

---

This is an electronic reprint of the original article.  
This reprint may differ from the original in pagination and typographic detail.

Duran, Rodrigo; Haaranen, Lassi; Hellas, Arto  
**Gender differences in introductory programming**

*Published in:*  
SIGCSE 2020 - Proceedings of the 51st ACM Technical Symposium on Computer Science Education

*DOI:*  
[10.1145/3328778.3366852](https://doi.org/10.1145/3328778.3366852)

Published: 26/02/2020

*Document Version*  
Peer reviewed version

*Please cite the original version:*  
Duran, R., Haaranen, L., & Hellas, A. (2020). Gender differences in introductory programming: Comparing moocs and local courses. In *SIGCSE 2020 - Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (pp. 692-698). (Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE). ACM. <https://doi.org/10.1145/3328778.3366852>

---

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Gender Differences in Introductory Programming: Comparing MOOCs and Local Courses

Rodrigo Duran  
rodrigo.duran@aalto.fi  
Aalto University  
Espoo, Finland

Lassi Haaranen  
lassi.haaranen@aalto.fi  
Aalto University  
Espoo, Finland

Arto Hellas  
arto.hellas@aalto.fi  
Aalto University  
Espoo, Finland

## ABSTRACT

We analyzed three introductory programming MOOCs and four introductory programming courses offered locally in a Finnish university. The course has been offered in all instances with roughly the same content, barring adjustments based on course feedback. We sought to understand how gender interacts with participating in the course in both instances. In particular, we looked at the differences in persistence, confidence, interest in CS, prior experience, and performance between men and women.

Overall, we found that men have more prior experience in both instances and have a higher interest in a CS degree. Furthermore, men perform slightly better on the MOOC while there was no significant difference in performance when it came to gender in the local instance. Aligned with prior research, we found a considerable gap in confidence between male and female students in both instances. At the same time, while women are still underrepresented in CS, we observe a considerable increase in women attending the MOOC. Unfortunately, women are also more likely to drop out early on in the MOOC than men.

## CCS CONCEPTS

• **Social and professional topics** → *Computing education; CS1; Student assessment; Gender*; • **Applied computing** → *Interactive learning environments; E-learning*.

## KEYWORDS

introductory programming, CS1, MOOC, gender, persistence, self-efficacy, confidence, interest, performance

### ACM Reference Format:

Rodrigo Duran, Lassi Haaranen, and Arto Hellas. 2020. Gender Differences in Introductory Programming: Comparing MOOCs and Local Courses. In *The 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20)*, March 11–14, 2020, Portland, OR, USA. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3328778.3366852>

## 1 INTRODUCTION

During the last century, the ability to write and read computer programs – *to program* – has evolved from a skill of a select few to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
SIGCSE '20, March 11–14, 2020, Portland, OR, USA

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-6793-6/20/03...\$15.00  
<https://doi.org/10.1145/3328778.3366852>

a skill common and wanted enough that it has even been discussed as the “literacy of today”<sup>1</sup>. At the same time, those working as programmers as well as those studying programming e.g. in a computer science program at a university are still mostly men [5, 9, 35]. This is an issue that has been acknowledged within the computer science education research community as well as within the broader science communities [32].

There is a strong push towards making programming and computer science (CS) studies accessible and available for all. Multiple national and international initiatives exist, both from the public and the private sector, and many countries offer CS as a part of their K-12 curricula [17, 19, 21, 26]. Recent efforts towards making programming available for all have been partially further fueled by the emergence of massive open online courses (MOOCs), which have been touted to make “education borderless, gender-blind, race-blind, class-blind and bank-account blind” [1]. Many MOOCs in programming are listed in (unofficial) lists of most popular MOOCs of all time<sup>2</sup>. In addition to teaching programming, MOOCs are used for introducing teachers to teaching CS [11, 40] and for recruiting students into CS [37].

While there has been a significant increase in the availability of programming courses and CS education, it is unclear whether this push has influenced the existing gender disparity in CS. Recent research suggests that there is no difference in gender in completion rates of MOOCs in general [20], but STEM subjects may suffer from a lower retention rate for women [33]. Similarly, a recent study suggested that using a MOOC in programming for recruiting students into CS, while yielding students who perform better in their CS studies, may also lead to increased gender disparity [24].

Seeking to understand the issue, we analyze introductory programming courses offered by a Finnish university. Focusing on three recent introductory programming MOOCs and four on-campus courses with the same content, we study differences in persistence. Our research question for this work is: *How does gender interact with taking a local or a MOOC version of an introductory programming course?* Specifically, we look at the differences men and women have in terms of *persistence, confidence, interest in CS, prior experience, and performance*.

The article is organized as follows. In the next section, we discuss the background for our work: persistence in MOOCs and women in CS and STEM. In Section 3, we outline the context, the course under study, and the data at our disposal. Results are presented in Section 4 and further discussed in Section 5. We conclude with our main findings in Section 6 and suggest future work.

