases (4 instances from 4 authors) happened when author order was ambiguous (often when people made similar levels of contribution). People with less experience

authoring papers were less likely to have had a bad experience, perhaps because problematic authorship does not happen on *most papers*, rather it occurs in *most careers*.

3.1.6 Co-authoring with long-term collaborators. When working with others on multi-publication projects, researchers can share prestigious spots over time. Some trade positions, while others keep the same relative author positions for all papers. Several interviewees mentioned making agreements with long-term collaborators to trade off the prestigious slot (StudM7, StudF9, StudF11, StudF12, StudM19, JFacF1, JFacM3, SFacF1, SFacM3), either in strict alternation (StudM7, StudF9, StudM19, JFacF1) or depending on who benefits most in each venue (JFacF1, SFacM3). Students are interested in trading first position, while faculty want to trade last position. In other cases, co-authors kept the same relative positions for all or most papers, due to a commonly advised student placing them in the same order (StudM7, SFacF1) or a preference of one co-advisor (JFacF2).

3.1.7 Learning about author order. In learning this complex process of author ordering, only three interviewees had any formal training: StudF1 took an authorship ethics class, JFacF2 received a rubric for authorship and ordering from a faculty mentor, and SFacF1 read papers on authorship. Informal guidance came from advisors (11 participants), discussion with peers (10 participants), observation (reading papers: 7 participants), and experience (*i.e.*, writing papers: 16 participants). StudM19 stated he never learned the meaning of author order. JFacM6 questioned, "[author order is] so straightforward. Is there anything to learn?"

3.2 Cultural differences in author order

There were some differences in the conceptualization of author order in groups of interviewees, with different cultures around authorship according to field, institution, or geography. Eight of the 15 ML researchers brought up "joint first author," where they note on a paper if any co-authors contributed equally. Since only one name can be listed first, people often paired joint first author with slot-swapping for multi-paper projects. No one in HCI mentioned joint first author.

Interviewees at the two institutions did not differ much in how they decided author order. Two Ph.D. students at U1 stated that their advisors make the decisions about author order (StudF8, StudM19), as did three at U2 (StudM14, StudM22, StudF23). The other 14 Ph.D. students at U1 and seven at U2 said they take an active role in determining author order.

Finally, while all participants currently work in the U.S., several held prior positions outside the U.S. and "relearned" the meaning of author order when they moved. PostF4 did her Ph.D. in Europe, where she said author order was not considered important, so she did not make sure she held positions that reflected her contribution to papers. Now that she has learned, she follows the prevalent U.S. author order model. Three Ph.D. students who did their undergraduate degrees internationally (StudF18, StudM20, StudM22) had had research advisors who followed a decreasing-seniority author order model. Now Ph.D. students, they follow the student-first, advisor-last model.

3.3 Differences associated with interviewee gender

Because determining and discussing authorship and author order is a social process, it is an arena in which gender is performed and influences interactions [84]. It is risky to draw population-wide conclusions about gender from our small number of interviewees (16 women, 22 men) and venues, but some differences did emerge. We saw that men conceptualized the author order process differently than women. Only men: claimed that people do not care about author order (StudM7, PostM2, JFacM3, JFacM4, SFacM3); called out the process as "really/very complicated" (StudM4, StudM5), "not always easy to determine" (SFacM2), "really *ad hoc*" (SFacM3), or "thorny" (StudM19); said it is better to err on the side of inclusivity (StudM4, StudM6, JFacM4); mentioned not wanting to

hurt people's feelings in negotiation (PostM2, SFacM2); and said they trust collaborators regarding authorship decisions (StudM6, SFacM3).

Women tended to discuss author order earlier than men. Five women (one student; one postdoc; three faculty) and four men (two students; two faculty) reported discussing author order very early, before writing begins (56% women compared to 42% female interview population). Nine women (six students, one postdoc, and two faculty) and six men (four students; two postdocs) discussed author order while writing the draft (60% women). Two women (students) and eight men (six students; two faculty) waited until immediately before submission to discuss author order (20% women). Only men (two students; two faculty) said they did not discuss author order until it was time to submit the camera-ready version of an accepted paper. In other words, 31% of women interviewed discussed author order very early, while only 18% of men discussed it very early; 56% of women discussed it early, compared to 27% of men; 13% of women discussed it late, compared to 36% of men; and no women discussed it very late, compared to 18% of men.

3.4 Discussion

Our 38 interviewees shed light on the complex process of author order. We discovered that people attach meaning, such as prestige or indication of amount of work, to certain author slots. Almost all participants agreed that the first author did or *should* have done the most work and writing, and the majority indicated that the last author was usually the lead advisor. In some cases, authors attached meaning to middle positions (particularly second author), but overall they viewed middle positions as unimportant. Participants who published outside the U.S. shared that norms were different at their previous institutions, and they relearned the meaning of author order in the U.S.

We found that people have different approaches to negotiating and deciding author order. Some broached the topic early, and some avoided what might be an uncomfortable discussion. Outcomes of negotiations (or lack thereof) did not always match participants' notions of fairness. For example, some people mentioned explicit arrangements for slot-swapping in long-term, multi-paper collaborations, and others expressed frustration over being consistently relegated to middle-author positions when co-authoring with another person. Individuals also had different ideas of what constitutes contribution and how much each type of contribution (*e.g.*, work, ideas) should count. This lack of consensus may result from the fact that very few interviewees received formal training in authorship and the meaning of author order. Past work has shown that when criteria are not specified upfront, women are disadvantaged [82], so the lack of formal training is likely not to work in women's favor.

A few trends arose along gender lines. Only men claimed people do not care about author order, the process is complicated, they trust their co-authors to be reasonable, and they try not to hurt co-authors' feelings when negotiating. These sentiments suggest men may be more likely than women to disregard author order (it is complicated, but nobody cares too much about it) and focus on relationships (not wanting to hurt feelings and trusting their co-authors to be kind to them).

