Assessing Gender Bias in the Software Used in Computer Science and Software Engineering Education

Lyndsey O'Brien, Tanjila Kanij and John Grundy

[a]HumaniSE Lab, Faculty of Information Technology, Monash University, Victoria, Australia

Abstract

Women are underrepresented in Computer Science (CS)/ Software Engineering (SE) and other technology related degrees. As undergraduates, they are also less likely to persist with CS/SE studies than men enrolled in those same courses. Gender correlated differences in personal characteristics, behaviour, and preferences mean that course design decisions may introduce unintended bias. To address this issue, we drew inspiration from the GenderMag method. GenderMag uses personas with evidence-based gender differences in problem-solving traits to detect usability issues in software. In this paper we investigate the personal qualities of CS and SE students, and how these influence their CS/SE learning journey. A series of persona development workshops were held to gather an extensive and unique qualitative dataset capturing the prior experiences, preferences, learning styles, motivations, goals, frustrations, and constraints of CS/SE students. Gender differences were used to construct preliminary male and female student personas. These personas were used in cognitive walkthroughs of software applications commonly used in education, and their performance compared to GenderMag's Tim and Abi. While the student personas were less effective and lacked specificity compared to Abi, they were able to identify issues not detectable with GenderMag. Furthermore, the findings show the utility of persona development workshops as a data collection method and introduce a comprehensive list of CS/SE student qualities that may inspire future investigations.

Keywords:
Gender inclusivity, education, persona, GenderMag, computer science, software engineering, software usability

1. Introduction

The traditional belief of "male being strong and rational" and "female being weak and emotional" has led to a gender disparity within Science, Technology, Engineering and Mathematics (STEM) [1]. There have, however, been advances towards gender parity within many STEM fields [2]. These changes are also realised in Australia, (15% enrolments in 2015 compared to 19% in 2021 in Engineering degrees) [3]. Enrolment to computing degrees has also seen some increment - 16% in 2015 compared to 21% in 2015 [3]. However, the figures indicate that women continue to comprise only one-fifth of enrolments in computing related degrees compared to men [3]. This figure comes from the STEM Equity Monitor, an annual report on the enrolment and course completion rates of women in Australian STEM degrees. Computer science (CS), Information Systems (IS), and Information Technology (IT) degrees are all broadly grouped as an information technology qualification [4]. However, CS itself is a wide area of study that also encompasses subjects like artificial intelligence, databases, and software engineering [5], and it is often a pathway for students into these more specific computing disciplines.

While lower enrolment rates are a major reason for women's under-representation in CS, they are not the only reason. In addition to lower enrolment rates, undergraduate women are also less likely to complete their CS studies than men enrolled in these courses [6], a troubling trend that has also been observed in both the United States [7] and Europe [8]. It has been suggested that this disparity is due, at least in part, to the presence of gender bias within CS courses [9, 10]. While direct discrimination is an issue, comparison with STEM fields where women are well represented reveals that alone it cannot explain their under-representation in CS [11]. Research into this gender imbalance has also found the presence of masculine cultures [11], negative assumptions about the capabilities of female students [12, 13], and biased course materials [9]. However, women's sense of belonging in CS may also be impacted by more subtle cues in the learning environment [14].

As stated, Software Engineering (SE) is a domain of Computer Science (CS) and is used alternatively in the literature. It is difficult to differentiate the findings in the existing literature between CS and SE education. Moreover, CS is considered a pathway for many SE students. For these reasons, we refer to computer science and software engineering teaching and learning environment as CS/SE in this article.

The decision to undertake a CS/SE course may be influenced by gender correlated factors including personality, values, interests, and computer self-efficacy [15]. Even in the absence of deliberate discrimination or unfairness, failing to account for gender differences in these traits may make CS/SE courses less appealing to women and introduce implicit bias. While learning and interacting with new software applications is an integral part of any CS/SE program, it is also an area where there are well-established gender differences [16]. These are often not considered by software developers who tend to design applications to suit their own needs [17], and this may negatively impact the experience of female students.

Gender Inclusiveness Magnifier (GenderMag) is a highly effective tool for the detection of gender bias within software that is based on five 'facets', or research supported differences in cognitive styles of problem-solving [16, 18]. These facets are: information processing style, learning style, computer self-efficacy, attitude toward risk, and motivations. The GenderMag method employs personas constructed using these facets to identify inclusivity issues in software applications [16]. Personas are fictitious but representative personal profiles that support the consideration of different user groups [19, 20]. GenderMag follows the InclusiveMag approach, a three-stage method for the development of personas tailored to a specific diversity dimension [21].

While the GenderMag personas have already been applied in an education context [22, 23], the facets they employ are based on evidence of gender differences in the general population. CS/SE students are a unique group whose gender expression may not be entirely consistent with this framework. Additionally, previous work looking at education software examined only a single Learning Management System (LMS) [22]. Using the InclusiveMag method it may be possible to develop representations of students that are more effective in detecting gender bias in CS courses, and that can be applied to a broader range of education software.

The present study investigated the characteristics of students that affect their experience in CS/SE education, as well as the extent to which their personal problem-solving traits align with the GenderMag framework. A series of workshops were conducted where students were asked to create and test personas representing themselves and their peers, and complete a survey assessing their personal cognitive styles. Thematic analysis of the deconstructed workshop data was used to identify gender related traits and personal qualities. These traits were used to construct gendered CS/SE student personas that were employed in cognitive walkthroughs of software applications commonly used in CS/SE education. Their performance in the detection of gender bias was then compared to the existing GenderMag personas.

The key contributions of this paper are: (1) A comprehensive list of traits and personal qualities that affect the study experience of CS/SE students, (2) evidence of gender differences in the CS/SE student population that are not captured in the existing GenderMag personas, and (3) an efficient and effective approach for collecting a rich qualitative dataset to use in the scoping stage of the InclusiveMag method.

The rest of the article is organized as follows; Section 2 presents a review of the relevant literature, Section 3 describes the problem statement and research questions under investigation, Section 4 elaborates research methodology, Section 5 illustrates our findings, Section 6 discusses these findings, Section 7 describes some of the limitations of our research, and finally Section 8 concludes the article.

2. Literature Review

Despite the increased participation of women in many STEM fields, CS/SE and other technology related careers continue to be male dominated [24, 25]. Efforts to address this issue are crucial, given evidence of enhanced performance and innovation in gender diverse organisations [26, 27]. In software development in particular, gender diverse teams have been associated with greater creativity and sounder decision making [28]. The shortage of women in these fields stems in part from their underrepresentation and lower course completion rates in technology related degrees [3, 6]. The broad themes of research on gender bias within CS/SE education is divided into investigating reasons and forming recommendations, gender bias within teamwork, gender expression and difference, self efficacy of students, and gender bias in CS/SE software. The following subsections present some key related research under each of these broad themes. It is noted that, there is a great body of research on gender in software engineering, [29, 30, 31, 32], however, since our focus is on education, we present literature focusing on gender in the CS/SE education context.