<sup>1</sup>e.g. European Commission (EC) policy “Coding - the 21st century skill” <https://ec.europa.eu/digital-single-market/en/coding-21st-century-skill>

<sup>2</sup>e.g. <https://www.onlinecourseport.com/the-50-most-popular-moocs-of-all-time/>

## 2 BACKGROUND

### 2.1 Persistence in MOOCs

In 2013 Agarwal discussed MOOCs and their potential for making education more inclusive [1]. Half a decade of experience and research later, this does not appear to be the case. Instead, MOOC participants seem to be highly educated, young, and male with the majority coming from developed countries [7]. Hone & El Said found no difference in completion rates between gender, level of study (undergraduate or postgraduate) or MOOC platform [20], although Rayyan et al. did note a “higher retention for males” on a course in a STEM field [33].

The high rate of dropouts is a common phenomenon in MOOCs. Clow coined the term ‘funnel of participation’ to describe the low persistence rate and presented empirical data to showcase it [8]. The funnel of participation has four steps: awareness of a MOOC, registration to the MOOC, activity on the course, and making meaningful learning progress. At each step, the drop off is large.

When it comes to predictors of completing a MOOC, Greene et al. found that prior experience of MOOCs, as well as, being older and more educated made dropping out less likely. The strongest predictor for completing a course that they found was a self-rated commitment to completion. [15]

In Finland, where this study is conducted, introductory programming MOOCs have been studied scarcely. Focusing on a MOOC in programming, there are some differences between age groups attending MOOCs, MOOCs have a clear drop-out trend, and women are less likely to attend the programming MOOC than men [23, 24].

### 2.2 Women in Computer Science and STEM

There is plenty of research and discussion on the underrepresentation of women in STEM in general (e.g. [2, 10, 16, 27]), as well as, specifically in CS [12, 28]. Confidence gap – men report higher confidence compared to women – is also a known issue in CS [3, 25].

Women have significantly lower involvement compared to men in CS. In the United States, for instance, in 2014 only 20% of CS A advanced placement exam takers were women [12]. Murphy et al. did find a lower pre-college programming experience with women [28]. However, they concluded that it did not correlate with success on the course. Additionally, women reported an almost equal level of mastery after the introductory programming course. Griffith observed that an increase in female graduate students also increased women in undergraduates [16].

Importantly, gender does not seem to play a part in student performance on CS courses. Werth [41] reports that there is no correlation between gender and course grades while Byrne and Lyons [4] reported female students having a slightly higher point average. Similarly, Ventura [36] found that gender did not affect programming assignments, exams, and overall course.

Beyer et al. [3] examined a large set of variables related to gender. The results did not show a difference between genders in quantitative ability, interest in CS, stereotypes, and knowledge of CS, or attitudes towards CS courses and instructors. However, they report female CS majors having lower confidence levels and having less previous experience with computers.

Stemming from the disparity of genders studying CS, women are also underrepresented when it comes to the research literature in

computing. Cohoon et al. [10] researched the gender of authors in ACM conference papers between 1967-2009. Of the authors those they could ascertain the gender, only 22% were women. However, the share of women authors rose from 7% in 1967 to 27% in 2009.

Women of color in STEM stand the *double bind* facing both racism and sexism at the same time. Ong et al. reviewed empirical research regarding the double bind and factors affecting women of color [29]. They highlight the multitude of issues, such as the overall STEM climate, lack of faculty role models, and funding for studies. For an illustrative case study of the double bind, see Foor et al. ethnographic study which highlights one of the core issues in the title “I Wish that I Belonged More in this Whole Engineering Group” [13].

## 3 CONTEXT & DATA

*Context.* The context of the study is an introductory programming course offered by a research-first University in Finland. There are two instances of the course: (1) for undergraduate students, the course is divided into two seven-week courses, and (2) for non-affiliated students, the course is offered as a full fourteen-week MOOC.

Undergraduate students take either only the first part of the course or both parts, depending on their major and preferences. Students attending the MOOC may, if they choose to do so, drop out in the middle with a diploma for the first seven weeks of the course. Both instances use the same course materials and assignments, and provide the same online support (synchronous and asynchronous discussion channels). Undergraduate students also have weekly lectures and walk-in labs, where they can ask for additional support. For additional details, see [24, 38, 39].

The MOOC serves also as an entrance exam to the university. Each year, approximately 50 of the students participating in the MOOC are admitted to a Finnish university as CS majors. Students can apply to study CS also using their high-school diploma or through a national CS-specific entrance exam. Even if students in the MOOC do not wish to apply for a CS study right at a Finnish university, completing the course has benefits: students in high-school often receive credits from completing the course or parts of the course from their high-school, and Universities across Finland typically give students credits for completing the course. Credits can be also received afterward, e.g. if the student has not enrolled in a University while working on the course. The course is also used by several companies for training their employers.

The course is offered in Finnish and, consequently, almost all of the students in the MOOC and the locally offered programming courses speak Finnish. Moreover, the students have attended the Finnish education system in their primary and secondary studies, meaning that the student population is rather homogeneous when compared to many international CS degrees given in English. Besides, studying in Finland is different from many other countries as there are no tuition fees and students receive financial aid from the government during their studies.

