Understanding Gender Differences to Improve Equity in Computer Programming Education

Jessica McBroom
School of Computer Science,
University of Sydney
Australia
jmcb6755@uni.sydney.edu.au

Irena Koprinska
School of Computer Science,
University of Sydney
Australia
irena.koprinska@sydney.edu.au

Kalina Yacef
School of Computer Science,
University of Sydney
Australia
kalina.yacef@sydney.edu.au

ABSTRACT

The increasing wealth of data available from large online programming courses offers exciting opportunities to improve equity in computer science education. In particular, it can provide insight into potential causes of differences between groups, informing future investigations, course development and educational policies. In this paper, we analyse differences between male and female high school and primary school students across five programming courses of different difficult levels (n=14,570), including block-based and traditional text-based courses. We examine differences in course enrolment, completion rates and student behaviour while solving the exercises, and consider the impact of grade level, enrolment and course type. We find that, while females are under-represented in the courses, performance differences only occur in specific circumstances, and in both directions. This suggests that the differences between male and female students are not fundamental, and that there is potential to address them by better understanding these circumstances.

CCS CONCEPTS

• **Social and professional topics** → **Gender**; • **Applied computing** → **Education**.

KEYWORDS

computer science education, gender equity, introductory programming, automated programming tutor, women in computing, student behaviour

ACM Reference Format:
Jessica McBroom, Irena Koprinska, and Kalina Yacef. 2020. Understanding Gender Differences to Improve Equity in Computer Programming Education. In *Proceedings of the Twenty-Second Australasian Computing Education Conference (ACE'20), February 3–7, 2020, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3373165.3373186

1 INTRODUCTION

While interest in Computer Science has been steadily increasing in recent decades, there are still far fewer women than men enrolled in Computer Science courses at university and, consequently, working in the field at large [5]. Numerous studies have analysed why women are not choosing or persisting Computer Science degrees and majors [2, 11–14]. Some of the main reasons include: the continuous perception that Computer Science is a "male" field; the perception that males are more naturally inclined towards computing than females; the females' lack of confidence in ability (self-efficacy); the small number of female role models; and the negative "geek" stereotype about computer scientists as lacking outside interests and personal skills, which females do not identify with. A recent study [3] explored the gender difference in academia by analyzing the perceptions of female postgraduate students and academics in Computer Science, showing that some of the negative perceptions persist even for this group of highly accomplished professionals who have stayed in Computer Science.

A detailed study [7], exploring the relationships between self-efficacy, goal orientation and metacognitive strategies in introductory programming courses by gender identified a self-efficacy feedback loop and found that female students had a different connection between self-efficacy and programming performance than male students: female students are affected by performance feedback early in the course, revising their self-efficacy beliefs earlier, which in turn affects their performance and motivation. This suggests that responses to early failures in Computer Science could be causing female students to disengage.

Since many gender differences have already emerged by the time students are enrolled in university, it is useful to find where these differences first appear. In this paper, we investigate gender differences in primary and high school students enrolled in 5-week online computer programming courses. We analyse student data during the programming courses, at a high level (enrolment) and finer learning progress level rather than analysing surveys to study perceptions as in previous work. In particular, we examine the differences in enrolment rate and student performance, and analyse the relationship between course type, enrolment type and school year level on enrolment and completion rates. We also investigate student progress in more detail in terms of viewed exercises and exercise progress levels (code that was saved, self-run, submitted for marking and that successfully passed all tests), identifying differences between males and females. This is in contrast to previous data-driven studies of student behaviour, which have focused on differences in other contexts (e.g. to predict course dropout and exam scores [1], to evaluate automated feedback [10] or to identify different student populations [8, 9]). Our work contributes towards understanding the gender differences in early Computer Science

ACE'20, February 3–7, 2020, Melbourne, VIC, Australia

ACM ISBN 978-1-4503-7686-0/20/02…$15.00
https://doi.org/10.1145/3373165.3373186

education and developing strategies for early intervention, to improve the learning experience, retention rate and gender balance in Computer Science.

2 DATA

Our data come from a 5-week programming challenge, the 2018 National Computer Science School (NCSS) challenge. This challenge consisted of 5 Python programming courses of varying difficulty levels (newbies, beginners blockly, beginners, intermediate and advanced), which were run online through the Grok Learning platform [6]. Each course involved a series of interleaved notes and programming exercises, released weekly. The advanced course additionally involved weekly programming tournaments. The newbies and beginners blockly courses introduced Python through a block-based environment, while the other courses were text-based. We considered all enrolments in these courses where students had a gender label and attempted Exercise 1 before the exercise deadline (n=14,570). Note that students could enroll in more than one course.

Students participating in the challenge were almost entirely from Australia, but a small number were from other countries too, including New Zealand and the United Kingdom. These included both primary and secondary students with grade levels ranging from Year 1 to Year 13.

3 ENROLMENT RATES

Out of 14,570 enrolments in the NCSS challenge, only 4738 (32.5%) were female. This rate is similar to rates found in previous studies for women entering IT degrees [4]. In order to gain insight into this imbalance, we consider the relationship between enrolment rates and the following factors:

(a) **Course type**. The programming courses varied in difficulty and delivery method (i.e. block or text based). Did these differences affect the rates of female enrolment?
(b) **Teacher or self-enrolment**. Students could be assigned to a course by teachers, or they could enroll themselves. Was the gender distribution different for self-enrolled students compared to teacher-enrolled students?
(c) **School year level**. The year levels of students participating in the challenge ranged from Year 1 to Year 13. Did the year level of a cohort impact on the enrolment gender distribution?

3.1 Course Type

The distribution of male and female student enrolments across the different NCSS challenge courses is given in Table 1. The proportion of female enrolments is not consistent across courses, with enrolment rates ranging from 14% to 39%. In general, there are disproportionately fewer female enrolments in courses requiring previous programming experience (i.e. intermediate and advanced, with 29% and 14% female enrolments respectively), and more female enrolments in block-based courses. In particular, the beginners blockly course, which is identical to the beginners course in all but the block environment, has a higher ratio of female to male students than the beginners course. This suggests that females are more likely than males to enroll in or be enrolled in block-based courses, and also less likely to enroll in more advanced courses.