Furthermore, men tended to determine author order later than women, with 87% of women and 45% of men discussing author order with their collaborators early or very early, and 13% of women and 55% of men discussing author order late or very late. These differences in discussion time corroborate the theory that men may be more likely than women to disregard author order. Perhaps men are less concerned about author order than women because they lack the lived experience of being a marginalized gender among their co-authors in situations of potential implicit bias. An alternative explanation is that women are more comfortable initiating such discussions. Studies have found that men are more likely to avoid conflict, while women are more likely to use collaborative conflict resolution, which would be consistent with this interpretation (*e.g.*, [14]).

Although our data evidence gender-specific trends, no one raised gender as a factor in authorship. The omission is not surprising for several reasons. First, the gendered power dynamic is so ingrained in conventional wisdom that it may become unquestioned [36, 42]. Further, women often evaluate fairness in relation to a perceived norm; *e.g.*, women determine the equity of division of household labor by comparing their time doing household labor to the time other women spend rather than an equitable division [75]). If women perceive receiving middle author positions as a norm for collaborative projects, this could explain why they did not cite more cases of dissatisfaction in authorship than men. Second, minorities may worry about being penalized for calling out sexism and racism [33, 38, 77], making them less likely to mention these effects.

Furthermore, when instances of prejudice are ambiguous, as they often are [24], recipients may experience cognitive burden, so they rationalize it away [64]. For example, when interviewees mentioned power differences and unfair outcomes, they brought up the power structure of academia—not gender, race, or any other differential that is not supposed to influence outcomes. However, our data show that the academic power structure does not equally represent people: seven of the eight instances of senior faculty inappropriately taking prestigious positions involved *male* senior faculty. Yet, interviewees with this experience (women and men) mentioned only seniority, not gender.

4 BIBLIOMETRIC ANALYSIS

Our interviews provided a qualitative look at the process behind author order. In them, we observed some differences in the conceptualization of author assignment along gender lines. These differences are important because past work has shown that even small biases in random processes can lead to notable gaps over time [50]. Therefore, we might expect small differences in how women and men determine authorship to have a quantifiable impact in a larger population. As a result, we conducted a bibliometric analysis of three HCI and ML conferences with different percentages of women, based on work showing that the percentage of women in a field impacts their success [22, 59].

4.1 Sample

We used the DBLP database [46], restricted to conferences from subfields in our interviews (ML, HCI) so we could interpret author order with respect to subfield norms. We further restricted analysis to three conferences with varying female presence: CSCW (35%), UIST (15%), and ICML (11%). This also limited the number of authors, which was important for our labor-intensive gender disambiguation. About 20% of names had unknown gender, which could bias results if unaddressed. Our final sample had gender for 97.1% of authors at 95.5% accuracy for ambiguous names.

Author disambiguation. Authors could be ambiguous in two ways. First, multiple people with the same name could be identified as one person. DBLP reduces these errors by appending a numeric identifier to authors sharing a name and describes its process as "at least as accurate as the data provided by the publishers."[1] In addition, the possibility of collisions is small within a subfield since authors who publish at the same conferences would probably disambiguate their own names, *e.g.*, by using middle initials. Second, ambiguity could occur if an author publishes under multiple names. With over 7000 authors in our dataset, manually inspecting all possibilities is challenging. Instead, we eliminated gender-specific author splitting by inspecting all duplicate first names among women with at least two papers to confirm all were distinct authors. Thus, we are confident disambiguation issues are not biasing our results due to gender differences.

Dataset limitations. DBLP lacks information that may also be relevant to measuring author impact: citations, patents, group size, gender composition of program committees, *etc.*

[1] https://dblp.org/faq/How+accurate+is+the+data+in+dblp

4.2 Metrics

For each paper, we extracted year, conference, and ordered author list to assign author position [solo, first, middle, last] from DBLP. We also needed gender and seniority, which were not in DBLP.

Gender. We used name as a proxy for gender. For most authors, we assigned gender according to first (or middle) names using the open-source Python package, SexMachine[2] [27]. 1356 authors had "androgynous" names, *i.e.*, ones that could not be resolved as "female" or "male." Some were truly ambiguous, like Casey, while others did not appear in the list of names from 55 groups of countries. We inspected these and marked ones we recognized from our knowledge of the field.

To confirm the accuracy of the automatically assigned genders, we manually checked 50 random names assigned as female or male. Of these 50 authors, 48 were online (*e.g.*, had personal websites), and their genders matched the name-inferred output. The two not found had names associated with the predicted genders. Thus, we are confident in these gender assignments.

For the 1356 authors of unknown gender, we asked workers on Mechanical Turk for evidence of gender given a link to an author's most recent publication. We instructed workers to use the paper to identify an author's institution and search for a photograph or text referring to them with gender pronouns. Workers provided URLs for evidence, a proposed gender, and a confidence rating. They were paid $1 for each requested name they looked up, regardless of confidence level; each name took around five minutes to look up. We checked 10% of the results, stratified by confidence. For workers who were somewhat, very, or extremely confident, accuracy was 95.5%, calculated by comparing workers' gender assignments to ours. Less confident suggestions were rejected, leaving authors androgynous. 1141 genders were resolved in this way. Androgynous authors are not shown in our analysis because they are so rare in our data (2.9%). Of the names resolved as female or male, the proportion of female names (18%) was similar to the whole dataset (20%), so the deletion of those authors should not introduce any further bias into our results.

Names cannot reliably indicate gender. Self-reported gender, which is information we do not have, is the only accurate way to determine authors' genders and could give more complete information than the binary assignments this name-as-proxy approach assumes. However, the impact of author order is determined by the reader, and most people still interpret gender as binary.