Students in the locally offered version have an end-of-course exam which they must take to pass the course, while students in the MOOC may take an online exam if they wish to receive official credits. The grading in both courses is based on the assignments

that students complete during the course and an end of course exam, where both amount for approximately the same amount of course points. To reach the best grade, the student needs to receive over 90% of the overall course points, and the lowest passing grade requires at least 50% of the overall course points. The student must receive at least half of the points in the exam to pass the course.

If a student wishes to apply for a study position, the student needs to complete at least 90% of the weekly assignments and attend a locally organized interview at the university campus. During the interview, students express their interests and motivations towards studying CS and work on programming assignments similar to the ones that they have completed during the course.

*Data.* We have collected student data from four locally offered 7-week introductory programming courses and three 14-week MOOCs. Our data contain information on students' completed assignments as well as data from a voluntary background survey. The survey asks for students' gender, year of birth, previous programming experience (in terms of hours programmed), interest in deepening CS knowledge or taking CS as a minor or major subject, and on whether the student is expecting to do well in the course that they are currently attending.

The voluntary survey is completed at the beginning of the course when students have not yet started to read course materials or work on course assignments. This means that interest in CS as well as expecting to do well in the course should not be influenced by the course at hand. In our data, about 1% of the MOOC participants and about 5% of the local participants chose not to disclose their gender or chose another option than male or female. In this analysis, we focus on the students who self-identified as female or male.

We measure persistence in terms of drop-out. Confidence, interest in studying CS, and prior experience are measured through the questionnaire given at the beginning of the course. Performance is measured using completed course assignments. We combined the data from MOOC instances and local versions of the course for the analysis.

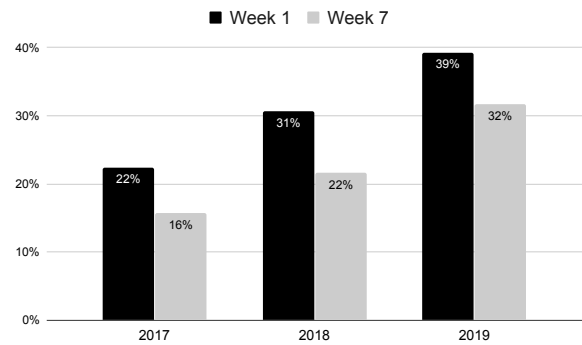
## 4 RESULTS

Overall, 4728 students answered the background survey, completed at least one assignment, and provided research consent (57% of the overall population). The MOOC version of the course had more men (64.3%) than women (35.7%) and men are marginally younger than women on average (33.29 versus 34.28 years old). In the local version of the course, the gender distribution is almost equal, with slightly more men (52.8%) than women (47.2%) with similar average ages (27.05 years for men versus 26.99 years for women).

In the past three years there has been an increasing proportion of women participating in the MOOC. Figure 1 shows the percentage of female students on the MOOC version on Week 1 and Week 7 for years 2017-2019.

### 4.1 Persistence

In the MOOC version, there is a difference in drop-out rates between men and women, especially in the early weeks as seen in Figure 2. 26% of women dropped the MOOC from Week1 to Week2 and 27% dropped out from Week2 to Week3. The corresponding numbers for men were 18% and 15%. Due to the option of getting a diploma



**Figure 1: Female students on the MOOC version on Week 1 and Week 7 between 2017-2019.**

in the middle of the MOOC, the drop-out is overall higher moving to Week 8, 26% for women and 16% for men.

Similarly, the weekly drop-out rates for the local course can be seen in Figure 2. The drop-out is considerably lower compared to MOOC, which is expected. Additionally, there is only a small difference in drop-out rates between men and women and they follow a similar pattern.

### 4.2 Prior Experience, Confidence, and Interest

Prior-experience was measured by the self-reported total number of hours dedicated to programming by the student. For the overall data, students can be considered beginners. However, there is a large variation in the reported hours, as presented in Table 1. Since the data is not normally distributed, a Mann-Whitney U test was performed to evaluate the difference in prior-experience between genders. *There is a significant difference in prior-experience between men and women both in the MOOCs ( $W = 2233608$ ;  $p < 0.00$ ), and local courses ( $W = 94834$ ;  $p < 0.00$ ), but the difference is small in the local courses ( $r = 0.22$ ) and moderate in MOOCs ( $r = 0.34$ ).* The  $r_{equivalent}$  from Rosenthal and Rubin [34] is used as effect size estimate for non-parametric tests in our analysis.