2.1. Gender Bias in CS/SE Education

A recent mapping of literature on "female inclusiveness" in CS/SE education reveals that a gender imbalance is still prevalent in this domain, although it is a knowledge and skill based profession, rather than a physical or labour intensive one [33]. There has been an increase in research on this topic since 2015, with a decline in 2020-21 with a possible reason of Covid-19 pandemic. Some significant findings of the systematic mapping study were - (1) a global imbalance in research on this topic, with the USA leading most of the research, (2) less research focusing on the persistence of women in the domain, and (3) a significant body of recommendations for educators.

Research into gender bias in education has traditionally investigated acts of direct discrimination [11], implicit beliefs that can lead some teachers to underestimate the competency of their female students [12, 13], or the use of gendered language or materials that reinforce negative stereotypes and create an unwelcoming learning environment [9].

Medel and Pournaghshband [9] examined the instructional materials used in CS education to identify gender differences in the representations, imagery, and language used. They found case study examples of gender bias including the disproportionate assignment of negative roles to characters with female names, persistence of a playboy centrefold picture as the standard stock image representing image processing, and perpetuation of stereotypes through the misuse of gendered pronouns. While their examples provide a compelling negative portrait of the portrayal of women within CS education materials, only a limited number were included. It has been demonstrated, however, that the creation of an unwelcoming learning environment can negatively affect the experience of women in CS [14].

Metaxa-Kakavouli et al.'s [14] study on the relationship between website design and ambient belonging tested the hypothesis that a perception that they do not belong decreases women's willingness to enrol in CS courses. When presented with an introductory CS course webpage designed to evoke masculine stereotypes, potential female students had lower confidence in their abilities, less sense of belonging, and a reduced intention to enrol in the course. While it is difficult to imagine that the choice of colours and pictures can have such a profound influence, the effect of masculine imagery on women's sense of belonging has also been observed offline in the design of CS classrooms [34]. Unfortunately, the implicit feeling of these students that they do not belong may be shared by some of their teachers [12].

In an extensive survey across CS and SE programs in the US, Cohoon [12] found a strong link between faculty attitudes towards diversity and women's confidence in their abilities and comfort asking questions in class. Surprisingly, half of professors in some faculties believed that efforts to increase the number of women would lead to a less capable student cohort. While the study also showed that, even in 2007, this belief was only held by a minority of teachers, negative assumptions about women's competence have also been revealed in more recent research within the broader STEM sector, and these assumptions may have harmful implications for their career prospects [13].

When presenting science faculty professors with the applications of otherwise identical students randomly assigned male or female names, Moss-Racusin et al. [13] found that students believed to be female were considered both less competent and less hireable than male students. This bias was present regardless of the professor's gender and resulted in them suggesting lower starting salaries and less career mentoring for students with feminine names. While not examining assumptions about competence directly, Wang and Redmiles [35] found a similar discrepancy in the hiring preferences of software engineers when presented with equally qualified men and women seeking leadership positions. Whether or not this sort of incidental discrimination can explain the underrepresentation of women in some STEM fields has been investigated by comparing them to fields with a more even gender balance [11].

From a series of studies with children and adolescents, Master et al. concluded that beliefs that "girls are less interested in computing and engineering" are formed early and cause gender disparity in later stages [36]. While it might be presumed that the presence of bias and discrimination provides a clear explanation for the underrepresentation of women in CS courses, comparison with other STEM fields reveals that the situation is more complex. While investigating potential reasons for the differing gender distributions across STEM fields, Cheryan et al. [11] found no evidence that formal gender discrimination is any more prevalent in CS than in courses where women are well represented. The authors conducted a broad literature review to identify factors that influence the participation of women in STEM courses before exploring the contribution of those factors to the discrepancies in women's representation across disciplines. Masculine culture and stereotypes, less experience with CS in early education, and lower self-efficacy were all identified as potential contributing factors. In contrast, while discrimination was present, it could not explain the greater underrepresentation of women in CS courses. It should be noted that the authors were looking at enrolment rates only and did not examine whether the discrimination identified negatively impacted educational experiences or student retention rates. Nevertheless, their findings suggest that there may be alternative explanations for the low participation of women in CS.

2.2. Gender Bias in CS/SE Teamwork

Investigating role selection in a final year student team projects, Nguyen-Duc and Jaccheri found that female students participated in project management and requirement engineering tasks more than architecture design or Scrum methods [37]. The authors also observed that the female students tended to engage in more light weight programming tasks over complex technical roles. However, they noted that these task assignments were more influenced by knowledge, and the previous experience and commitment of team members, than the gender identity of the students. Al-Taharwa, on the other hand, experimented with student teams with female leadership and regular membership of the team, and concluded that female-led teams followed good software engineering practices, however, achieved low prediction models [38]. In large software engineering projects, Laura and Marie found that female students were less likely to assign themselves to technical roles and consequently received lower peer review scores for their contributions to team projects [39].

Apart from team projects, gender has also been studied in pair programming settings within SE education. Choi et al. found that in terms of quality of programming, there was no significant difference between same and mixed gender pairs, however communication and compatibility was better in the same gender pairs [40]. Similarly, Gómez et al. did not find any significant difference in productivity between same and mixed gender pairs [41]. Productivity, however, was determined by lines of code written within a fixed timeframe. The authors did find that productivity varied a lot in mixed pair groups. In a randomized controlled trial with a large sample of undergraduate computer science students, Jarrat et al. investigated the influence of gender of the partner in a pair and found that having a female partner was associated with higher confidence in the outcome [42]. The authors also noted that female students had lower perceptions of their competence relative to their partners.

2.3. Gender Expression and CS/SE Education

While the term gender has often been used interchangeably with biological sex in academic research and viewed binarily [43], this fails to account for the true variability in human gender expression and may obscure interesting findings related to people who do not exist at either end of the masculine/feminine spectrum [44]. The present paper is concerned with gender expression, which relates both to how the individual sees themselves along that spectrum and their presentation and behaviour in relation to social norms [43]. While no person will demonstrate all the characteristics typically associated with either masculinity or femininity, there are differences in how individuals experience CS studies that can be correlated with their gender identity [15].

In a survey of first year students in the US, Beyer [15] found significant gender differences in student's confidence in their computer skills, interest in pursuing CS studies, and the values and motivations that shaped their career goals. Men were actually slightly more likely to hold negative beliefs about computer use, but the effect size was small, and they were still more likely to undertake a CS course than the women surveyed. While there was no relationship between gender and academic performance, lower confidence in their computer skills was evident even among the women who chose to undertake CS studies.

2.4. Gender Difference and Self Efficacy of Students

Gender differences in CS student's sense of self-efficacy are well documented in the literature [15, 11]. While investigating the effects of self-regulated learning techniques in a beginner CS course, Lishinski et al. [45] also found that self-efficacy is positively correlated with academic performance. Perhaps more interestingly however, they observed gender differences in the malleability of the student's sense of self-efficacy over time. While men's view of their abilities continued to change in response to feedback, women's self-efficacy was fixed earlier in the course, leaving them more affected by early challenges. Given the relationship between self-efficacy and performance outcomes, this study highlights possible issues with the presence of steep learning curves at the beginning of CS courses. It suggests that gender inclusive education should avoid overwhelming students early when women may be disproportionately affected by negative experiences and feedback. One of the challenges to this is the necessity of introducing new and potentially challenging software programs to students early in their CS studies.