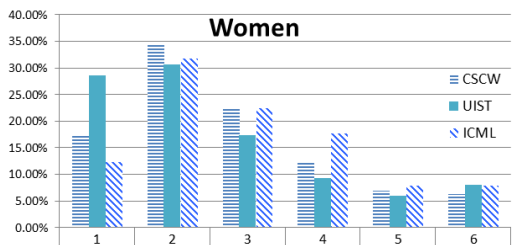
Seniority. We estimated seniority by looking at how long authors had been publishing. To improve accuracy, we included all DBLP publications. We defined an author's rank as 1 in their first year of publishing, 2 in years 2-5, 3 in years 6-10, 4 in years 11-15, 5 in years 16-20, and 6 after 20 years. This distinguishes between new, possibly one-time, authors (rank 1) and more established young authors (ranks 2-3). This method assigns everyone to rank 1 at their first DBLP paper. However, some authors may have already been senior when DBLP's dataset begins, perhaps publishing in other fields before shifting to CS. To address this, we removed the first ten years from analysis. This helps reduce seniority errors, since anyone publishing in the first of those ten years would be at least rank 3 at the start of our analysis. Our final data include years 1996 through 2016.

4.3 Results

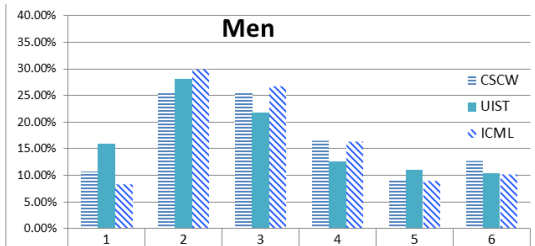
Our dataset has 4355 papers; 1614 (37%) of them have at least one female author. The only statistically significant difference in output is that women published fewer papers than men in UIST and ICML. However, this difference is not significant when restricted to authors with multiple papers (Table 3).

A common explanation for gender inequity is the pipeline [10, 17, 81], which would be reflected in relative ranks and differences in publication rates. For women and men in all conferences, the dropoff as rank increases is exponential (Figure 1).

[2]This name equates sex and gender and emphasizes sexuality, which can make women feel unwelcome in male-dominated areas. After receiving feedback, the creator of a Ruby version renamed it [54].



	1	2	3	4	5	6
CSCW	17.24%	34.65%	22.44%	12.39%	7.00%	6.28%
UIST	28.67%	30.67%	17.33%	9.33%	6.00%	8.00%
ICML	12.33%	31.72%	22.47%	17.62%	7.93%	7.93%



	1	2	3	4	5	6
CSCW	10.73%	25.55%	25.44%	16.48%	9.18%	12.61%
UIST	15.97%	28.14%	21.80%	12.67%	11.03%	10.39%
ICML	8.16%	29.87%	26.70%	16.31%	8.83%	10.13%

Fig. 1. Percentage of women (left) and men (right) at each rank. Rank 1 corresponds to the first year of publishing (across a larger sample of conferences) and each higher rank corresponds to the next 5 years of publishing. Results are restricted to 2010 to 2016 to illustrate the current landscape.

Conf.	Paper count	Auth. count	# of auth/paper	Average # of papers for each author				Average # of papers per year for each author			
				All		Multi-publishers		All		Multi-publishers	
				Women	Men	Women	Men	Women	Men	Women	Men
CSCW	1119	2354	3.41 (1.58)	1.65 (1.62)	1.62 (1.76)	3.44 (2.35)	3.46 (2.78)	0.97 (0.26)	0.96 (0.27)	0.88 (0.50)	0.85 (0.52)
UIST	675	1530	3.82 (1.82)	**1.39 (1.13)**	**1.75 (2.25)**	3.02 (1.85)	3.84 (3.65)	0.95 (0.23)	0.96 (0.26)	0.72 (0.45)	0.85 (0.48)
ICML	2561	3781	2.76 (1.13)	**1.62 (1.67)**	**1.93 (2.35)**	3.51 (2.57)	3.77 (3.37)	0.95 (0.27)	0.94 (0.31)	0.81 (0.52)	0.83 (0.52)

Table 3. Publication output for women and men. "All" includes all authors in a conference. "Multi-publishers" indicates authors with at least two papers in that conference. (Left) Basic statistics. (Middle) The average number of papers published. (Right) To adjust for the higher raw output of people who have published longer, we show authors' average paper count per year of their publication history. Standard deviations are in parentheses. Significant differences between women and men are shown in bold ($p < 0.05$). Authors in multiple conferences are included in multiple rows.

4.4 Gender composition of papers across conferences

Given that the percentage of women differs across conferences, we would expect the gender composition numbers for papers (*e.g.*, number of papers with zero women, number of papers with one woman) to differ, too. We quantified this using a binomial test to compare the number of papers with female authors to the expected number, if authors were selected randomly from the author pool in that conference. Figure 2 shows the percentage differences between the observed frequency of paper gender compositions and the expected frequency, for papers with at least two authors (only 7% of papers had only one author). For example, given the number of authors on UIST papers and UIST's percentage of women, we would expect 55% of UIST papers with at least two authors to have no women authors, if authors were selected randomly from the UIST author pool (15% women). However, 63% of these UIST papers have no women, which is significantly higher than chance and represents a 14% increase from the expected prevalence of UIST papers with no women.

All three conferences have significantly more papers with no women authors than expected (positive bars in the leftmost section of Figure 2) and significantly fewer papers with one woman author than expected (negative bars in the middle section of Figure 2). The only significant difference for papers with multiple women authors happens for UIST, which has 25% fewer papers with multiple women authors than would be expected by chance.

4.5 Author position

We observe differences in the author positions women and men occupy. At junior ranks (1-2), all are more likely to be first or middle author than last across all conferences (Figure 3). The

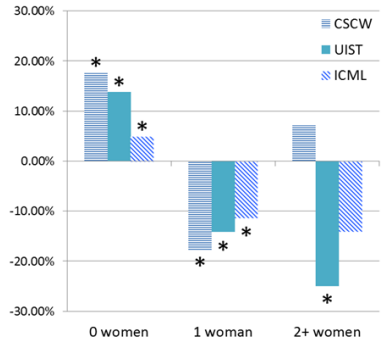


Fig. 2. Percentage differences between the expected and observed frequencies of gender compositions on papers with at least two authors (0 women, 1 woman, and 2 or more women). Significant differences, as determined by a binomial test, are indicated by * ($p < 0.05$).