**Table 1: Prior-experience descriptive statistics. Experience is reported as total hours programmed.**

	MOOC (N = 3904)			Local (N = 824)		
	Men	Women	Overall	Men	Women	Overall
Mean ( $\mu$ )	963.64	409.70	768.07	429.12	1139.56	765.60
Median	40	1	20	20	1	10
Std. Dev ( $\sigma$ )	4833.06	3784.59	4498.22	3771.57	7084.18	5598.19

*There is a significant difference in self-efficacy by gender (Mann-Whitney U test  $W = 3465726$ ,  $p$ -value  $< 0.00$ ) overall, with a moderate effect size ( $r = 0.31$ ). A similar pattern emerges when analyzing MOOCs and local courses, both with significant differences between men and women regarding self-efficacy with moderate effect sizes (MOOCs  $r = 0.30$ , local course  $r = 0.34$ ). Answers to the self-efficacy question in the survey are summarized in Figure 3. Overall, men rated their confidence in performing well on the course higher both*

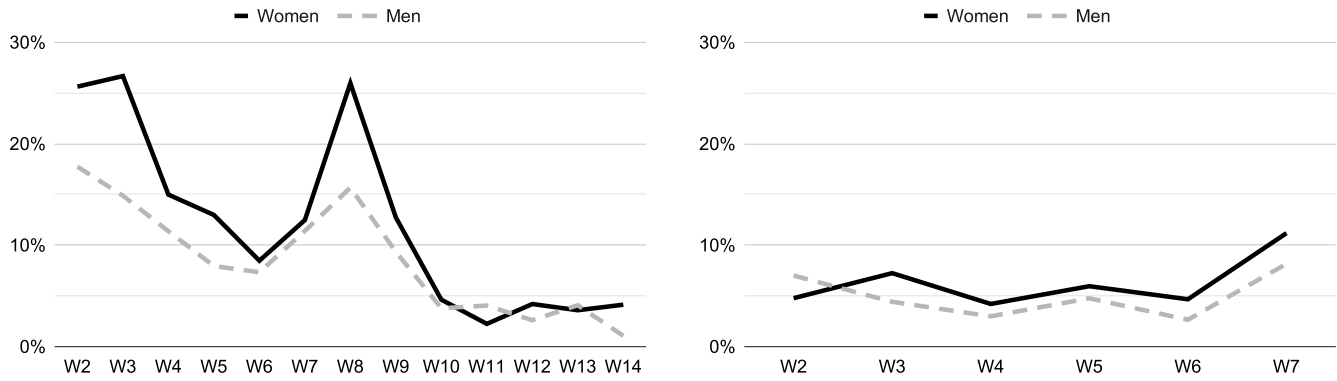


Figure 2: Left: Weekly drop-out percentage for the MOOCs. Right: Weekly drop-out percentage for the local version.

on the MOOC and local versions with approximately 75% of them at least somewhat agreeing with the statement ‘I believe I will do well on this course’. In contrast, about half of the women were either neutral or disagreed with the statement at least to some degree.

In general, the interest in CS is high (see Table 2) both in MOOCs and local courses, with men more interested in CS than women. A Mann-Whitney U test shows that *there is a significant difference in interest between men and women both in MOOCs ( $W = 1282367$ ,  $p$ -value  $< 0.00$ ) and local courses ( $W = 80424$ ;  $p$ -value  $< 0.00$ ) but the effect size is small ( $r = 0.11$  for MOOCs;  $r = 0.14$  for local courses).*

Table 2: Interest in CS descriptive statistics. Interest is reported as a yes (1) or no (0) question.

	MOOC (N = 3904)			Local (N = 824)		
	Men	Women	Overall	Men	Women	Overall
Mean ( $\mu$ )	0.84	0.74	0.81	0.81	0.68	0.75
Median	1	1	1	1	1	1
Std. Dev ( $\sigma$ )	0.37	0.44	0.40	0.40	0.47	0.44

### 4.3 Performance

Given the structure of the MOOC, which provided a clear exit point after week 7, we investigated partial grades until week 7 as a proxy for performance. In the subsequent analysis of performance, we considered students who submitted coursework in any of the weeks between weeks 1 and 7. Table 3 shows that students who persisted in the MOOC, both male and female participants had higher means and smaller standard deviations in grades compared to the local courses. In the MOOC version, a Mann-Whitney U test shows that *there is a significant difference in performance between men and women ( $W = 314687$ ;  $p$ -value  $< 0.00$ ) but the effect size is small ( $r = 0.14$ ). The local courses do not show a significant difference between men and women ( $W = 44098$ ,  $p$ -value =  $0.054$ ,  $r = 0.08$ ).*

### 4.4 Interactions

To investigate which factors could affect course assignment grades (as partial grades on week 7), summarized in Table 3, we analyzed the interactions between gender, self-efficacy, and interest in CS. Figure 4 shows a trend where increasing self-efficacy levels are

Table 3: Assignment grades (ranging from 0 to 100) of students who submitted coursework between weeks 1 and 7.

	MOOC (N = 1632)			Local (N = 570)		
	Men (N = 1178)	Women (N = 454)	Overall	Men (N = 309)	Women (N = 261)	Overall
Mean ( $\mu$ )	96.92	95.11	96.42	89.78	87.46	88.72
Median	100	98.74	99.66	94.21	93.31	93.94
Std. Dev ( $\sigma$ )	7.67	8.91	8.07	12.45	13.72	13.09

associated with higher medians of assignment grades. Since the assumptions required by a parametric ANOVA test (experimental errors are normally distributed, evaluated by a Shapiro-Wilks test on the residuals, and equal variances between treatments, evaluated by Levene’s test) were not met, we used the non-parametric variant of the ANOVA Wald-Type Statistic (WTS) [14].