Table 1: Proportion of male and female enrolments in each course of the 2018 NCSS challenge.

Course	Female Enrolments		Male Enrolments		Total
Advanced	43	14%	258	86%	**301**
Intermediate	1067	29%	2574	71%	**3641**
Beginners	1975	32%	4183	68%	**6158**
Beginners Blockly	686	39%	1053	61%	**1739**
Newbies	967	35%	1764	65%	**2731**
All Courses	**4738**	**33%**	**9832**	**67%**	**14570**

Table 2: Proportion of male and female self and teacher enrolments in the 2018 NCSS challenge.

Enrolment Nature	Female Enrolments		Male Enrolments		Total
Self	2564	33%	5296	67%	**7860**
Teacher	2174	32%	4536	68%	**6710**
All	**4738**	**33%**	**9832**	**67%**	**14570**

3.2 Teacher or Self Enrolment

The proportion of male and female self and teacher enrolments is given in Table 2. A teacher enrolment indicates that a student was enrolled in a course through a teacher's account, whereas a self-enrolment indicates that they enrolled through their own account. It is possible, though, that some teachers may have directed students to enroll in courses without assigning them themselves, meaning that some self-enrolments could be teacher enrolments. While this could lessen apparent differences between the groups, we did not consider this a major issue, since there were many students and any significant trends could not be completely hidden.

From Table 2, overall, there no significant difference between the percentage of teacher enrolments and the percentage of self-enrolments that were female (32% and 33% respectively). This suggests that teacher enrolments were not more gender-balanced than self-enrolments, and that there could be opportunities to increase female participation through classes.

Differences between the groups do emerge, however, when courses are also considered. Figure 1 shows the percentage of self and teacher female enrolments across the different competition courses. Notably, the female self-enrolment rate across the first four courses is quite consistent, with about one third of enrolments being female in all cases. The teacher enrolment rates, however, are different across these courses. In particular, higher proportions of newbies and beginners blockly teacher enrolments are female (38% and 53% respectively) compared to the self-enrolments, with the opposite being true for the beginners and intermediate courses.

It is particularly interesting to consider the beginners and beginners blockly enrolments, since these courses were equivalent in all but the programming environment (text vs block). Since the proportion of female self-enrolments in these courses was consistent, this suggests that female students were just as likely to choose the

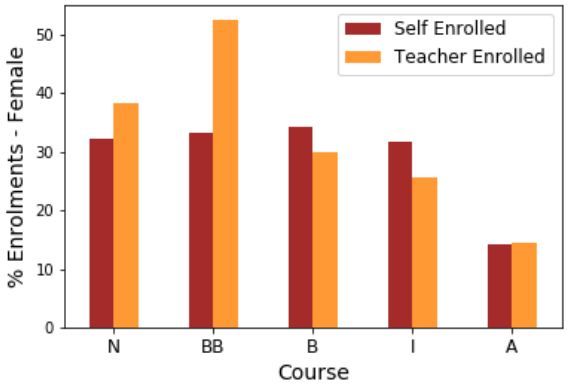

Figure 1: Visualisation of female enrolment rates across courses and based on enrolment type. Each bar shows the percentage of enrolments in each category that are female. For example, more than 50% of teacher-enrolled beginners blockly students are female. (N: newbies, BB: beginners blockly, B: beginners, I: intermediate, A: advanced).

text-based courses as males when deciding themselves. However, the proportion of female teacher enrolments differs across these courses, with more than half the beginners blockly enrolments being female, suggesting that teachers are more likely to enroll females in block-based courses than they are to enroll males. If this is the case, it is important to ensure block and text-based courses lead to equal learning outcomes.

3.3 School Year Level

The grade distribution of enrolments in the NCSS challenge is given in Table 3. The majority of enrolments are from students in Years 5 to 11, with Years 9 and 10 being the largest cohorts[1].

The ratio of male to female enrolments across year levels is visualised in Figure 2. In general, the enrolments tend to be more gender balanced in younger years, and less balanced in later years. This may indicate that females participate less in programming as they become older, or it could indicate that programming classes are becoming more gender balanced, but there has been insufficient time for this to spread to later year groups. It would be interesting to compare this distribution to that of later years to determine which is the case.

4 COMPLETION RATES

In the previous section, the proportions of male and female enrolments in the 2018 NCSS challenge were considered with respect to course type, enrolment type and school year level. While there was an imbalance in the numbers of male and female students enrolling in the courses, we wondered if this imbalance extended to student performance during the course itself. As such, in this section, we

[1]Note that the grading systems of countries differ. For example, Year 10 New Zealand students are generally younger than Year 10 Australian students. However, since almost all students in the dataset are Australian, this is unlikely to have a noticeable impact on the results.

Table 3: Proportion of male and female self and teacher enrolments in the 2018 NCSS challenge.

Year Level	Female Enrolments		Male Enrolments		Total
1	3	38%	5	63%	**8**
2	7	41%	10	59%	**17**
3	46	47%	52	53%	**98**
4	175	61%	111	39%	**286**
5	524	48%	572	52%	**1096**
6	584	47%	646	53%	**1230**
7	835	43%	1119	57%	**1954**
8	570	32%	1217	68%	**1787**
9	941	27%	2514	73%	**3455**
10	735	23%	2505	77%	**3240**
11	221	20%	881	80%	**1102**
12	82	29%	198	71%	**280**
13	4	100%	0	0%	**4**
Unknown	11	85%	2	15%	**13**
All Years	**4738**	**33%**	**9832**	**67%**	**14570**

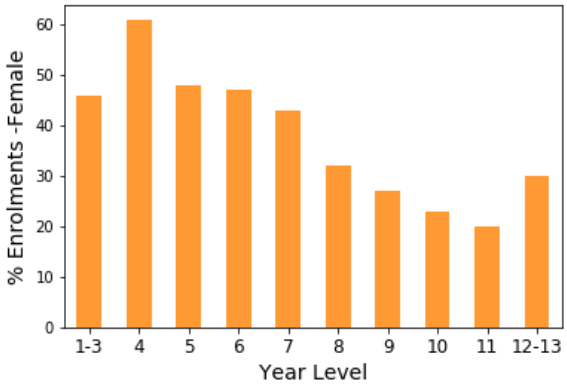

Figure 2: Percentage of enrolments in each year level that are female. Note that Years 1-3 and 12-13 are grouped together due to low numbers.

consider the number of programming exercises completed by male and female students during the challenge to gain insight into how they progressed through the courses.