2.5. Gender Bias in Software

To select software that meets the needs of all students, it is important to recognise that most software is not designed with inclusivity in mind. In a systematic mapping study reviewing current approaches for fostering gender inclusivity in software design, Nunes et al. [17] found three recurring issues- (1) creators tend to assume that software, and the common methods for developing it, are gender neutral even where there is no gender diversity in the development team, (2) gender is looked at binarily and without consideration of how to address any gender bias that is identified, and (3) there is a lack of information and guiding methods available to help teams navigate gender inclusivity issues. In view of these findings, it is unsurprising that the authors also found that developers tend to design according to their own preferences and requirements, failing to account for the needs of a diverse potential user base. This bias may create usability issues that disproportionately affect users who are demographically different from the developers. If these inclusivity issues are to be addressed, they must first be detected. This can be done without the involvement of end users using expert-based testing methods like cognitive walkthroughs or heuristic evaluations [46].

Nielson [47] describes five elements of software usability: learnability, efficiency, memorability, low error rate, and subjective satisfaction. Subjective satisfaction is an aspect where individual differences in personal characteristics, some of which will be gender correlated, could be particularly impactful. The Cognitive walkthrough is a usability inspection method able to capture this internal experience of users [48], that has been found to be particularly sensitive to issues related to learnability [49, 50], another crucial element in the education context.

2.6. Education software

While learning and using new software systems is an integral part of any CS/SE program, much of the work related to their design and selection has not considered student diversity. While Oliwa [51] sought the opinions of students, teachers, and administrative staff on the required functions of an online learning platform, they did not consider demographic differences within those groups. In a case study looking at the needs and expectations of students, Şahin and Yurdugül [52] found that a competitive environment, including elements like leaderboards, is a desirable feature of online learning platforms. However, their study also examined the needs of students collectively, and did not consider how personal characteristics like a user's age, cultural background, or gender could affect the suitability of certain features. This is problematic considering there is research that suggests more competitive settings may actually decrease the educational performance of women [53]. While Lim et al. [54] found gender differences in the features used and frequency of interactions with a LMS when surveying university students, the reasons for those differences and how they might be mitigated through better software design, were left unexamined.

In contrast, while not directly considering demographic differences, Kolekar et al. [55], proposed an addition to an existing LMS that would allow it to adapt to student's learning styles. Based on usage data, students are categorised into one of eight groups. The system then presents different user interfaces and learning materials based on the presumed needs and preferences of that user group. While the authors found that the system improved the performance of students, their sample was small, and detailed results, including the significance of the difference found, were not reported in the paper. That said, it provides an interesting example of how education software could be made more inclusive.

In a similar small study, Yalamu et al. [56] created a LMS prototype that incorporated aspects of Papua New Guinea indigenous culture. While most of their participants preferred to view the interface in English over their local language, the incorporation of cultural icons and motifs improved the experience of many users. Feedback comments from 14% of the participants indicated that seeing elements of their own culture made them feel a greater sense of belonging. This provides an interesting contrast to the work of Metaxa-Kakavouli et al. with a masculine website design [14]. Building inclusivity into a user interface could involve the addition of elements that students can relate to, not just the removal of imagery that could make them feel excluded.

In addition to LMSs, CS/SE students are exposed to a broad range of software packages including programming environments, database management systems, and project management software. Selecting which software to use in a CS course is a unique challenge as technology is constantly changing, and the tools used are often the same as those used in industry [57]. While gender differences in software use have been thoroughly investigated in the literature [58], as evidenced above, they have often been overlooked in the selection and design of education software. As a further example, Parker [57] published a guide for educators to use when selecting software for use in information technology courses. While the author included ease of use and learnability as factors for consideration, there was no discussion of the differing needs of individual students.

2.7. The GenderMag Approach

GenderMag is a research supported tool for identifying gender bias in software using personas and a custom cognitive walkthrough method [16]. Burnett et al. assembled five facets (information processing style, learning style, computer self-efficacy, attitude toward risk, and motivations) of software use that are known to correlate with the user's gender when developing a new tool for the detection of gender bias. Women are generally more risk averse, more likely to adopt a comprehensive information processing style, and less likely to tinker with new software features. As with CS/SE education, they also tend to have lower self-efficacy than men. These facets are represented in the GenderMag personas, 'Tim', 'Pat' and 'Abi'. A persona is a descriptive profile of a fictional person used to represent a class of potential users while conducting usability testing [19]. These personas are used in a systematic cognitive walkthrough process to identify usability problems in problem-solving software interfaces.

The GenderMag method has been shown to be highly accurate in its identification of software usability issues [16], and applying it to increase software inclusivity can lead to design improvements that benefit all users, not just women [18]. Shekhar and Marsden [22] employed the GenderMag method to investigate gender bias in LMSs with information technology and software engineering students and professionals. While they modified the demographic traits of the GenderMag personas, they used the pre-set facet values. Their study showed that the method was effective for detecting usability issues, and that the female persona 'Abi', in particular, could help participants consider the needs of users different from themselves.

In the education context, GenderMag has already been used in the assessment of both online courseware [23] and a LMS [22], and there is growing support for the use of personas and other user experience (UX) tools in STEM course development [59]. However, GenderMag inquiries are limited to five facets based on gender differences observed in the general population [16]. While limiting the number and nature of facets in this way makes the method more feasible [21], it also narrows its scope. At present, there have been no investigations into whether these facets can accurately capture the experiences of CS/SE students when learning new software.

CS/SE students are a unique group. They show notable personality differences even when compared with students in other STEM disciplines [60, 61]. Beyer [15] found that personal qualities like low conscientiousness and low openness to experience are predictors of whether a student will take a CS/SE course regardless of gender. Low openness refers to avoiding changes or resisting new ideas, and low conscientiousness refers to not preferring structure, schedule or organization [62]. While there are gender correlated differences that affect the experience of CS/SE students [15, 63], like personality, these may not entirely mirror traits seen in the general population. This means that there may be relevant usability issues that are not detectable with the existing GenderMag facets.

Mendez et al. [21] introduced InclusiveMag as a method for developing new inclusivity tools, similar to GenderMag, that support underserved populations. Their method has three stages: scope, derive, and apply. In the first stage researchers set the domain and conduct research to identify relevant facets. Stages two and three involve the creation and use of personas based on those facets. In the present paper we take a narrow focus, examining gender differences found specifically within the CS/SE student population.

3. Problem Analysis

While the utility of the GenderMag personas has been thoroughly investigated [18, 23, 64], the applicability of their facets to CS/SE students as a distinct population has not. To directly address the lower course completion rates of female students, it is necessary to investigate the qualities and experiences that make this group unique. Employing personas that capture these qualities in cognitive walkthroughs of education software may provide a more accurate depiction of the difficulties and frustrations faced by these students. Furthermore, prior work in education was limited to LMSs [22] or online courseware [23]. Consideration of a broader range of applications, including those developed for use in industry, will improve our understanding of the usability issues encountered by CS/SE students during their degree.