The MOOCs ANOVA results show that *self-efficacy ( $WTS = 21.29$ ;  $p = 0.01$ ) is the only significant effect on assignment grades with no significant interactions. The local course ANOVA results show no significant main effects with a significant interaction between gender and interest in CS ( $WTS = 4.32$ ;  $p = 0.04$ ).*

Evaluating students who submitted coursework from weeks 1 to 7, and answered the questionnaire (N = 1777), the *partial correlation between grades and gender, with self-efficacy as a component is  $r = -0.13$ ;  $p < 0.00$ . The correlation between grades and self-efficacy with gender as a component is  $r = 0.15$ ;  $p < 0.00$ . The correlation between grades and gender with prior-experience as component is  $r = -0.12$ ;  $p < 0.00$ . Analyzing the subset of students who attended the MOOCs (N = 1301), the partial correlation between grades and gender with self-efficacy as a component of the partial correlation is  $r = 0.10$ ;  $p < 0.00$ . For those who attended the local version of the course (N = 476) the partial correlation is  $r = -0.04$ ;  $p = 0.34$ .*

## 5 DISCUSSION

Here we summarize the main results of our analysis:

- Prior-experience: Significant difference between genders. Moderate effect size in the MOOC, small in the local course.
- Self-efficacy: Significant difference between genders. Moderate effect size across the board.

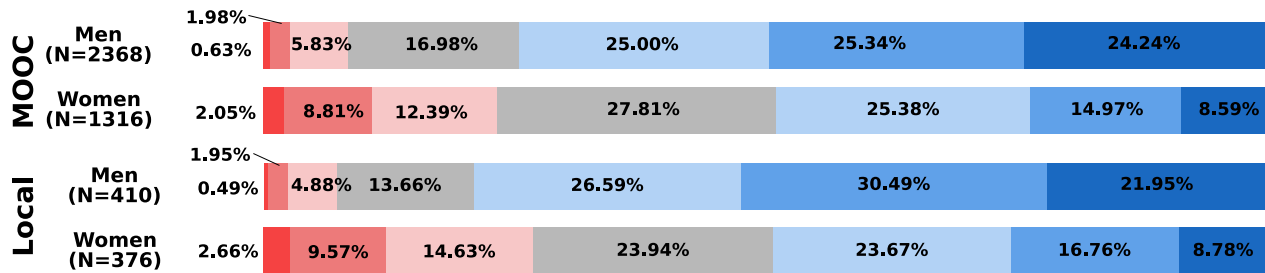


Figure 3: Self-efficacy for the course instances: Students were asked if they agreed with statement ‘I believe I will do well on this course’. Options: strongly disagree (left), disagree, somewhat disagree, neutral, somewhat agree, agree, and strongly agree (right).

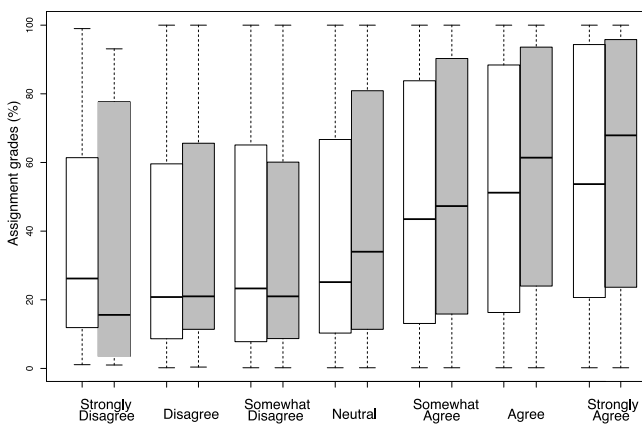


Figure 4: Box plot of assignment grades (ranging from 0 to 100) for men (gray) and women (white) categorized by their answer to the self-efficacy question asking whether the participant expects to do well in the course.

- Interest in CS: Significant difference between genders. Small effect size across the board.
- Performance: Significant difference between genders in MOOCs, small effect size.
- Interactions: Self-efficacy is a significant effect on MOOCs. Interaction between gender and interest is significant in local courses. All partial correlations, even when significant, are small.

### 5.1 Confidence Gap and Performance

Overall, we found a considerable gap in confidence between men and women on both versions of the course as seen in Figure 3. These results align with earlier research on self-efficacy and gender in CS [3, 25]. It is also worth noting that men display a similar pattern of high confidence regardless of whether they are participating in the MOOC or are enrolled students participating in the local course. Likewise, the pattern of responses for women is similar on MOOC and local versions.

When it comes to performance, measured using completed assignments, we did observe a significant difference between men and women with a small effect size on the MOOC version. Notably,

this was not the case on the local course where there was no significant difference in performance between men and women, which is aligned with prior research [4, 36, 41]. One possible explanation for the difference in performance on the MOOCs might be the interest in pursuing a CS degree. The difference in interest was significant between men and women on the MOOC version but not significant on the local course.

Finally, we studied the interaction between gender, self-efficacy, interest, and course performance. Within the locally offered course, the only significant interaction was between gender and interest, suggesting that men are more interested in pursuing CS further, which then also reflects on performance. In the MOOC version, the only significant interaction was identified between the self-efficacy and course performance. This implies that, on the MOOC, the main contributor towards the lower performance of women is the lower self-efficacy.