We focus on all courses except the advanced course from now on, because exercise completion was not always well defined in the advanced course (submissions to tournament exercises could keep being improved) and the other courses had significantly more enrolments to facilitate a deep analysis of differences.

The overall differences in male and female exercise completion rates are shown in Figure 3. Note that these graphs must be interpreted with care, since the male and female populations are distributed differently. For example, the female average age is generally lower than the male average, and the graphs do not account for this difference. Despite this, however, the graphs show some

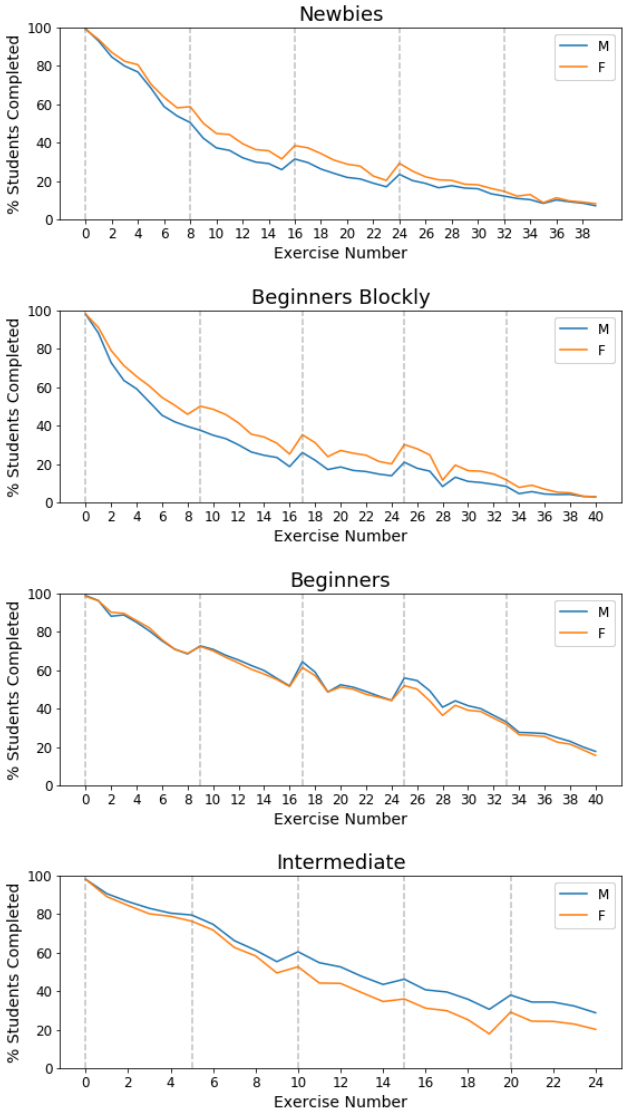

Figure 3: Overall male and female exercise completion rates during the 2018 NCSS challenge courses. The solid lines indicate the percentage of male and female students who completed each exercise. Exercises are ordered by their appearance in the courses, and the dashed lines mark the start of each week.

interesting patterns, which can be useful starting points for further investigation. Namely:

(1) The male and female completion lines in the beginners course and at the start of the intermediate course are almost identical, but they diverge abruptly midway through the intermediate course. Even if there were differences between the male and female populations, it is surprising that this gap would form in one text-based course, but not the other, and only midway through the course.

(2) Females appear to complete more exercises of the block-based courses than males, but this difference vanishes towards the end of the course, with both groups converging to low completion rates overall. Clearly, many students find these courses challenging (which is problematic if females are more likely to be enrolled in them), but is there a reason why females tend to do better, or is this just an artifact of the different population distributions?

In the next sections, we examine these two effects more carefully, controlling for factors such as grade level and enrolment type, to determine whether they can be explained by differences in population distribution.

4.1 Intermediate Course Gap

In this section, we examine more closely the gender gap that appears to form midway through the intermediate course, but not in the beginners course. In particular, we determine whether it still persists after controlling for two important population factors that could potentially explain the effect:

(a) **Enrolment Type**. Of the 1067 female intermediate students, 368 (34%) were teacher-enrolled. This is in contrast to the intermediate males, where 1075 / 2574 (42%) were teacher-enrolled. Since a higher proportion of intermediate males were teacher-enrolled, they may have had more teacher support, explaining the increase in completion rates.

(b) **School Year Level**. Female intermediate students were younger on average than male intermediate students, with male enrolments peaking in Year 10, and female enrolments peaking in Year 9. Since the male students were older, they may have found the exercises easier and thus completed more.

4.1.1 Controlling for Enrolment Type. We divided male and female populations into self and teacher-enrolled groups, then compared the exercise completion rates of the four groups. If the male students had indeed completed more exercises only because more were teacher-enrolled, then we would expect no differences among self-enrolled students, and no differences among teacher-enrolled students.

The comparison between self and teacher-enrolled male and female groups is shown in Figure 4. Clearly, teacher-enrolled groups did indeed complete more exercises than their corresponding self-enrolled groups. However, this does not explain the gender gap, since the gap persists even after controlling for enrolment type.

Interestingly, from this graph, enrolment type has an immediate impact on completion rates, with self and teacher-enrolled students diverging at the start of the course. However, the gender gap, in both enrolments, only occurs from midway through the course, around exercises 9 and 10.

4.1.2 Controlling for School Year Level. We controlled for school year level similarly to enrolment type, by grouping students of similar year levels together. These groups, along with the numbers of enrolments, are shown in Table 4. Note that, since there were few primary school students (Years 1-6), we grouped all of these together. Similarly, we grouped Years 7-8 and 11-13 together to avoid small group sizes.