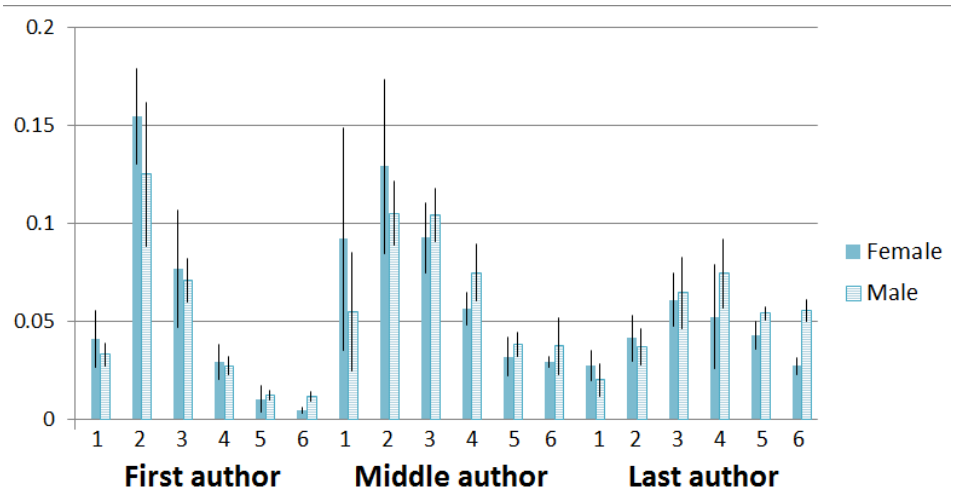


Fig. 3. The average distribution of genders by rank and author position. Y height is scaled by author slot and gender and averaged across conferences. Darker bars are female; lighter (striped) bars are male. Error bars show standard deviation. Solo-author papers are not included.

effect is stronger for junior women than men as a percentage of all author slots for each gender. Established authors (ranks 3-6) are more likely to be middle or last than first. Authors who have been publishing the longest (ranks 5-6) are more likely to be last than in any other slot. This effect is stronger for men than women. In CSCW only, a smaller fraction of women than men are middle authors. Also, in CSCW, a larger fraction of women than men first author two-author papers, while other conferences are the reverse. The result that high-rank authors are more likely to be last is consistent with input from many interviewees, who said the last slot goes to the advisor, typically the most senior member of the work team. Some authors may use alphabetical order in author lists. However, since there is no correlation in our dataset between gender and a name's alphabetical position, any differences observed across genders are not influenced by alphabetical ordering.

	First	Middle	Last
(Control) Intercept	**-0.53****	**-0.42****	**-1.20****
(Control) Rank	**-1.17****	**-0.16****	**1.14****
Intercept	**0.40****	**-0.82****	**0.49****
Higher-ranked co-authors	**0.48****	-0.06	**-0.33****
Lower-ranked co-authors	**-0.15****	**-0.08***	**0.38****
Percent female authors	**-0.12***	**0.12***	-0.01
Number of authors	**-0.13****	**0.25****	**-0.14****
Is female	-0.02	0.22	-0.19
Higher-ranked x female	0.03	-0.04	0.01
Lower-ranked x female	0.04	0.14	**-0.18***
Percent female x female	0.11	**-0.34***	**0.24***
Number authors x female	**-0.03***	-0.02	**0.04***

Table 4. Significant factors in predicting author position are **bold** (* indicates $p < .05$; ** indicates $p < .001$).

These data suggest that, while women are first authoring papers, they are not transitioning into the prestige slot associated with more senior roles (last author) at the same rates as men.

4.6 Predictive analysis of author position

We next used logistic regression to determine which aspects of collaboration and co-author characteristics contribute to author position. We predicted the position for a particular author (first, middle, last), controlling for rank because of rank-based differences in author ordering. Predictors were chosen based on our interviews and quantitative analysis. Due to the structural impact of author count on number of middle author slots, we included *number of authors*. Due to conference differences, we added *conference*. Due to gendered differences, we included *author gender*. Due to power differences, we included *percentage of higher- and lower-ranked co-authors*. Due to gendered co-author differences, we included *fraction of female authors* and crossed each predictor with author gender. Table 4 shows significant factors for predicting slots. Results confirm expected general trends. For example, higher rank makes first authorship less likely and last authorship more likely.

Our results verify the general trends described in the interviews. Table 4 shows that first author is more likely with a high percentage of higher-ranked co-authors and last author is more likely with a high percentage of lower-ranked co-authors. However, these effects *reverse for women authors*. Not only that, for a woman author, having more female co-authors makes last author more likely. That is, when a woman publishes with other women, she is more likely to hold the prestigious last author slot. These interactions between gender and other factors are important evidence that gender impacts authorship.

4.7 Discussion of bibliometric results

We found several trends in our bibliometric analysis (4355 papers by 7376 authors in three HCI and ML conferences over 20 years). First, women produce fewer papers than men in UIST and ICML, where they are underrepresented (the author populations at UIST and ICML are 15% and 11% women, respectively). There are no significant differences in output by multi-publishers or in yearly publication rates.

While women publish at rates similar to men, papers' gender compositions often do not reflect the gender balance of authors in each conference. For example, all three conferences have more papers with no women authors than would be expected by chance, were authors randomly selected.

In addition to the appearance of women on papers, we examined author order. During our qualitative work, most interviewees said the first author (often a student) did the most work, the last author (often an advisor) provided "intellectual and practical resources," and middle authors contributed to a lesser degree than first or last authors. For all three conferences, women are more often first author and men are more often last author. This can be partially explained by the fact

that more men than women have been publishing a long time and take advisory roles on papers they produce. We therefore controlled for rank in a predictive analysis of author position.