The study has been designed to address the following research questions:

- RQ1: What are the facets of CS/SE students that affect the way they experience CS/SE education?
- RQ2: Do CS/SE students express gender differences in problem solving consistent with the GenderMag framework?
- RQ3: Can personas developed by students in collaborative workshops be adapted to detect gender related usability issues in software?
- RQ3: Are custom student personas more effective at detecting usability issues in CS/SE education software than the GenderMag personas?

4. Methodology

The study was conducted in three phases (Figure 2). In the first, a series of persona development workshops were held to investigate the personal traits of CS/SE students (RQ1) and assess their problem-solving traits under the GenderMag framework (RQ2). The second phase was a thematic analysis of the workshop data to see if gendered CS/SE student personas could be extracted from the participant responses (RQ3). In the final phase these personas were used in a series of cognitive walkthroughs of education software, and their performance compared to GenderMag's Tim and Abi (RQ4).

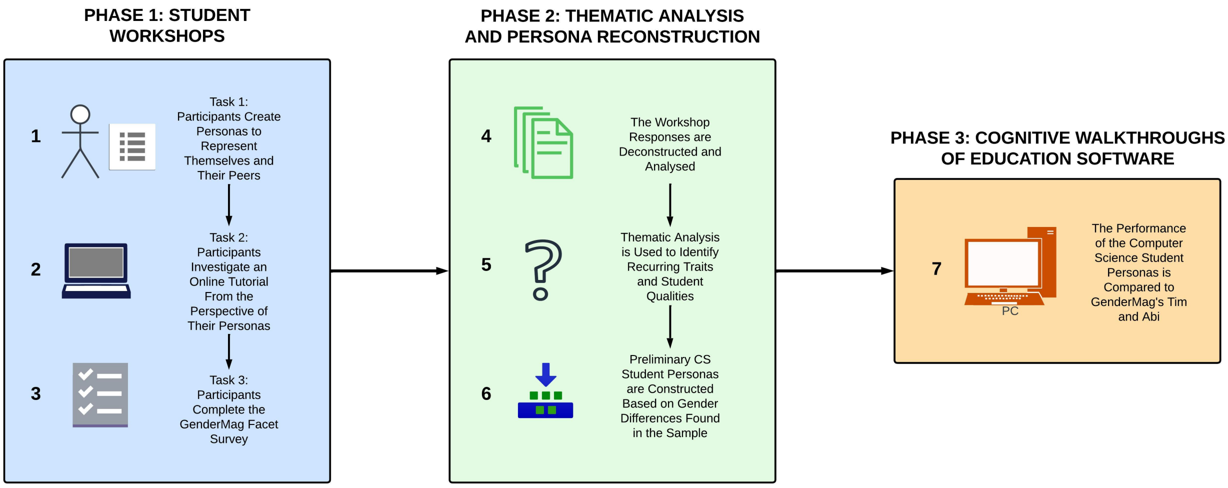


Figure 1: Methodology Overview

4.1. Recruitment

Monash University students over the age of 18 currently studying a CS/SE related course were recruited through convenience sampling. This means that the participant recruitment was based primarily on availability and interest over a more purposive approach [65]. The study was advertised through social media, student clubs and societies, and leaflets distributed around the campus. Potential participants were also invited to participate in person or contacted via email. While convenience sampling reduces the generalisability of the findings [66], the limited time available, and scheduling requirements of the study design, made it the most appropriate option for the project.

4.2. Persona Creation Workshops

Students were invited to attend a one-hour workshop designed for groups of three to five participants. Although multiple participants attended the workshops at the same time, they were not particularly collaborative in nature. A workshop approach was adopted to allow for the collection of a larger number of participant responses within the time constraints of the project and collate a more representative data set [65]. Allowing each participant to create individual responses ensured that the perspective of less extroverted students would still be captured in our results.

The workshops were divided into three parts. In the first part, the participants created one or more CS/SE student personas. This approach was selected over a traditional interview method to allow persona development to be guided by CS/SE students themselves. These tasks made no reference to the GenderMag facets and gave participants significant freedom over the content and structure of their answers. This limited the influence of the researcher's familiarity with GenderMag on participant responses, and increased the likelihood of identifying novel insights into gender differences that are not captured under the GenderMag framework. In the second task, participants were instructed to investigate an online tutorial from the perspective of the personas they created. Finally, participants were introduced to GenderMag to complete their third task. An online survey rating their personal cognitive styles according to the five facets.

4.2.1. Participants

A series of eight workshops were held with groups of one to four participants. While the initial intention was to conduct all workshops with groups of three to five, difficulties with scheduling and participant retention meant that four were held with only a single participant. A total of 15 participants from diverse cultural backgrounds attended. Of those participants, 14 disclosed their gender (seven male, seven female) and most were aged 18-21. A summary of the participant's course enrolments can be seen in Table 1, with the majority (73%) undertaking degrees in CS/SE or Information Technology. All participants were on-campus students, with 14 of the 15 participants studying full-time.

Table 1: Summary of Participant Enrolments

Course	AQF Level	No. Participants
Computer Science	Bachelor	5
Data Science	Masters	2
Information Technology	Bachelor	4
Information Technology	Masters	1
Information Technology	Unknown	1
Other	Bachelor	1 Engineering (Specialising in Software) 1 Business (Currently taking an IT subject)

4.2.2. Task 1:Student Personas

Participants were asked to create one or more CS/SE student personas based on themselves or their peers. Human factors, commonly present in personas across different domains and identified by Karolita et al. [67] in their persona taxonomy, were used as examples to encourage the students to create realistic and detailed student descriptions. These factors included personal characteristics, skills and experience, and group characteristics.

Participants were provided with the following prompts when beginning the task:

> Who are they?
> What are their goals?
> What motivates them?
> What are their constraints?
> *Is there anything that could make learning or using new software difficult for them?*

They were given freedom over the format of their responses, as well as whether they included a picture, but were asked to add a one-line summary at the end. Karolita et al. [67] found that 'taglines', or one-line statements, are a commonly included feature of personas used in the software development domain. Summaries were requested in this case to gain insight into which characteristics the students considered most important. Most participants chose to create a single persona based on themselves.

Pilot Workshop Outcomes

In a pilot workshop, two participants wrote a group description in their response instead of an individual persona. To limit this in future workshops, age, gender, skills, and experience were added to the instructions as examples of "Who are they?". This encouraged the participants to think of the personas as individuals and provide more personalised responses.

4.2.3. Task 2:Persona Testing

Participants selected one of two introductory courses in MATLAB or SQL for the testing task. They were asked to reflect on the following questions as they investigated the tutorial:

> What is your persona thinking as they go through the steps?
> Do they get stuck at all?
> Is there anything that bothers them?

Participants were told they would use their personas for this task before they began creating them. This was to provide context for the participants and improve the relevance of their responses to their educational experiences and use of software. MATLAB and SQL were selected because they are widely used for different units within SE. However, since not all students complete all units, and some participants may not be familiar with these languages, the courses selected started at an introductory level.