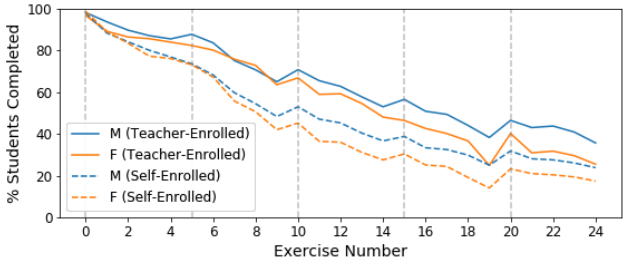

Figure 4: Comparison of intermediate course exercise completion for self and teacher-enrolled male and female students.]

Table 4: Number of intermediate male and female enrolments for each grade bucket.

Year Level	Female Enrolments	Male Enrolments
Primary	88	68
Years 7-8	236	259
Year 9	343	727
Year 10	256	1028
Years 11-13	141	490
Unknown	3	2
All	**1067**	**2574**

Figure 5 shows the completion rates of male and female students, separated by year levels. Similarly to before, if year level were the explanation for the difference between male and female students, we would expect the effect to vanish after controlling for year level. Interestingly, with the notable exception of Year 10, we indeed find little gender difference for most year groups.

4.1.3 Controlling for School Year Level and Enrolment Type. To be certain that the Year 10 pattern could not be explained by enrolment type, we controlled for enrolment type and year level simultaneously, with the results shown in Figure 6. Since the gender gap forms in both self and teacher-enrolled cases, this suggests it cannot be explained by enrolment type.

This pattern in year 10 intermediate is clearly very interesting. In particular, it is surprising that both younger and older year groups behave similarly, irrespective of gender, but this one does not. In addition, it is peculiar that the gap forms midway through the course when other gaps form immediately, such as the gap between self and teacher-enrolled students. We further investigate this effect in Section 5.

4.2 Block-Based Gap and Beginner Patterns

In this section, we examine the previously identified trend of females completing more exercises than males in block-based courses. In addition, we investigate whether a difference between year 10 students and other year groups is also present in the beginners course.

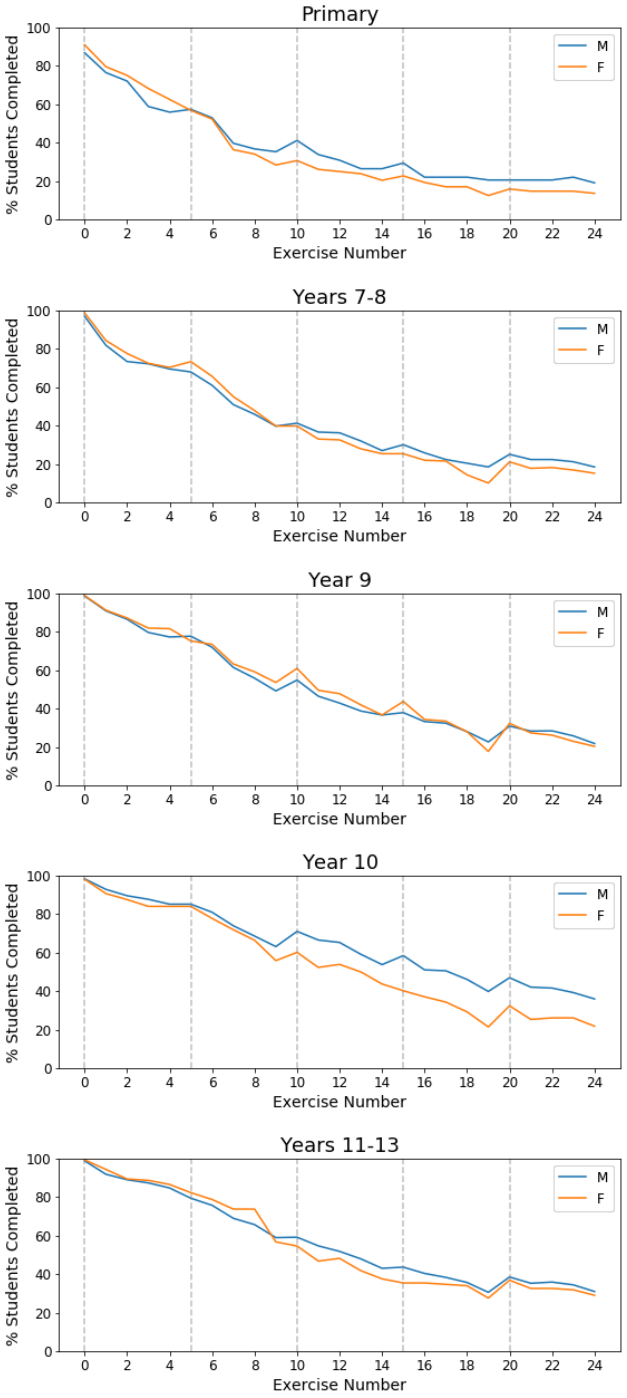

Figure 5: Intermediate course comparison of exercise completion for students in different year levels. There is little difference in all cases except year 10.

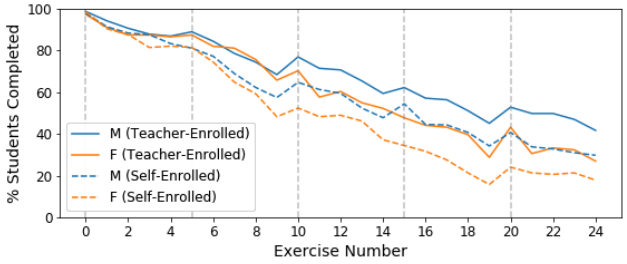

Figure 6: Year 10 intermediate, self and teacher-enrolled female and male exercise completion rates.