As expected, if an author has many lower-ranked co-authors, they are more likely to be last author and less likely to be first. However, if a woman has many lower-ranked co-authors, she is less likely to be last author. Perhaps senior women are more likely than men to assume cooperative work does not require a prestigious author position (*e.g.*, [62]), or maybe students are more likely to put a male co-advisor as last author: several faculty interviewees mentioned inappropriate author order placement by students as a cause for dissatisfaction in author order.

Another possible explanation could be our inference of rank from time publishing. This metric may result in inflated estimates of women's true ranks due to gender differences in time to tenure or promotion (*e.g.*, [66]). However, our interviews do not corroborate this, since no one spoke of *faculty rank* as important in author order decisions. Rather, the important rank difference is student or postdoc *versus* faculty, which should be accurately captured thanks to our removing the first ten years of data.

5 DISCUSSION

These quantitative results suggest that women assume different author positions than men. While we cannot verify the fairness of author order, the differences we find merit further study. Given that the distribution of women is similar to men, women's underrepresentation as last author cannot be explained simply by a glut of junior women. Our qualitative findings indicate other reasons: (1) women are not successfully negotiating for positions equivalent to men's, (2) co-authors are not putting women in the same slots as men, or (3) women are doing more first-author work on papers. Our qualitative study also found gender differences in interviewees' conceptualization of the authorship process, discussed in Section 3.4. Overall, men seemed less concerned than women about author order and its implications. Only men claimed people do not care about author order, called the process complicated, and said they trust their co-authors to be reasonable. Additionally, men tended to determine author order later in the paper-writing process than women.

We began this work motivated by personal experience and the belief that HCI can be more inclusive. We, as female authors, have a range of ranks (senior faculty to undergraduate), fields, and authorship experiences. Our recommendations for a more fair and open authorship process are derived chiefly from our interview study, which revealed researchers' process for deciding and discussing author order.

From these data we learned that our experiences and impressions about problematic aspects of authorship were not unique. While gender is not the only factor affecting authorship [21, 90], it does impact author position and the credit women receive for their work. The problem extends throughout the pipeline: women are not transitioning from positions where they do the most work (first author) to positions where they are seen as intellectual leaders (last author) at the same rate as men (Sections 4.5 and 4.6). Thus, simply getting more junior women in the door is unlikely to solve this problem.

The impact of differences in author position should not be understated. Ph.D. admission, faculty positions, and tenure are mediated by credit received through publication and author order. These effects may be exaggerated by tendencies to attribute more credit to men for their work [66], which can impact both author order and its interpretation.

That said, our work has several limitations. Our interview sample may not have fully represented all authorship approaches within HCI and ML. Intervention-based studies are needed to complement and confirm the theories in this paper. In our bibliometric study, rank was inferred from time publishing and thus does not capture true rank (institutional position). Gender was also inferred

(chiefly from first names) and therefore is error-prone. Additionally, our metrics did not address impact (*e.g.*, citation count) or author order fairness.

We cannot know about authors omitted from papers. We lack data on how, why, and by whom author order is assigned. For example, we cannot determine whether women are in middle slots when men are last more than the reverse because women are less likely to lead teams with men, or because decision makers are more likely to put women in middle slots than men. We also lack data on *submissions*. Gender differences may happen in review, such as women's papers receiving more scrutiny [66] or fewer women reviewers [44]. Finally, our analyses did not include race, socioeconomic status, sexual orientation, religion, *etc.* Gender is not the only axis on which people can be privileged or disadvantaged [21, 90], and it would be enlightening to include other power structures in an analyses of authorship and influence in research. Despite these flaws and uncertainties, we believe action is merited and unlikely to cause harm. Every author should assume responsibility for making fair choices about authorship. Too often, author order is not discussed or is discussed late, as interviewees shared.

Senior authors can initiate discussion on author order with students and other junior authors. Most, but not all, faculty interviewees said they already initiate discussions with students. They can teach co-authors about the meaning of author positions and the impact of author order on careers. By making criteria explicit, they can help avoid bias [82]. Senior authors can monitor efforts to mitigate the deflation and inflation that can occur in how women and men discuss relative contributions [31] and ensure all contributors meriting authorship receive it. Faculty can teach junior colleagues to stand up for themselves by example. Finally, authorship discussion is not a one-time event: it should occur as a paper evolves. For example, we added two authors to this paper after earlier submissions.

Junior authors can also initiate discussion. Although they often receive no formal training on author order according to our interviewees, students can ask faculty about this area of professional conduct. Students can also make women's contributions more visible to decision makers [26]. It is important for privileged authors to advocate for marginalized groups, as minorities may be penalized for speaking up about diversity while the dominant group is not [33].

As a field, we can establish clearer standards for the meaning of author positions. Many journals and professional organizations do not state authorship requirements [12]. For example, the TOCHI submission page says nothing about authorship requirements [2]. ACM's authorship policy includes authorship criteria—namely, that authors "made substantial intellectual contributions," "participated in drafting and/or revision," *etc.*—but provides no guidance on author order [1]. The National Science Foundation (NSF) requires researchers receiving NSF funding to complete "Responsible Conduct of Research" training, which in recent years has included authorship ethics [56]. It is unclear how many U.S. researchers have taken this training and if non-U.S. funding bodies require similar training.

As a field, we can also provide tools to make authorship discussions more productive. For example, the American Psychological Association website posts "scorecards" to determine who merits an author slot [7] and their position [8], based on a paper about author order [87]. The HCI and ML communities have offered no explicit instructions. Their discussions of authorship instead address disclosing contributions [6, 49], as is standard in medicine [5].

Finally, as a field, we can work harder to ensure that we appropriately assign credit to middle authors in career-affecting decisions. Disclosing individual author contribution in papers, as a field standard, would reduce reliance on imprecise meanings of author order. For example, our author contribution statement in the Acknowledgements explains what aspects of this project each author worked on. All these initiatives would work to eliminate implicit bias, as well as addressing the inadequacy of author order as a measure of contribution as academics write papers collaboratively.