### 5.2 Weekly Drop-out Rates

As seen in Figure 2, both courses have noticeable weekly drop-out rates. The higher drop-out rates in the MOOC version is in line with previous work on MOOCs, which suggests that MOOCs are started by many but finished by few [8]. When considering gender, women are more likely to drop out from the MOOC at the beginning and the midpoint; towards the end of the course, there is no noticeable difference between men and women. As for the similar retention towards the end of the course, it is possible some of it could be explained by e.g. the sunk cost fallacy [30]; as students have already invested plenty of effort to the course, they do not wish to quit. As for the differences in retention at the beginning of the course, we do not know *why* this happens.

When considering the locally offered course, there is almost no difference in terms of drop-out between men and women. From the organizational perspective, the significant difference between the MOOC and the locally offered course is that the local course has lectures and walk-in labs. It is possible that some of the differences could be explained by a sense of belonging; previous research suggests that a sense of belonging may have a stronger relationship with success in women than in men [18] and thus the local community may influence the sense of belonging. Similarly, it is possible that the MOOC and the online support available transmit stereotypical cues, leading to decreased participation in women [6].

### 5.3 Trends in Participation

When studying the participation in the MOOC across the years 2017–2019, we observe a trend towards gender parity (Figure 1). While in 2017, approximately 22% of the participants at the beginning of the course were women, in 2019, approximately 39% of the participants at the beginning of the course were women. There is also an increase in terms of persistence – at the end of the 7th week of the MOOC, in 2017 16% of the participants were women, while the corresponding number for 2019 is 32%.

When comparing these results with previously reported numbers on participation in a programming MOOC organized in Finland, there is a substantial improvement. In an article published in 2015, Kurhila and Vihavainen reported that 5.2% of participants in a programming MOOC identified themselves as women [23]. At the same time, when the data in the above-mentioned article was collected, there were significantly fewer opportunities for CS studies in primary and secondary education. In general, one could state that the recent efforts in making CS available for all have also been visible in Finland, and those efforts may have influenced the participation rate in the MOOC under study. Our result further suggests that studies that use participant background data should be re-evaluated periodically with data on new participants as the backgrounds of participants may change over time.

### 5.4 Limitations

Here, we discuss some of the limitations of our work. First, there is a possibility of sampling bias. As discussed in Section 4, only a subset of all participants were included in the study (57% of the whole population). Moreover, the data comes from a specific course offered as a MOOC and as a local version organized in a specific country. We do not claim that the presented results would hold in other contexts or countries. On the contrary, we are hoping that others perform similar analyses on their courses.

Second, we did not study the reasons for attending the courses. Some students in the MOOC are applying for a study right, some are participating in the course to e.g. attain a high-school course, and some are participating in the course for fun. Similarly, some students in the local course are studying CS as a major, some CS as a minor, and some are considering whether they would like to take CS as a major or a minor. We only considered the courses and did not further distinguish between the subpopulations; doing so could have provided additional insight into the topic.

Third, when comparing our results with other similar studies, we reported several statistically significant results, even though most of the effect sizes were small to negligible. One of the reasons for this may be the higher number of participants in this study when compared to many of the existing related studies: simply put, statistically significant results are easier with a larger sample size.

Finally, our study does not address issues related to the double bind. One of the issues affecting research on double bind in CS is that basic data is not collected [29]. Unfortunately, the background survey given at the start of the course did not include questions on students' race or ethnicity, and therefore, we could not investigate potential issues with the double bind.

## 6 CONCLUSIONS

In this work, we studied gender and persistence in introductory programming courses offered both locally and as MOOCs.

Our answer to our research question, *How does gender interact with taking a local or a MOOC version of an introductory programming course?*, is as follows. First, women are more likely to drop out from the MOOC than men. Within the local course, the interaction between gender and drop-out is subtler. Second, the confidence gap is visible within both the MOOC and the locally offered course, suggesting that women are less confident in their performance in the course (even if there is no real difference in the actual performance). Third, men are more likely to be interested in CS than women, and fourth, men are more likely to have prior programming experience when entering the MOOC. Finally, there is no meaningful difference in terms of performance (completion of course assignments) in the locally offered course, while in the MOOC, there is a significant albeit very weak difference in performance in favor of men completing more course assignments. The majority of the differences, while statistically significant, were typically small.

When analyzing interactions between the variables and course performance, no statistically significant factors were found except the self-efficacy in the MOOC and the interaction between gender and interest in CS in the local course. Further, when analyzing trends in MOOC participation, we observed that there is a clear increase in women starting the programming MOOC (an increase from 22% to 38% during 2017-2019, nearly doubling the number of women who start the course). This suggests that there is a clear increase in interest in learning to program. We also discussed possible reasons for this change, including the recent push towards making CS accessible for all. To our knowledge, there have been no significant efforts towards making the course better for women, but there has been continuous course improvement in terms of making the course better for all participants.