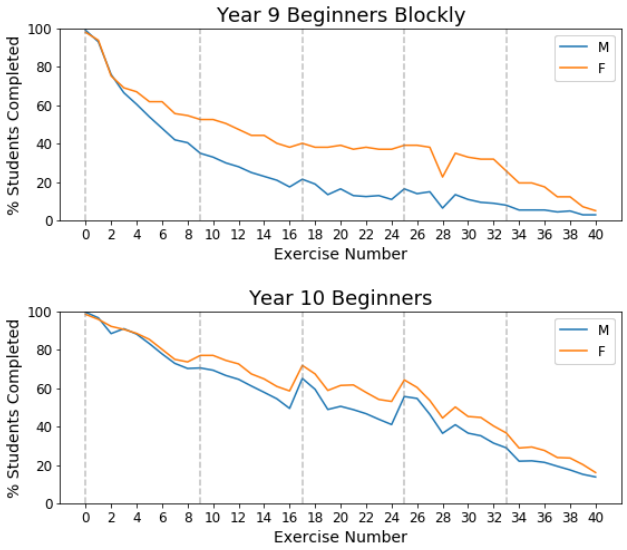

Figure 7: Cases where male and female exercise completion rates differ after controlling for year level (in addition to year 10 intermediate). The graphs show the percentage of students completing each exercise in the year 9 beginners blockly group (top) and year 10 beginners group (bottom).

As in Section 4.1, we controlled for year level by comparing students of similar grades in the newbies, beginners blockly and beginners courses. We found little difference between male and female students across all year groups and courses, with only two exceptions. Specifically, female students in year 9 beginners blockly and year 10 beginners completed noticeably more exercises than the corresponding male groups, as shown in Figure 7.

To explore this further, we took a similar approach to in Section 4.2 by controlling for enrolment type. We found that only the year 10 beginners pattern was present in both self and teacher-enrolled groups (see Figure 8, bottom). In contrast, year 9 teacher-enrolled beginners blockly students completed more exercises than all other groups, as shown in Figure 8, top.

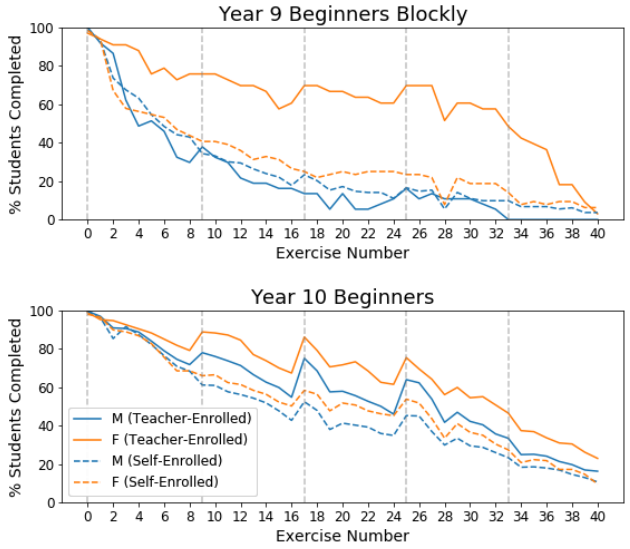

Figure 8: Comparison between self and teacher enrolled male and female students in the year 9 beginners blockly (top) and year 10 beginners (bottom) courses. The legend is shared.

4.3 Overall Differences

These results suggest that, in general, females do not complete more block-based exercises than males. Specifically, when controlling for year level, the male and female newbies students, and most beginners blockly students, were very similar. It was only in a specific circumstance that females clearly completed more exercises: only in year 9 beginners blockly, and only in teacher-enrolled classes.

In addition, the results indicate that females do not complete less exercises than males in the other courses. When controlling for year level and enrolment type, we found that almost every male and female group completed a similar number of exercises. It was only in year 10 that we observed differences: females completed more beginner exercises and males completed more intermediate exercises.

This suggests that, overall, there is little difference between male and female school students when completing programming exercises. In particular, differences only appear in very specific situations. In the next section, we will examine two of these situations more closely: year 10 beginners, and year 10 intermediate. While the third situation (year 9 beginners blockly) is also interesting and worth further investigation, the pattern does not apply to self-enrolled females, suggesting it could be due to class-related factors, which we unfortunately do not have the data to investigate further.

5 INVESTIGATING YEAR 10 BEGINNER AND INTERMEDIATE COMPLETION DIFFERENCES

In the previous section, we identified that in almost all cases, female and male programming students completed a similar number of exercises. However, one important exception to this was the year

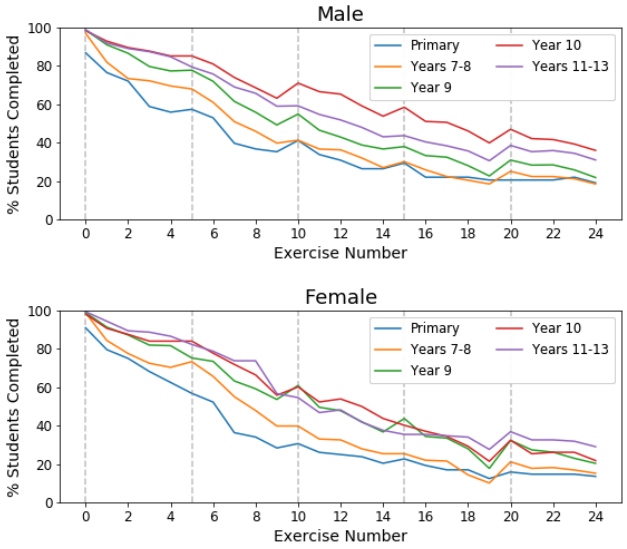

Figure 9: Relationship between year level and exercise completion in the intermediate course for female and male students. The year 10 male students complete more exercises than all other male groups.

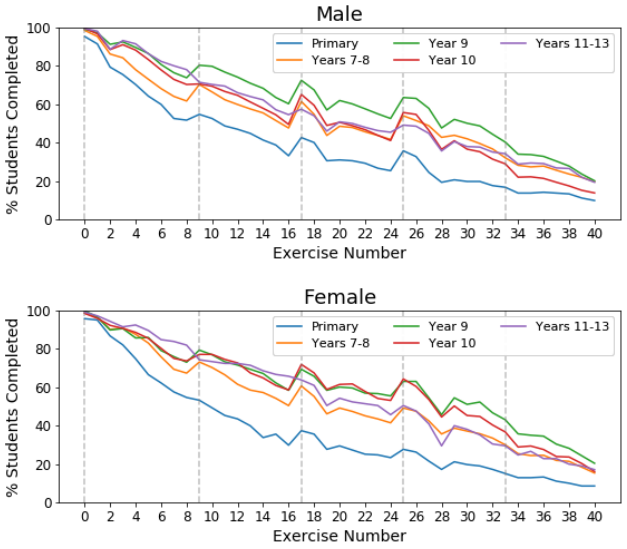

Figure 10: Relationship between year level and exercise completion in the beginners course for female and male students.