4.2.4. Task 3:GenderMag Facet Survey

Their last task was to complete a survey, available through Open Educational Resources Commons [68], where they rated their cognitive styles across the five GenderMag facets. Scores were used to evaluate their similarity to the three GenderMag personas. The 'Tim' persona has facet values most commonly seen in men, 'Abi' those seen in women, and 'Pat' an amalgamation of the two [16]. In general, people not showing extreme values are referred to as "Pat".

The five facets are information processing style, learning style, computer self-efficacy, attitude toward risk, and motivations (Figure 2). Individual facet scores range from -10 to +10. A very negative score on a facet indicates a cognitive style similar to the Tim persona, while very high scores suggest similarities with Abi. Total scores range between -50 and +50. Average scores by gender were compared to evaluate the extent to which the sample's problem-solving traits are consistent with the GenderMag framework. Since the authors did not develop this survey, adopting it from the Open Educational Resources Commons [68], we used the same scoring mechanism (resulting in positive (+) and negative (-)) provided by the source.

4.3. Analysis

4.3.1. Deconstruction and Thematic Analysis of Participant Responses

A thematic analysis of the participant's written responses was used to identify and categorise personal qualities that affect student's experiences in CS/SE education. Thematic analysis is a flexible approach that aids in the discovery and reporting of relevant patterns within a qualitative dataset [69, 70]. Braun and Clarke [70] describe a six-phase model of thematic analysis that begins with the researcher familiarising themselves with the data before developing the initial codes.

Familiarisation and Initial Coding

Memos were taken to record the researcher's thoughts and reasoning as initial impressions of the individual participants, created personas, and complete dataset were developed. Preliminary codes were appended to each passage of text produced during the persona and test activities. In light of the additional context provided by the association between the two tasks, and records of participant comments within the workshops, a latent coding process was chosen over a semantic approach. Latent coding attempts to ascertain meaning in data that goes beyond literal description or summary of participant words [70].

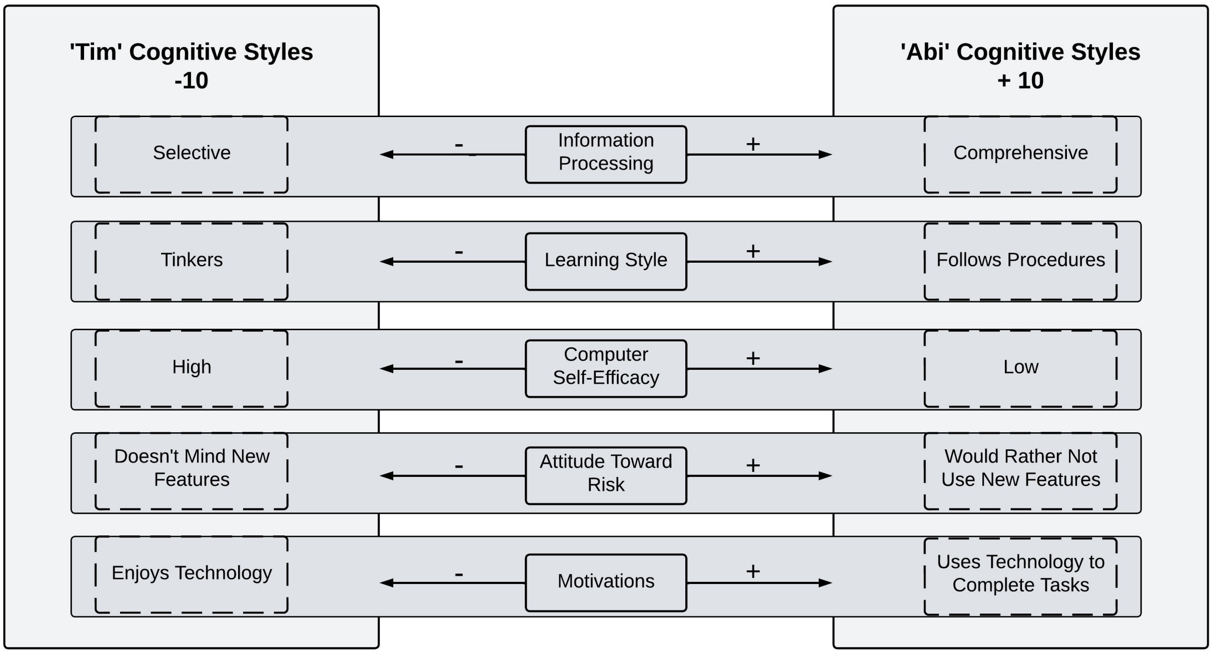


Figure 2: The Five GenderMag Facets. Adapted from the Cognitive Styles Survey [68]

Searching for Patterns

The initial codes were categorised based on emerging patterns and the workshop prompts. The data was restructured using the recurring attributes: gender, age, culture, group identity, skills/experience, background, preferences, learning style, motivations, goals, frustrations, and constraints. Figure 3 presents a snippet of coding. These attributes were used to create affinity diagrams grouping and connecting related text segments. Associated and heavily overlapping attributes, like motivations and goals, were merged during this process (Figure 4). Colours were used to denote the gender of the persona each extract was taken from (blue, green and yellow were used to denote female, male and unspecified or non-binary, respectively). These diagrams were used to identify recurring traits within the attributes and iteratively refine the code list. This refined code list was used to recode the complete dataset before further analysis. For example the female persona description - "work in (G)oogle", male persona attribute - "have a successful career in AI" and an unspecified persona attribute "getting into... IT companies" were coded under "opportunity" which was merged with other codes such as "industry", "lifestyle" and "personal fulfillment" to form the group "Career", indicated in Figure 4.

When reviewing the use of UX design methods in STEM education research, Minichiello et al. [59] found that in most papers, manual qualitative clustering methods like affinity diagramming were used to organise data during the persona development process. Affinity diagramming is also strongly recommended by Adlin and Pruitt in their guide to building personas because of its speed and ease of use [19].

4.3.2. Persona Reconstruction

For each attribute and its associated codes ('student qualities'), relevant text passages in the dataset were collected and grouped by gender. Comparisons were made based on the presence/absence of traits, the student qualities represented, and the language used by the participants. Notable differences were used to create preliminary CS/SE student personas. Since our target is to understand the female and male facets displayed by CS/SE students to detect