We acknowledge the fact that this article has been written by an "all-male panel" with little insight into the issues related to researching and reporting gender-related topics in STEM fields. Our research solely points out a problem but does not provide directions to remedy it. The studied course is an introductory level course with no background requirements, and the data was collected before students had worked on the course. Thus, we have no recommendations on what could be done to alleviate the situation within tertiary education.

As a part of our future work, we are looking into interviewing students in future MOOCs to determine factors that contribute to students dropping out from the course. While such work has been previously conducted within locally offered programming courses [22, 31], it is meaningful to study whether the causes generalize to MOOCs and whether some of the causes leading to students dropping out could be remedied. We are also looking into the difference between students who apply to study CS through the MOOC and students who attend the MOOC but do not apply to study CS.

### Acknowledgments

This research has been partially supported by the grant 201365/2015-4 provided by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## REFERENCES

- [1] Anant Agarwal. 2013. Online universities: it's time for teachers to join the revolution. Available online: <https://www.theguardian.com/education/2013/jun/15/university-education-online-mooc>.
- [2] David N Beede, Tiffany A Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E Doms. 2011. Women in STEM: A gender gap to innovation. *Economics and Statistics Administration Issue Brief* 04-11 (2011).
- [3] Sylvia Beyer, Kristina Rynes, Julie Perrault, Kelly Hay, and Susan Haller. 2003. Gender Differences in Computer Science Students. In *Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education (SIGCSE '03)*. ACM, New York, NY, USA, 49–53. <https://doi.org/10.1145/611892.611930>
- [4] Pat Byrne and Gerry Lyons. 2001. The effect of student attributes on success in programming. *Acm sigcse bulletin* 33, 3 (2001), 49–52.
- [5] Scott Carlson. 2006. Wanted: Female Computer-Science Students. *Chronicle of Higher Education* 52, 19 (2006).
- [6] Sapna Cheryan, Victoria C Plaut, Paul G Davies, and Claude M Steele. 2009. Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of personality and social psychology* 97, 6 (2009), 1045.
- [7] Gayle Christensen, Andrew Steinmetz, Brandon Alcorn, Amy Bennett, Deirdre Woods, and Ezekiel Emanuel. 2013. The MOOC phenomenon: who takes massive open online courses and why? Available at SSRN 2350964 (2013).
- [8] Doug Clow. 2013. MOOCs and the funnel of participation. In *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 185–189.
- [9] J McGrath Cohoon. 2001. Toward improving female retention in the computer science major. *Commun. ACM* 44, 5 (2001), 108–114.
- [10] Joanne McGrath Cohoon, Sergey Nigai, Joseph Kaye, et al. 2011. Gender and computing conference papers. *Commun. ACM* 54, 8 (2011), 72–80.
- [11] Barbara Ericson, Mark Guzdial, Briana Morrison, Miranda Parker, Matthew Moldavan, and Lekha Surasani. 2015. An eBook for teachers learning CS principles. *ACM Inroads* 6, 4 (2015), 84–86.
- [12] Barbara J Ericson, Miranda C Parker, and Shelly Engelman. 2016. Sisters Rise Up 4 CS: Helping female students pass the advanced placement Computer Science A exam. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 309–314.
- [13] Cynthia E Foor, Susan E Walden, and Deborah A Trytten. 2007. "I wish that I belonged more in this whole engineering group:" Achieving individual diversity. *Journal of Engineering Education* 96, 2 (2007), 103–115.
- [14] Sarah Friedrich, Frank Konietzschke, Markus Pauly, et al. 2017. GFD: An R package for the analysis of general factorial designs. *Journal of Statistical Software* 79, 1 (2017), 1–18.
- [15] Jeffrey A Greene, Christopher A Oswald, and Jeffrey Pomerantz. 2015. Predictors of retention and achievement in a massive open online course. *American Educational Research Journal* 52, 5 (2015), 925–955.
- [16] Amanda L Griffith. 2010. Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review* 29, 6 (2010), 911–922.
- [17] Shuchi Grover and Roy Pea. 2013. Computational thinking in K–12: A review of the state of the field. *Educational researcher* 42, 1 (2013), 38–43.
- [18] Bonnie M Hagerty, Reg A Williams, James C Coyne, and Margaret R Early. 1996. Sense of belonging and indicators of social and psychological functioning. *Archives of psychiatric nursing* 10, 4 (1996), 235–244.
- [19] Fredrik Heintz, Linda Mannila, and Tommy Färnqvist. 2016. A review of models for introducing computational thinking, computer science and computing in K-12 education. In *2016 IEEE Frontiers in Education conference (FIE)*. IEEE, 1–9.
- [20] Kate S Hone and Ghada R El Said. 2016. Exploring the factors affecting MOOC retention: A survey study. *Computers & Education* 98 (2016), 157–168.