10 beginner and intermediate students. Specifically, for beginner, year 10 female students completed more exercises than year 10 male students, with the converse being true for intermediate. In this section, we investigate this situation further to gain insight into potential reasons for this difference.

5.1 Relative Performance

When investigating why the year 10 students behaved differently, it is useful to consider their performance relative to other cohorts of the same gender. Specifically, we can observe whether some groups appear to have overperformed or underperformed relative to these cohorts. For example, perhaps the year 10 female students completed more beginner exercises because they performed unexpectedly well and this motivated them to do more or, in contrast, because the year 10 males completed less exercises than expected.

In Figure 9, we compare the intermediate exercise completion rates of year 10 male and female students with all other students of the same gender. While the completion rates of females appear to roughly increase with age, the same is not true for the male students. Indeed, the year 10 males complete more exercises than every other male group, including older males. This suggests that the difference between male and female intermediate students could be caused by the males in this particular group overperforming, rather than the females underperforming.

To determine whether the difference in beginners could also be explained in this way, we repeated the process for the beginners course, with the results shown in Figure 10.

Interestingly, in both of these graphs, the year 9 students appear to complete the most exercises. In the male case, however, the year 10 students complete significantly less exercises than the year 9 students, while the completion rate is much closer for the year 10 female students. This suggests that the year 10 males could be underperforming relative to the other year groups with the same not being true for females, leading to the gender gap.

5.2 Relative Enrolment Numbers

In the previous section, we found that the gender gaps in year 10 beginners and intermediate may have been caused by year 10 males performing differently from expected. In particular, year 10 males seemed to overperform in intermediate and underperform in beginners relative to other male year groups. In this section, to better understand why this might be the case, we explore another difference between year 10 males and other male groups: their relative enrolment numbers.

One interesting observation from the graphs in the previous section is that older student groups do not always complete more exercises than younger groups. Indeed, we found multiple cases where younger students completed more exercises than older groups:

(a) Year 10 intermediate males completed more exercises than all other male groups, including years 11-13.
(b) Year 9 males in beginners completed more exercises than all other male groups, including year 10 and years 11-13.
(c) Year 9 and 10 females in beginners completed a similar number of exercises, with both completing more than years 11-13 females.

These cases are interesting because one would generally expect student ability to increase with age, allowing students in higher year levels to more easily complete exercises.

One possible explanation for why younger students sometimes complete more exercises than older students could be that the older students have more school-related commitments. For example, students in years 11-13 may have had less time for the challenge

because they were preparing for end-of-school exams. However, this does not explain why we observe different patterns in different courses. Indeed, if year 10 males completed less beginner exercises because they had more school commitments, we would not expect them to complete the most intermediate exercises.

To further understand these patterns, we investigated whether there was any relationship between exercise completion rates and enrolment rates. A relationship between such factors could help to provide further insight into the gender gaps in the year 10 beginners and intermediate courses.

Figure shows the grade distribution of male and female students in the intermediate and beginners courses. Interestingly, there appears to be a relationship between the groups completing the most exercises and their relative numbers of enrolments:

(a) Year 10 males, who completed the most exercises among all intermediate male groups, also had the highest proportion of intermediate enrolments.
(b) Year 9 males, who completed more exercises than all other male groups in beginners, also had the highest proportion of enrolments in beginners
(c) Year 9 and 10 females in beginners, who completed a similar number of exercises, had a similar number of enrolments. In addition, they both completed more exercises and had more enrolments than the years 11-13 female group. (Note, though, that there were also a high proportion of year 7-8 enrolments, but this group did not show the same patterns.)

From these, it seems that there is some relationship between the male and female groups with the highest exercises completion rates, and their relative proportions of enrolments in the courses.

One possible explanation for the correspondence between exercise completion and enrolment rates could be that they are both artifacts of student motivation. For example, perhaps year 9 male students are the male group most interested in the beginners course, so they have the highest enrolment rate and completion rate among all other males.

Another possible explanation for the relationship could be that year groups with high enrolment rates tend to be comprised of students trying the course for the first time, while older year groups have a mix of new and previous students. For example, if many male students try the beginners course in year 9, then by year 10, the strongest ones will have moved on to the intermediate course, while the weaker ones may have stayed to repeat the beginners course. This would create a situation where year 9 students are on average stronger than year 10 students, and thus could explain why they complete more exercises.

While it is difficult to draw concrete conclusions about these relationships without further information, it seems likely that the gender gaps in the year 10 beginners and intermediate courses are not independent of the year 10 enrolment rates. In future, it would be useful to more deeply investigate these factors with data from consecutive years to gain further insight into the relationship.

5.3 Behavioural Patterns

In the previous sections, we found that the year 10 gender gap appear to be caused by males overperforming in the intermediate course and underperforming in the beginners course. In addition,

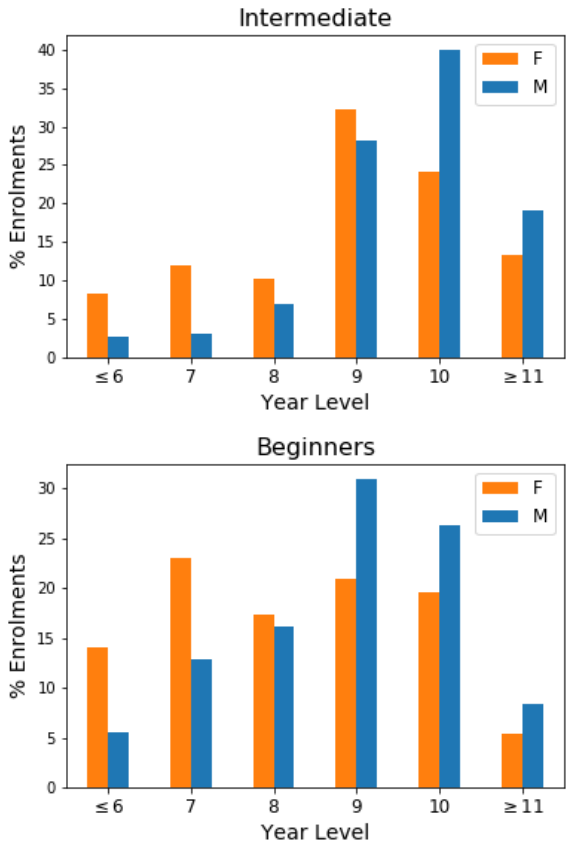

Figure 11: Intermediate (top) and beginners (bottom) male and female enrolment distribution.

this effect appeared to be related to the relative enrolment numbers of these male groups. However, this did not explain why:

(1) the gender gap in the immediate course only formed midway through the course. This is particularly odd, since the gap between self and teacher-enrolled students formed immediately
(2) the beginner gap closed while the intermediate gap did not

In this section, we examine more detailed student behaviour throughout the courses to better understand these differences.