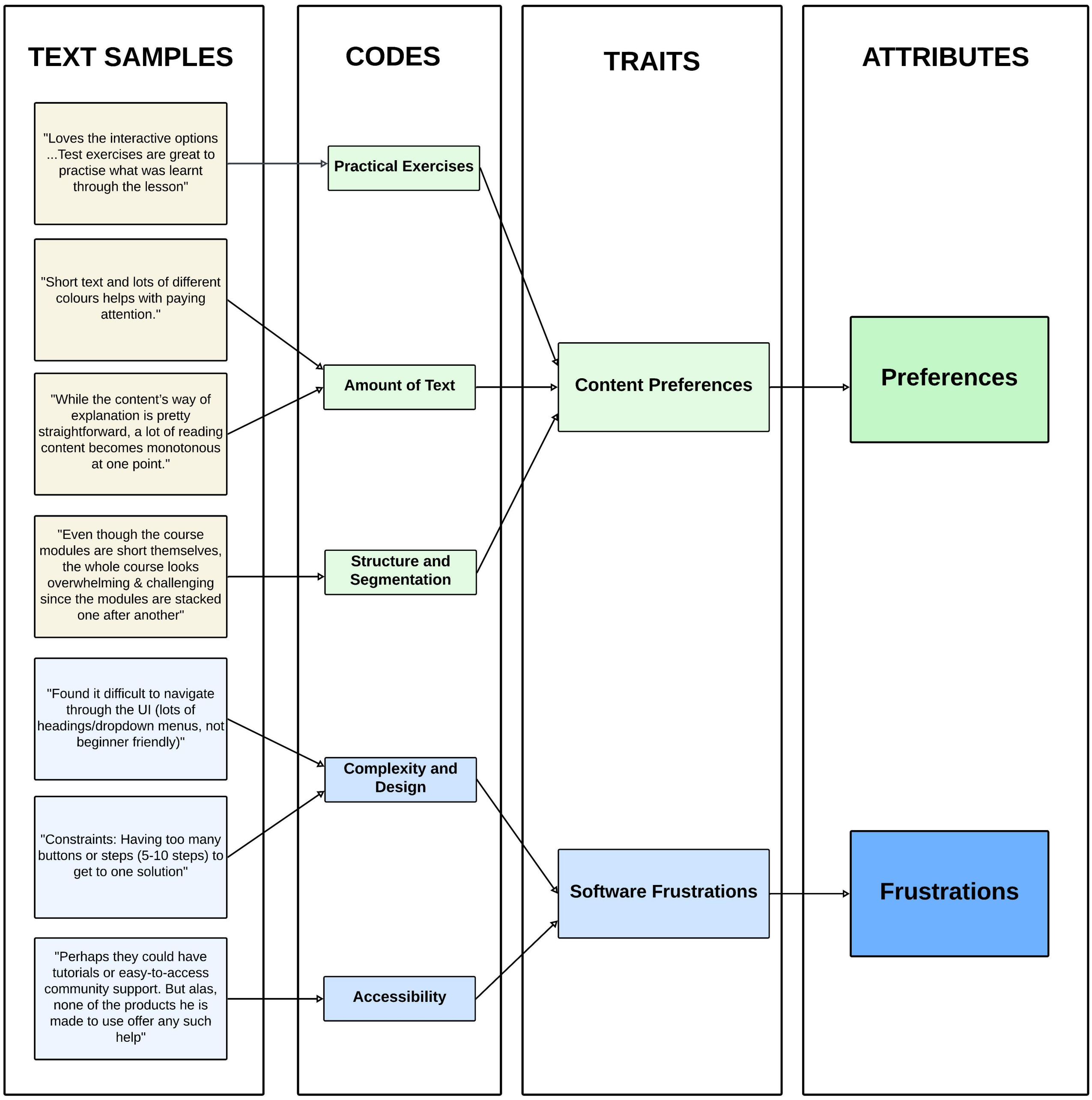


Figure 3: Example of coding

potential bias within education software, we developed "Abi" and "Tim" personas only.

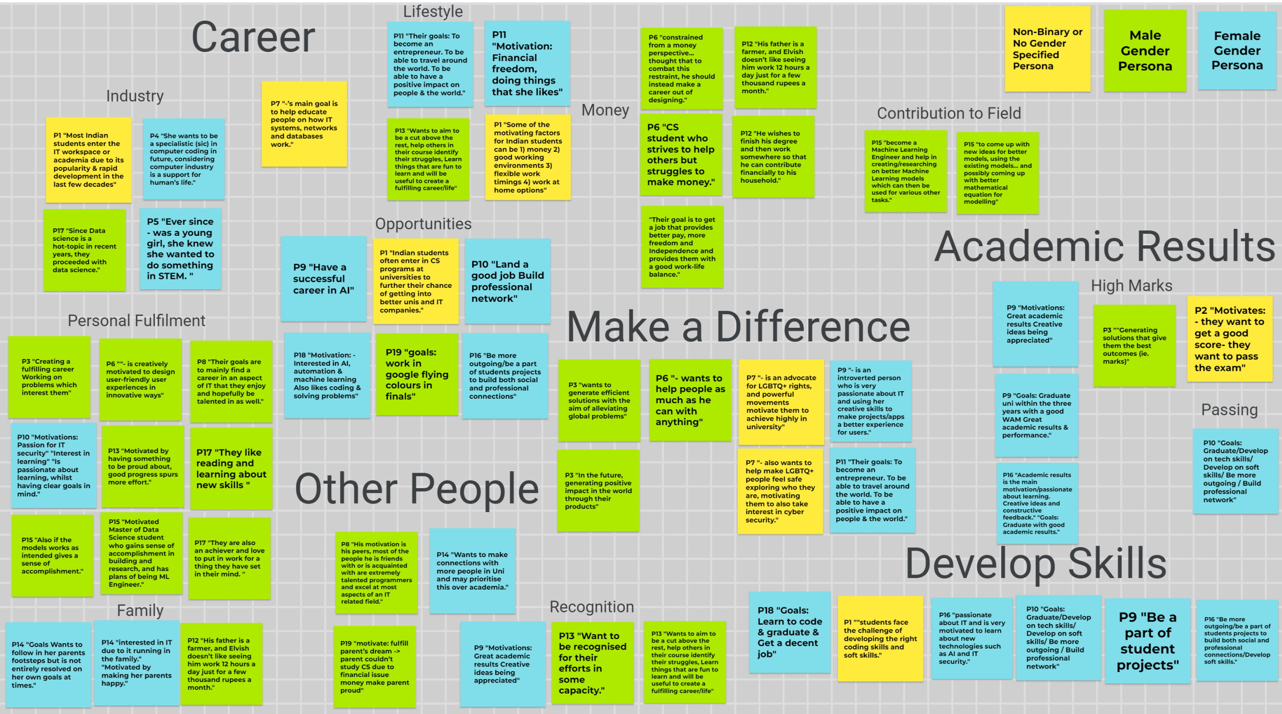


Figure 4: Affinity Diagram: Motivations and Goals

4.4. Cognitive Walkthroughs

While Minichiello et al. [59] found that assessment of new personas is not always reported in usability research, they nevertheless described it as the important last step in the persona development process. The preliminary CS/SE student personas were assessed by employing them in cognitive walkthroughs and comparing their performance to GenderMag's Tim and Abi.

Four desktop or web-based software applications were chosen for evaluation from the following categories: database management, procedural programming, project management and learning management. These applications were selected because they are used for multiple units taught in CS/SE. The selected applications cover a breadth of application types in CS, especially majoring SE students will encounter over the course of their degree. The selection, outlined in Table 2, includes an application designed for use in education, and others drawn from industry. It includes both text-based and more visual tools, as well as variety in the complexity of the user interfaces. This was to expand upon the prior work examining a single LMS [22], and explore the experience of student users when learning diverse application types. For each application, two multi-stage tasks were selected to investigate its learnability and ease of use.