This project has proven invaluable for all authors of this paper, who have changed their practices on authorship decisions. One now tracks how often she is last author compared to co-authors to ensure she speaks up if necessary. Another has raised the question of how to trade last position when students are co-advised. All faculty authors train their students to be proactive about authorship discussions, including providing scripts for requesting credit. While we have found these strategies to be transformational, further research on mitigating bias in authorship is important future work.

A INTERVIEW QUESTIONS

Discussions with interviewees in Section 3 were guided by the following questions and prompts:

(1) How do you determine author order?
(2) How do you discuss author order with your collaborators?
(3) When does this discussion happen?
(4) Have you ever started with one order and then changed?
(5) What factors do you consider when deciding author order?
(6) How did you learn what author order means?
(7) Identify a paper where you were happy about author order. Talk about it.
(8) Identify a paper where you weren't happy about author order. Talk about it.
(9) How do you think [co-author] would describe the author order decision process?
(10) Were you ever included as an author when you did not want to be?
(11) Were you ever excluded when you thought you should be an author?
(12) Has one of your collaborators ever been upset about or disagreed with author order? How did you handle this situation?

ACKNOWLEDGMENTS

The first author (Early) did the majority of the data analysis and writing. The last author (Mankoff) initiated the project and provided funding and insight to the project as a whole. The middle authors (alphabetical order) contributed to the whole project, with Hammer guiding data analysis and providing a background in writing about gender and academia, Hofmann conducting half the interviews, Rode providing knowledge of feminist theory, and Wong contributing actively to the data analysis process. This work was supported by the Okawa Foundation.

REFERENCES

[1] ACM Transactions on Computer-Human Interaction. 2016. Policy on authorship. https://www.acm.org/publications/policy-on-authorship
[2] ACM Transactions on Computer-Human Interaction. n.d.. Author information. https://tochi.acm.org/authors/
[3] Claire Ainsworth. 2015. Sex redefined. https://www.nature.com/news/sex-redefined-1.16943
[4] Gerald V. Barrett and Scott B. Morris. 1993. The American Psychological Association's *amicus curiae* brief in *Price Waterhouse v. Hopkins*: The values of science versus the values of the law. *Law and Human Behavior* 17, 2 (1993), 201–215.
[5] Tamara Bates, Ante Anić, Matko Marušić, and Ana Marušić. 2004. Authorship criteria and disclosure of contributions: Comparison of 3 general medical journals with different author contribution forms. *Journal of the American Medical Association* 292, 1 (2004), 86–88.
[6] Christine Bauer and Afsaneh Doryab. 2016. Solving the battle of first-authorship: Using interactive technology to highlight contributions. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 609–620.
[7] Bharati Belwalkar and Steven Toaddy. 2014. Authorship determination scorecard. http://www.apa.org/science/leadership/students/authorship-determination-scorecard.pdf
[8] Bharati Belwalkar and Steven Toaddy. 2014. Authorship tie-breaker scorecard. http://www.apa.org/science/leadership/students/authorship-tie-breaker-scorecard.pdf
[9] Jerome T. Bentley and Rebecca Adamson. 2003. *Gender differences in the careers of academic scientists and engineers: A literature review.* Technical Report.

[10] Sue E. Berryman. 1983. *Who will do science? Trends, and their causes in minority and female representation among holders of advanced degrees in science and mathematics.* Technical Report.
[11] Mohit Bhandari, Thomas A. Einhorn, Marc F. Swiontkowski, and James D. Heckman. 2003. Who did what? *The Journal of Bone and Joint Surgery* 85, 8 (2003), 1605–1609.
[12] Lana Bošnjak and Ana Marušić. 2012. Prescribed practices of authorship: Review of codes of ethics from professional bodies and journal guidelines across disciplines. *Scientometrics* 93, 3 (2012), 751–763.
[13] Hannah Riley Bowles, Linda Babcock, and Lei Lai. 2007. Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes* 103, 1 (2007), 84–103.
[14] Sheryl D. Brahnam, Thomas M. Margavio, Michael A. Hignite, Tonya B. Barrier, and Jerry M. Chin. 2005. A gender-based categorization for conflict resolution. *Journal of Management Development* 24, 3 (2005), 197–208.
[15] Michael Burawoy. 1998. The extended case method. *Sociological Theory* 16, 1 (1998), 4–33.
[16] Judith Butler. 1990. *Gender Trouble: Feminism and the Subversion of Identity.* Routledge.
[17] Tracy Camp. 1997. The incredible shrinking pipeline. *Commun. ACM* 40, 10 (1997), 103–110.
[18] Stephen J. Ceci and Wendy M. Williams. 2011. Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences* 108, 8 (2011), 3157–3162.
[19] J. McGrath Cohoon and William Aspray (Eds.). 2006. *Women and Information Technology: Research on Underrepresentation.* Vol. 1. The MIT Press.
[20] Jonathan R. Cole, Burton Singer, et al. 1991. A theory of limited differences: Explaining the productivity puzzle in science. In *The Outer Circle: Women in the Scientific Community*, Harriet Zuckerman, Jonathan R. Cole, John T. Bruer, et al. (Eds.). Norton, New York, 277–310.
[21] Kimberle Crenshaw. 1989. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum* 1989, 1 (1989), 139–167.
[22] Nilanjana Dasgupta, Melissa McManus Scircle, and Matthew Hunsinger. 2015. Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academy of Sciences* 112, 16 (2015), 4988–4993.
[23] Derek J. de Solla Price and Donald Beaver. 1966. Collaboration in an invisible college. *American Psychologist* 21, 11 (1966), 1011–1018.
[24] John F. Dovidio and Samuel L. Gaertner. 2000. Aversive racism and selection decisions: 1989 and 1999. *Psychological Science* 11, 4 (2000), 315–319.
[25] The Economist. 2016. Why research papers have so many authors. https://www.economist.com/news/science-and-technology/21710792-scientific-publications-are-getting-more-and-more-names-attached-them-why
[26] Juliet Eilperin. 2016. White House women want to be in the room where it happens. https://www.washingtonpost.com/news/powerpost/wp/2016/09/13/white-house-women-are-now-in-the-room-where-it-happens/
[27] Ferhat Elmas. 2013. Sex Machine. https://pypi.python.org/pypi/SexMachine/
[28] Mary Frank Fox. 2005. Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science* 35, 1 (2005), 131–150.
[29] Mary Frank Fox and Sushanta Mohapatra. 2007. Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. *The Journal of Higher Education* 78, 5 (2007), 542–571.
[30] Sebastian Frische. 2012. It is time for full disclosure of author contributions. *Nature* 489, 7417 (2012), 475.
[31] Michelle C. Haynes and Madeline E. Heilman. 2013. It had to be you (not me)! Women's attributional rationalization of their contribution to successful joint work outcomes. *Personality and Social Psychology Bulletin* (2013), 956–969.
[32] Madeline E. Heilman and Michelle C. Haynes. 2005. No credit where credit is due: Attributional rationalization of women's success in male-female teams. *Journal of Applied Psychology* 90, 5 (2005), 905–916.
[33] David R. Hekman, Stefanie K. Johnson, Maw-Der Foo, and Wei Yang. 2016. Does Diversity-valuing Behavior Result in Diminished Performance Ratings for Nonwhite and Female Leaders? *Academy of Management Journal* (2016).
[34] Erin Hengel. 2017. Publishing while female: Are women held to higher standards? Evidence from peer review. (2017). https://doi.org/10.17863/CAM.17548
[35] Gary Holden, Gary Rosenberg, and Kathleen Barker. 2005. Bibliometrics: A potential decision making aid in hiring, reappointment, tenure and promotion decisions. *Social Work in Health Care* 41, 3-4 (2005), 67–92.
[36] Janet Holmes. 2005. Power and discourse at work: Is gender relevant? In *Feminist Critical Discourse Analysis: Gender, Power and Ideology in Discourse*, Michelle M. Lazar (Ed.). Springer, 31–60.
[37] Shulamit Kahn. 1993. Gender differences in academic career paths of economists. *The American Economic Review* 83, 2 (1993), 52–56.
[38] Cheryl R. Kaiser and Carol T. Miller. 2001. Stop complaining! The social costs of making attributions to discrimination. *Personality and Social Psychology Bulletin* 27, 2 (2001), 254–263.