For each exercise, we determined how much progress students made towards solving exercises. When completing exercises, students could run their own code to test it, and could also submit the code for marking after self-testing. Their work was also periodically autosaved. We used this information to measure student progress on the exercises, by calculating the percentage of students who passed each exercise, submitted their work for marking, ran their code, had their work autosaved and, finally, viewed the exercise. Note that each progress level was a subset of the next one so, for example, all students who submitted their work also tested their code. The results are shown in Figure 12.

One interesting difference between the male and female beginners graphs in Figure 12 is that the female graph is fairly smooth at

the beginning, while the male graph has an irregularity at exercise 2, with less students viewing this exercise than surrounding exercises (see location A in the figure). Even if exercise 2 were more difficult than the surrounding exercises (which it did not seem to be from manual inspection), it is surprising that some male students were entirely missing this exercise.

When we viewed the course from the perspective of a student account, we found that this exercise was positioned slightly differently from other exercises in the course. Specifically, in the course navigation pane, there were various topic headings, which were generally followed by a series of note page links, then an exercise link at the end. As such, most exercises stood out in the navigation pane. However, the link to exercise 2 was positioned between a group of note pages, with exercise 3 coming at the end instead. If students were reading through all of the note pages before attempting each exercise, then they would have been very unlikely to miss seeing this exercise. However, if the students were skipping past note pages using the navigation pane, then it would have been much easier to miss seeing this exercise. This suggests that some of the year 10 male students may not have been reading some of the course notes, potentially resulting in more difficulty later in the course. This could explain some of the observed gender gap.

Another interesting difference between the male and female graphs is in the intermediate course, where there are spikes in the percentage of male exercise views at the starts of weeks 3 and 4 (location B), but not in the female views. Note that, unlike in the previous case, the exercises before these were not easy to miss. In addition, the view increases came at the start of new weeks, not in the middle. This suggests that some male students chose not to attempt some exercises, possibly because they lost interest or were too busy, but were happy to try again at the start of the next week. However, this did not seem to occur for intermediate female students.

Focusing on location C in Figure 12, another interesting difference between the female and male intermediate students is that the female students appeared to have slightly more difficulty at the very beginning of the course. Notice that most female students who viewed the starting exercises also wrote and ran some code, since there is little difference between the run and view lines. However, some of these students did not submit their program for marking, since there is a gap between the run and submit lines. Since this gap is much smaller in the corresponding male graph, this suggests the female students may have been experiencing more difficulty from early on.

These graphs also suggest why the intermediate gender gap seemed to form midway through the course, instead of at the beginning. Notice in the exercises leading up to the split (location D) that both male and female students seemed to experience difficulty, since there is a gap between the submit and run lines. This is the time when students began learning about loops, so perhaps the gap formed here due to a change in course difficulty. This does not, however, explain why the gap between self and teacher enrolled students formed immediately – perhaps the intermediate gender gap was related more to student motivation than course difficulty, in line with the findings of Lishinski et al. [7] who found that male and female students revise their self-efficacy differently in response

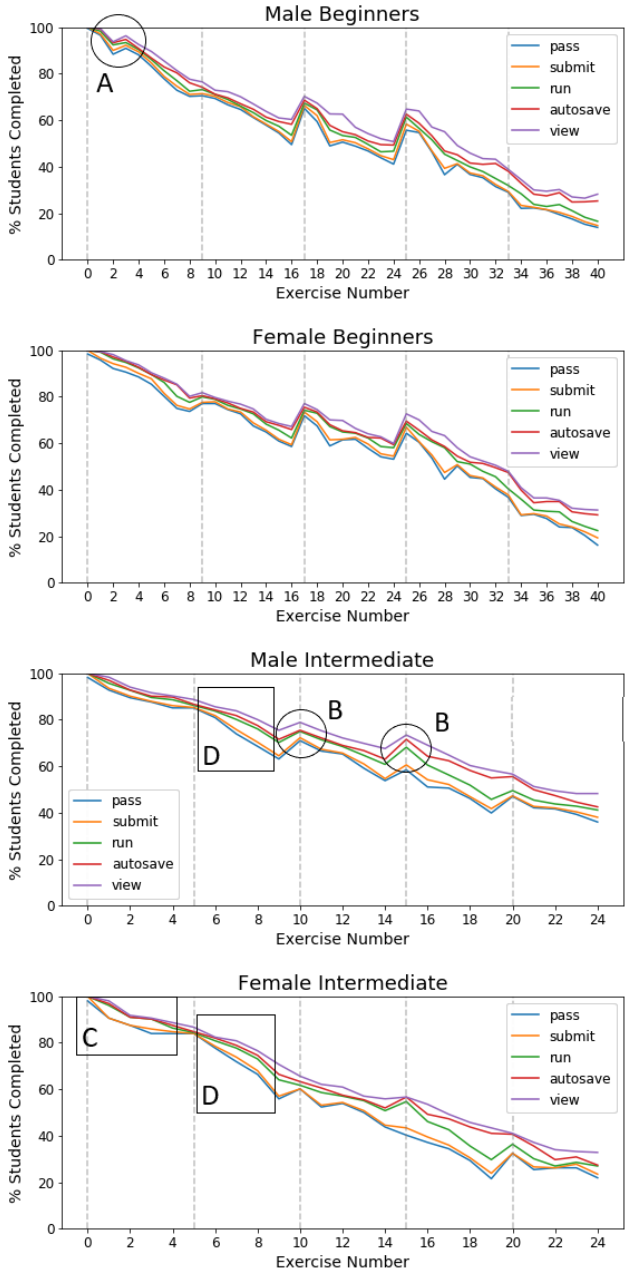

Figure 12: Differences in male and female exercise progress in the beginners and intermediate courses. For each exercise, the lines show the percentage of students who passed the exercise, submitted their program for marking, self-tested their program, had their work autosaved and viewed the exercise. Letters A-D refer to points of interest discussed in the text.

to early failures in computer programming, which may cause more female students to disengage.