- [21] Peter Hubwieser, Michail N Giannakos, Marc Berges, Torsten Brinda, Ira Diethelm, Johannes Magenheimer, Yogendra Pal, Jana Jackova, and Egle Jasute. 2015. A global snapshot of computer science education in K-12 schools. In *Proceedings of the 2015 ITiCSE on working group reports*. ACM, 65–83.
- [22] Päivi Kinnunen and Lauri Malmi. 2006. Why students drop out CS1 course?. In *Proceedings of the second international workshop on Computing education research*. ACM, 97–108.
- [23] Jaakko Kurhila and Arto Vihavainen. 2015. A Purposeful MOOC to Alleviate Insufficient CS Education in Finnish Schools. *Trans. Comput. Educ.* 15, 2, Article 10 (April 2015), 18 pages. <https://doi.org/10.1145/2716314>
- [24] Juho Leinonen, Petri Ihanola, Antti Leinonen, Henrik Nygren, Jaakko Kurhila, Matti Luukkainen, and Arto Hellas. 2019. Admitting Students Through an Open Online Course in Programming: A Multi-year Analysis of Study Success. In *Proceedings of the 2019 ACM Conference on International Computing Education Research (ICER '19)*. ACM, New York, NY, USA, 279–287. <https://doi.org/10.1145/3291279.3339417>
- [25] Mary A Lundeberg, Paul W Fox, Amy C Brown, and Salman Elbedour. 2000. Cultural influences on confidence: Country and gender. *Journal of Educational Psychology* 92, 1 (2000), 152.
- [26] Linda Mannila, Valentina Dagiene, Barbara Demo, Natasa Grgurina, Claudio Mirolo, Lennart Rolandsson, and Amber Settle. 2014. Computational thinking in K-9 education. In *Proceedings of the working group reports of the 2014 on innovation & technology in computer science education conference*. ACM, 1–29.
- [27] Catherine Mavriplis, Rachel Heller, Cheryl Beil, Kim Dam, Natalya Yassinskaya, Megan Shaw, and Charlene Sorensen. 2010. Mind the gap: Women in STEM career breaks. *Journal of technology management & innovation* 5, 1 (2010), 140–151.
- [28] Laurie Murphy, Brad Richards, Renée McCauley, Briana B Morrison, Suzanne Westbrook, and Timothy Fossum. 2006. Women catch up: gender differences in learning programming concepts. In *ACM SIGCSE Bulletin*, Vol. 38. ACM, 17–21.
- [29] Maria Ong, Carol Wright, Lorelle Espinosa, and Gary Orfield. 2011. Inside the double bind: A synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics. *Harvard Educational Review* 81, 2 (2011), 172–209.
- [30] Roch Parayre. 1995. The strategic implications of sunk costs: A behavioral perspective. *Journal of Economic Behavior & Organization* 28, 3 (1995), 417–442.
- [31] Andrew Petersen, Michelle Craig, Jennifer Campbell, and Anya Tafilovich. 2016. Revisiting Why Students Drop CS1. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research (Koli Calling '16)*. ACM, New York, NY, USA, 71–80. <https://doi.org/10.1145/2999541.2999552>
- [32] Francisco O. Ramirez and Naejin Kwak. 2015. *Women's Enrollments in STEM in Higher Education: Cross-National Trends, 1970–2010*. Springer International Publishing, Cham, 9–49. [https://doi.org/10.1007/978-3-319-08629-3\\_2](https://doi.org/10.1007/978-3-319-08629-3_2)
- [33] Saif Rayyan, Daniel T Seaton, John Belcher, David E Pritchard, and Isaac Chuang. 2014. Participation and Performance in 8.02 x Electricity and Magnetism: The First Physics MOOC from MITx. 2013. In *Physics Education Research Conference Proceedings*. 289–292.
- [34] Robert Rosenthal and Donald B Rubin. 2003. equivalent: A simple effect size indicator. *Psychological methods* 8, 4 (2003), 492.
- [35] Linda J Sax, Kathleen J Lehman, Jerry A Jacobs, M Allison Kanny, Gloria Lim, Laura Monje-Paulson, and Hilary B Zimmerman. 2017. Anatomy of an enduring gender gap: The evolution of women's participation in computer science. *The Journal of Higher Education* 88, 2 (2017), 258–293.
- [36] Philip R Ventura Jr. 2005. Identifying predictors of success for an objects-first CS1. (2005).
- [37] Arto Vihavainen, Matti Luukkainen, and Jaakko Kurhila. 2012. Multi-faceted support for MOOC in programming. In *Proceedings of the 13th annual conference on Information technology education*. ACM, 171–176.
- [38] Arto Vihavainen, Matti Luukkainen, and Jaakko Kurhila. 2012. Multi-faceted Support for MOOC in Programming. In *Proceedings of the 13th Annual Conference on Information Technology Education (SIGITE '12)*. ACM, New York, NY, USA, 171–176. <https://doi.org/10.1145/2380552.2380603>
- [39] Arto Vihavainen, Matti Paksula, and Matti Luukkainen. 2011. Extreme Apprenticeship Method in Teaching Programming for Beginners. In *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education (SIGCSE '11)*. ACM, New York, NY, USA, 93–98. <https://doi.org/10.1145/1953163.1953196>
- [40] Rebecca Vivian, Katrina Falkner, and Nickolas Falkner. 2014. Addressing the challenges of a new digital technologies curriculum: MOOCs as a scalable solution for teacher professional development. (2014).
- [41] Laurie Honour Werth. 1986. *Predicting student performance in a beginning computer science class*. Vol. 18. ACM.