[39] Jofish Kaye. 2015. A rose by any other name. http://sigchi.tumblr.com/post/127563985260/a-rose-by-any-other-name
[40] Svein Kyvik and I. Marheim Larsen. 1994. International contact and research performance. *Scientometrics* 29, 1 (1994), 161–172.
[41] Svein Kyvik and Mari Teigen. 1996. Child care, research collaboration, and gender differences in scientific productivity. *Science, Technology & Human Values* 21, 1 (1996), 54–71.
[42] Michelle M. Lazar. 2005. Politicizing gender in discourse: Feminist critical discourse analysis as political perspective and praxis. In *Feminist Critical Discourse Analysis: Gender, Power and Ideology in Discourse*, Michelle M. Lazar (Ed.). Springer, 1–28.
[43] Erin Leahey. 2006. Gender differences in productivity research: Specialization as a missing link. *Gender & Society* 20, 6 (2006), 754–780.
[44] Jory Lerback and Brooks Hanson. 2017. Journals invite too few women to referee. *Nature* 541 (2017), 455–457.
[45] Sarah-Jane Leslie, Andrei Cimpian, Meredith Meyer, and Edward Freeland. 2015. Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 6219 (2015), 262–265.
[46] Michael Ley. 2002. The DBLP computer science bibliography: Evolution, research issues, perspectives. In *International Symposium on String Processing and Information Retrieval*. Springer, 1–10.
[47] Juan M. Madera, Michelle R. Hebl, and Randi C. Martin. 2009. Gender and letters of recommendation for academia: Agentic and communal differences. *Journal of Applied Psychology* 94, 6 (2009), 1591–1599.
[48] Paula Mählck. 2001. Mapping gender differences in scientific careers in social and bibliometric space. *Science, Technology & Human Values* 26, 2 (2001), 167–190.
[49] Esperanza Marcos, Juan Manuel Vara, and Valeria de Castro. 2012. Author order: What science can learn from the arts. *Commun. ACM* 55, 9 (2012), 39–41.
[50] Richard F. Martell, David M. Lane, and Cynthia Emrich. 1996. Male-female differences: A computer simulation. *American Psychologist* (Feb 1996), 157–158.
[51] Elba Mauleón and María Bordons. 2006. Productivity, impact and publication habits by gender in the area of Materials Science. *Scientometrics* 66, 1 (2006), 199–218.
[52] Peggy McIntosh. 2009. White privilege and male privilege: A personal account of coming to see correspondences through work in women's studies. In *Privilege and Prejudice: Twenty Years with the Invisible Knapsack*, Karen Weekes (Ed.). Cambridge Scholars Publishing, 7–19.
[53] Corinne A. Moss-Racusin, John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. 2012. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences* 109, 41 (2012), 16474–16479.
[54] Brian Muller. 2014. Why I'm renaming a gem. http://findingscience.com/ruby/2014/11/17/why-im-renaming-a-gem.html
[55] Jennifer C. Nash. 2008. Re-thinking intersectionality. *Feminist Review* 89, 1 (2008), 1–15.
[56] National Science Foundation. 2009. Responsible conduct of research. https://www.nsf.gov/bfa/dias/policy/rcr.jsp
[57] Paula Nicolson. 2015. Gender, power and organization: A psychological approach.
[58] Joshua M. Paiz, Elizabeth Angeli, Jodi Wagner, Elena Lawrick, Kristen Moore, Michael Anderson, Lars Soderlund, Allen Brizee, and Russell Keck. 2014. APA style: In-text citations. https://owl.english.purdue.edu/owl/resource/560/03/
[59] Emma Pierson. 2014. In science, it matters that women come last. http://fivethirtyeight.com/features/in-science-it-matters-that-women-come-last/
[60] Christina Richards, Walter Pierre Bouman, Leighton Seal, Meg John Barker, Timo O. Nieder, and Guy T'Sjoen. 2016. Non-binary or genderqueer genders. *International Review of Psychiatry* 28, 1 (2016), 95–102.
[61] Jennifer A. Rode. 2011. Reflexivity in digital anthropology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 123–132.
[62] Luisa Martín Rojo and Concepción Gómez Esteban. 2005. The gender of power: The female style in labour organizations. In *Feminist Critical Discourse Analysis: Gender, Power and Ideology in Discourse*, Michelle M. Lazar (Ed.). Springer, 61–89.
[63] Mareva Sabatier, Myriam Carrere, and Vincent Mangematin. 2006. Profiles of academic activities and careers: Does gender matter? An analysis based on French life scientist CVs. *The Journal of Technology Transfer* 31, 3 (2006), 311–324.
[64] Jessica Salvatore and J. Nicole Shelton. 2007. Cognitive costs of exposure to racial prejudice. *Psychological science* 18, 9 (2007), 810–815.
[65] Jeffrey C. Sandler and Brenda L. Russell. 2005. Faculty-student collaborations: Ethics and satisfaction in authorship credit. *Ethics & Behavior* 15, 1 (2005), 65–80.
[66] Heather Sarsons. 2017. Recognition for group work: Gender differences in academia. *American Economic Review* 107, 5 (2017), 141–45.
[67] Allison K. Shaw and Daniel E. Stanton. 2012. Leaks in the pipeline: Separating demographic inertia from ongoing gender differences in academia. *Proceedings of the Royal Society of London B: Biological Sciences* 279, 1743 (2012),