One further observation from these graphs is that almost all students who submitted their code for marking also ended up solving the exercise. This suggests that any intervention targeted at reducing the gender gap would need to come before students submit their code.

In summary, the differences in student progress through the course exercises provide some insight into the gender gaps in these courses. In particular, it seems male students may not have been reading as many notes in the beginners course as female students, potentially contributing to the gap. In addition, females seemed to experience difficulty from the very beginning of the intermediate course, which then became more difficult in the second week, and were potentially more concerned about missing exercises than male students. While this does not explain why the groups behave differently, it does suggest the potential for early interventions to help bridge these gaps, particularly by targeting students before they submit their code.

6 CONCLUSION

In this paper, we have investigated differences between male and female school students across five programming courses of different difficult levels (n=14,570). The most significant difference was in the number of enrolments, with twice as many male enrolments as female. In general, this difference was largest in more advanced courses and as school year level increased. In addition, teachers were more likely to enroll female students in block-based courses than female students were to self-enroll.

With respect to exercise completion rates, we found very little difference between male and female students, except in very specific circumstances. In particular, a gender gap formed in the year 10 beginner and intermediate courses, with males completing more intermediate exercises and females completing more beginners exercises. In addition, female teacher-enrolled students in year 9 completed more exercises in the beginners blockly course than other year 9 students. Since there was little difference in all other cases, this suggests that there is no fundamental difference in the programming ability of female and male students, but that gender gaps can arise in particular environments. As such, to improve equity in programming education, it is important to identify and further investigate these situations to understand and address potential causes of gender gaps.

In future, we intend to investigate whether the identified gender gaps were specific to the analysed year groups by observing whether they persist or change over time. In addition, we hope to more closely study the behaviour of students in groups with gender gaps to gain further insight into potential causes. Such study will assist in ongoing work to overcome disadvantage in computer science education and reduce the gender gap.

ACKNOWLEDGMENTS

We would like to thank Grok Learning [6] for the provision of data.

REFERENCES

[1] Michael Mogessie Ashenafi, Giuseppe Riccardi, and Marco Ronchetti. 2015. Predicting students' final exam scores from their course activities. In *2015 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9. https://doi.org/10.1109/FIE.2015.7344081

[2] Sylvia Beyer. 2014. Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education* 24, 2-3 (July 2014), 153–192. http://www.tandfonline.com/doi/abs/10.1080/08993408.2014.963363

[3] Katrina Falkner, Claudia Szabo, Dee Michell, Anna Szorenyi, and Shantel Thyer. 2015. Gender gap in academia: perceptions of female computer science academics. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*. ACM, 111–116. https://doi.org/10.1145/2729094.2742595

[4] Judy Kay and Jenny Edwards. 2001. A sorry tale: a study of women's participation in IT higher education in Australia. *Journal of Research and Practice in Information Technology* 33, 4 (2001), 329–335.

[5] Maria Klawe, Telle Whitney, and Caroline Simard. 2009. Women in Computing—take 2. *Commun. ACM* 52, 2 (Feb. 2009), 68–76. https://doi.org/10.1145/1461928.1461947

[6] Grok Learning. 2019. https://groklearning.com

[7] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. 2016. Learning to program: Gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. In *Proceedings of the 2016 ACM Conference on International Computing Education Research*. ACM, 211–220. https://doi.org/10.1145/2960310.2960329

[8] Jessica McBroom, Bryn Jeffries, Irena Koprinska, and Kalina Yacef. 2016. Exploring and Following Students' Strategies When Completing Their Weekly Tasks.. In *Proceedings of the 9th International Conference on Educational Data Mining*. International Educational Data Mining Society, 609–610. http://www.educationaldatamining.org/EDM2016/proceedings/paper_195.pdf

[9] Jessica McBroom, Bryn Jeffries, Irena Koprinska, and Kalina Yacef. 2016. Mining Behaviours of Students in Autograding Submission System Logs.. In *Proceedings of the 9th International Conference on Educational Data Mining*. International Educational Data Mining Society. http://www.educationaldatamining.org/EDM2016/proceedings/paper_172.pdf

[10] Jessica McBroom, Kalina Yacef, Irena Koprinska, and James R Curran. 2018. A Data-Driven Method for Helping Teachers Improve Feedback in Computer Programming Automated Tutors. In *the 19th International Conference on Artificial Intelligence in Education*. Springer, Springer International Publishing, Cham, 324–337. https://doi.org/10.1007/978-3-319-93843-1_24

[11] Phoenix Moorman and Elizabeth Johnson. 2003. Still a stranger here: Attitudes among secondary school students towards computer science. In *ACM SIGCSE Bulletin - Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education*, Vol. 35. ACM, 193–197. https://doi.org/10.1145/961290.961564

[12] Ilias O Pappas, Trond Aalberg, Michail N Giannakos, Letizia Jaccheri, Patrick Mikalef, and Guttorm Sindre. 2016. Gender Differences in Computer Science Education: Lessons Learnt from an Empirical Study at NTNU.. In *Proceedings of the 2016 Norsk Informatikkonferanse*.

[13] Eileen M Trauth, Susan H Nielsen, Liisa A Von Hellens, et al. 2003. Explaining the IT gender gap: Australian stories for the new millennium. *Journal of Research and Practice in Information Technology* 35, 1 (2003), 7.

[14] Brenda Cantwell Wilson. 2002. A study of factors promoting success in computer science including gender differences. *Computer Science Education* 12, 1-2 (2002), 141–164. https://doi.org/10.1076/csed.12.1.141.8211