The GenderMag cognitive walkthrough requires the evaluator to answer a series of questions for each subgoal and subsequent actions a persona will take [16]. This includes whether the persona would have formed the subgoal and whether the persona will know what they need to do at each step [16]. While these questions are highly relevant when assessing usability in industry, they make less sense in an education context where students are often provided with very specific instructions. In light of this, the GenderMag walkthrough was adapted to evaluate tasks appropriate for students learning a new software. The structure of this walkthrough is presented below:

For each instruction/set of instructions consider:

1. Would the persona have trouble understanding this instruction or step?
2. Would the persona encounter any frustrations or difficulties when taking this step?
3. Would the persona have any trouble knowing that they have completed this step successfully?

The second question aims to capture the usability element of subjective satisfaction as defined by Nielson [47]. The walkthroughs looked not only at whether the fictional student could successfully complete the tasks, but also whether they would find the application pleasant to use.

Table 2: Applications Selected for Cognitive Walkthroughs

Application	Category	Domain	Presentation	Complexity
SQL Developer	Database Management	Industry	Text-Based	High
BlueJ	Procedural Programming	Education	Visual	Low
IntelliJ	Procedural Programming	Industry	Text-Based	High
Trello	Project Management	Industry	Visual	Low
Canvas	Learning Management	Education	Visual	Low

5. Results

5.1. Workshop Responses

5.1.1. Personas

A total of 19 personas were created. Eighteen were created by individual participants, and one was created by two female participants in collaboration. Of the 19 personas, 17 included gender. Eight were male, eight were female, and one was non-binary. Two were members of the LGBTQ+ community. A summary of the participants and their personas can be seen in Table 3. Fifteen personas included age (M = 22.67, SD = 4.20), and seven included cultural background. In each case where age and cultural background were included, these matched the personal demographics of the participant who created the persona.

All but two personas included a one-line summary. Demographic information, most commonly gender, was included in over half of these. Of the seven male summaries, four referred to personal constraints or difficulties that affected their studies, and six included goals or motivations. In contrast, goals or motivations were only included in one female summary, while personal constraints were included in five.

There were 17 responses to the testing task. Two male participants chose to go through the coding tutorials with two different personas. The testing task responses provided examples of behaviour, preferences and frustrations that supplemented the persona descriptions during the analysis. Hereafter a ‘persona’ created by the participants refers to both the task 1 description and content of any corresponding test task.

5.1.2. GenderMag Scale

All 15 participants rated themselves using the GenderMag Cognitive Styles survey. Their total scores were all in the -25 to +25 range indicating that none of the participants were particularly Tim or Abi like in their gender presentation. As the participant scores were not normally distributed, and only a small sample size was available, Mann-Whitney U tests were

Table 3: Persona and Participant Demographic Information

No	Persona Gender	Participant Gender	Persona Age	Participant Age	Persona Cultural Background	Participant Cultural Background
1		Male		18-21	Indian	Indian
2		Female		26-30		Chinese
3	Male	Male	18	18-21	Caucasian	Caucasian
4	Female	Female	30	26-30		Chinese
5	Female	Female	20	18-21		
6	Male			22-25		Australian
7	Non-Binary			22-25		Australian
8	Male	Male	19	18-21		
9	Female	Female	23	22-25	Sri Lankan	Sri Lankan
10	Female	Female	23	22-25		Sri Lankan
11	Female	Female	20	18-21		Cambodian
12	Male	Male	20	18-21	Indian	Indian
13	Male	Male	19	18-21		Australian
14	Female	Male	19	18-21		Australian
15	Male	Male	24	22-25	Indian	Indian
16*	Female	Female	23	22-25	Sri Lankan	Sri Lankan
17	Male	Male	30	22-25		Indian
18	Female	Female	20	18-21		Malaysian
19	Male	Male	22	18-21		Malaysian

* Created in collaboration by the participants who created Personas 9 and 10

used to compare the scores of the male and female participants. The results found no significant differences between the total scores ($U = 18$, n1 = n2, = 7, $p > 0.05$ two-tailed), or the scores of any individual facets. A summary of the mean values by gender can be seen in Tables 4 and 5, below. Very positive values indicate trait expressions more commonly seen in women, and very negative values those seen in men.

Table 4: Mean Total Scores by Gender

Gender	Total Score				
	N	Min	Max	Mean	Std
Men	7	-18	17	1.29	14.48
Women	7	-16	9	-3.00	9.68

Table 5: Mean Facet Scores by Gender

Gender	Information Processing			Learning Style			Computer Self-Efficacy			Attitude Toward Risk			Motivations		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
Men	4	10	6.43	-6	10	2.29	-9	5	-3.86	-9	4	-1.57	-9	10	-2.00
Women	-9	8	2.71	-10	9	3.29	-10	8	-4.71	-10	10	-3.43	-10	8	-0.86

5.2. Thematic Analysis

The analysis revealed 18 traits that capture the full range of qualities present in the written responses. These are outlined in Table 6. Traits or qualities where the participants in our sample displayed notable gender differences are denoted with "(*f)" and "(*m)" signs - to indicate majority appearance in female and male personas, respectively.

The personas described variable learning styles, but no gender patterns were identified within this attribute. Notably, statements suggesting a preference for a slower learning pace, including concerns with being overwhelmed with too much information at once, were present in a third of the personas, but this did not appear to be gender related.

5.2.1. Skills, Experience and Background

Existing CS Skills. Most of the responses (15) provided some description of the personas level of experience in CS. Eight personas, four male and four female, had existing skills or knowledge. All of these experienced personas had some programming knowledge, and there were examples of both male and female personas with skills in data analysis or UI design. Male and female personas were also equally likely to be described as being inexperienced.

While there were no gender differences in the types of existing CS skills, there were differences in the language used to describe them. While the words "experience" or "background" were used in three of the four experienced male personas, female personas only referenced "skills" or "knowledge". While two of the male personas were described as being "talented" or "good with", the only measure of skill included in a female persona was "fundamental"

> *Male Persona Example: "Talented UI designer and data analyst, some experience from high school as well as personal use of different programs"*
> *Female Persona Example: "Skills: fundamental knowledge of Python, Java, HTML, MATLAB and CSS..."*

Communication and Teamwork. The ability to work with others was the most referenced non-CS related skill. More than half of the female personas (5) included communication, teamwork, or people skills. In contrast, these types of skills were not included in any of the male personas. This greater interest in collaboration and social connections was also reflected in the goals of some of the female personas.

5.2.2. Motivations and Goals

Money and Lifestyle. Money or lifestyle aspirations were motivating factors for six of the student personas (four male, one female, one ungendered). Financial difficulties motivated two of the male personas. They wanted to earn enough money to support themselves, or their families. Others hoped that their CS studies would lead to job opportunities with higher pay, increased flexibility and freedoms, or a better work life balance.

> *Male Persona: "1) money 2) good working environments 3) flexible work timings 4) work at home options"*
> *Ungendered Persona: "Their goal is to get a job that provides better pay, more freedom and Independence and provides them with a good work-life balance."*

Interest/Passion The female personas were more likely to be motivated by passion. While four female and two male personas expressed goals or motivations based on interest in CS, the descriptions used in the female personas were more emotive. Three used the words "passion for" or "passionate about", while the male personas only referenced enjoyment or interest.