3736–3741.
[68] Reena Sidhu, Praveen Rajashekhar, Victoria L. Lavin, Joanne Parry, James Attwood, Anita Holdcroft, and David S. Sanders. 2009. The gender imbalance in academic medicine: A study of female authorship in the United Kingdom. *Journal of the Royal Society of Medicine* 102, 8 (2009), 337–342.
[69] Gerhard Sonnert and Gerald Holton. 1996. Career patterns of women and men in the sciences. *American Scientist* 84, 1 (1996), 63–71.
[70] Steven J. Spencer, Claude M. Steele, and Diane M. Quinn. 1999. Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* 35, 1 (1999), 4–28.
[71] Steven Stack. 2002. Gender and scholarly productivity: The case of criminal justice. *Journal of Criminal Justice* 30, 3 (2002), 175–182.
[72] Claude M. Steele and Joshua Aronson. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69, 5 (1995), 797–811.
[73] Rhea E. Steinpreis, Katie A. Anders, and Dawn Ritzke. 1999. The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles* 41, 7-8 (1999), 509–528.
[74] Paula E. Stephen. 1987. *Demographic and Economic Determinants of Scientific Productivity.* Policy Research Program, College of Business Administration, Georgia State University.
[75] Joanne Hoven Stohs. 2000. Multicultural women's experience of household labor, conflicts, and equity. *Sex Roles* 42, 5-6 (2000), 339–361.
[76] Derald Wing Sue. 2010. *Microaggressions in Everyday Life: Race, Gender, and Sexual Orientation.* John Wiley & Sons.
[77] Janet K. Swim and Lauri L. Hyers. 1999. Excuse me–What did you just say?!: Women's public and private responses to sexist remarks. *Journal of Experimental Social Psychology* 35, 1 (1999), 68–88.
[78] Ezequiel Tacsir, Matteo Grazzi, and Rafael Castillo. 2014. *Women in science and technology: What does the literature say?* Technical Report. Inter-American Development Bank.
[79] Frances Trix and Carolyn Psenka. 2003. Exploring the color of glass: Letters of recommendation for female and male medical faculty. *Discourse & Society* 14, 2 (2003), 191–220.
[80] Teja Tscharntke, Michael E. Hochberg, Tatyana A. Rand, Vincent H. Resh, and Jochen Krauss. 2007. Author sequence and credit for contributions in multiauthored publications. *PLoS Biol* 5, 1 (2007), e18.
[81] Lori Turk-Bicakci and Andrea Berger. 2014. Leaving STEM: STEM Ph.D. Holders in Non-STEM Careers. Issue Brief. *American Institutes for Research* (2014).
[82] Eric Luis Uhlmann and Geoffrey L. Cohen. 2005. Constructed criteria redefining merit to justify discrimination. *Psychological Science* 16, 6 (2005), 474–480.
[83] Virginia Valian. 2005. Beyond gender schemas: Improving the advancement of women in academia. *Hypatia* 20, 3 (2005), 198–213.
[84] Candace West and Don H. Zimmerman. 1987. Doing gender. *Gender & Society* 1, 2 (1987), 125–151.
[85] Jevin D. West, Jennifer Jacquet, Molly M. King, Shelley J. Correll, and Carl T. Bergstrom. 2013. The role of gender in scholarly authorship. *PloS one* 8, 7 (2013), e66212.
[86] Joan Williams, Katherine W. Phillips, and Erika V. Hall. 2014. Double jeopardy?: Gender bias against women of color in science. http://www.uchastings.edu/news/articles/2015/01/williams-double-jeopardy-report.php
[87] Roger B. Winston. 1985. A suggested procedure for determining order of authorship in research publications. *Journal of Counseling & Development* 63, 8 (1985), 515–518.
[88] Alison Wylie. 2011. What knowers know well: Women, work, and the academy. In *Feminist Epistemology and Philosophy of Science.* Springer, 157–179.
[89] Yu Xie and Kimberlee A. Shauman. 1998. Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review* (1998), 847–870.
[90] Iris Marion Young. 2011. *Justice and the Politics of Difference.* Princeton University Press.

Received April 2018; revised July 2018; accepted September 2018