Motivated by Others. Of the eight male personas, six included motivations related to other people. Two of these wanted to exceed or match the success of their peers, two were motivated by family, and two aimed to improve other's lives through their future work. The non-binary persona similarly had altruistic goals, wanting to promote online safety for LGBTQ+ people. In contrast, only one of the female personas was primarily motivated by others. They were

Table 6: Traits and Qualities of CS/SE Students

Attribute	**Traits**	**Student Qualities**
Skills, Experience and Background	Level of Experience	Inexperienced in CS Not Sufficiently Prepared by Prior CS Studies Prior Skills/Knowledge in CS
	Non-CS Related Skills	Communication and Teamwork (*f) Other Skills
	Harmful Gender Related Experience	Stereotypes (*f) Harassment
Preferences	Content Presentation	Appealing Presentation Attention (*f) Visual Learning (*f)
	Content	Structure and Segmentation Amount of Text Practical Exercises(*m)
	Software Usability(*m)	Clarity Navigation
	Independence	Prefers Support Works Alone
Learning Style and Behaviour	Attitude Towards Learning	Enjoys Learning Learns Everything They Can Learns Only What's Necessary
	Ease of Learning	Easily Picks Up New Content Has Difficulties Learning Perseverance Through Difficulties
	Learning Pace	Works Efficiently Slower Learning Pace
Goals and Motivations	Career	Money and Lifestyle (*m) Interest/Passion Recognition or Accomplishment (*m)
	Study	Academic Achievement(*f) Skill Development Social(*f)
	Other People(*m)	Family Peers Altruistic Goals
Frustrations and Constraints	Software(*m)	Accessibility Complexity and Design
	Teaching Approach	Pacing of Content
	Content	Lacks Connection to Future Career Lacks Clarity Poor Presentation(*f)
	Personal Barriers	Struggles with Emotions or Mental Health (*m) Financial or Resource Constraints (*m) Time Management and Motivation (*f) Low Confidence in Abilities (*f)
	Interpersonal	Lack of Support Gender Stereotypes(*f) Social Difficulties

taking a CS course because of family expectations.

Recognition or Accomplishment. Similarly, a wish to be exceptional or have recognition for their talents was common in male personas (3), but almost entirely absent from the female personas.

> *Male persona: "Wants to aim to be a cut above the rest... Motivated by having something to be proud about, good progress spurs more effort. Want to be recognised for their efforts in some capacity."*

While many female personas included goals related to career success, none included the desire to make a unique contribution in CS or be more successful than their peers.

Social Goals. While other people were not primary motivations in the female personas, many described personal social goals that were not present in any of the male personas. Half (4) aspired to be more extroverted, participate in student projects, or build their personal or professional connections.

Academic Goals. Academic goals were present in both male and female personas but these tended to be more explicit, and were more likely to be repeated multiple times, within the female personas. In one female persona, high academic achievement was their first listed motivation, goal, and constraint.

> *Female persona: "Worrying too much about academic results which restricts her from trying new things. Eg. studying electives from other faculties"*

5.2.3. Preferences

Presentation and Attention. Almost half of the personas (9) included aesthetic preferences for content presentation or UI designs. These included the use of colour and interesting fonts, length of text passages, and general creativity and character of the overall design. Male and female personas were equally likely to be bothered by a bland or boring presentation, but in female personas this was more often linked to motivation and attention.

> *"Believes that a nicer UI would help in keeping student's interest on the page"*
> *"Rather boring site - might lose interest quickly"*

A preference for visual learning, with the inclusion of video or picture content, was also found in two of the female personas.

Content and Structure. The length and structure of text passages was another element frequently linked to attention and motivation. Both male and female personas showed a preference for shorter text passages and examples, or the inclusion of summaries and overviews. In female persona's this was expressed in their criticisms of the online tutorials used in the second task.

> *Female Persona Extracts: "no overview showing all components of SQL & how they relate to each other" "no summary given after each topic"*
> *"While the content's way of explanation is pretty straightforward, a lot of reading content becomes monotonous at one point."*

While examples were seen to be useful, they could be viewed negatively when not sufficiently realistic. One of the career-oriented male personas criticised a lack of connection between the examples available and their future work.

> *"Real world examples are present but still not similar to what would be seen in career."*

Inclusion of practical exercises was viewed positively in many personas, but both the male personas and real-life participants had mixed reactions to the quizzes present in the online tutorials. While they were viewed as being useful, they were also a source of stress.

> *Male Persona: "If this quiz were assessed - would be under a lot of stress and pressure making some of his answers incorrect as he cannot think straight."*

Software Usability. Preferences related to usability were found in seven of the personas (Four male, two female, one ungendered). There were preferences for clear icons and layouts, simple menu structures, and navigation that doesn't require students to take an excessive number of steps to reach their goal.

This wish for software that is easier to use was also reflected in many personas' frustrations. A lack of accessibility options, like tutorials, and poorly designed interfaces were common concerns.

5.2.4. Frustrations and Constraints

The constraints and frustrations of most female personas focused only on personal limitations, while the male persona's constraints were equally likely to be based on personal limitations and external factors outside of their control.

Low Confidence in Abilities. In some female personas, the descriptions seemed to show a lack of confidence in the student's ability to complete their coursework. While there were also examples of male personas who were struggling, none seemed to imply that the student could not learn if given additional time or resources.

> *Male Persona: "doesn't always pick things up immediately, requires access to resources that make it easy to look back on and fix misconceptions or that help you apply your learning."*
> *Female Persona:: "She tried so hard, but it's not working." "is a helplessful (sic) student"*

Personas were not just affected by their own perceptions, but also by how they were seen by others. Two female personas were concerned with being perceived negatively by their peers because of their gender. This was expressed as being "looked down on" or "considered as having low potential". In one case this was a result of personal exposure to negative stereotypes in STEM. Similarly, one of the male personas described being negatively impacted by their teacher's impressions.

> *"and his professors sideline him as they think he's not smart enough to complete the degree."*

Emotions and Mental Health. Descriptions of emotional difficulties were present in four of the personas. Three of those personas were male and one was non-binary. Two personas had an existing mental health condition (one male, one non-binary). The others were stressed or overwhelmed by assessment or heavy workloads.

Time Management and Motivation. Six personas (four female, two male) had difficulties with time management. There were examples of both male and female personas who struggled to meet assignment deadlines or keep up with their course content. Reasons for this included unmanageable workloads, problems with motivation and, in the case of two female personas, struggles with competing commitments and work-life balance.

> *"Sometimes failing to balance work/professional/personal life while being a student"*
> *"Has other commitments that can take up a lot of time so her attention is often split on various goals."*

Financial or Resource Constraints. Descriptions of personal financial struggles or a lack of access to required resources were present in half of the male personas. Two of these were unable to access software they needed because of OS or hardware limitations.

5.3. Personas

Gender differences identified in the sample were used to construct the preliminary personas shown in Figure 5 below (Hereafter CSTim and CSAbi). While individual students are likely to display a range of these qualities, and may behave in more "Pat" style, we considered the extremes of the traits found in our sample. As such only "Tim" and "Abi" personas were created. Since our main motivation was to detect bias in education software, two personas with the extreme qualities of each spectrum were sufficient for the congnitive walkthroughs.

5.4. Cognitive Walkthroughs

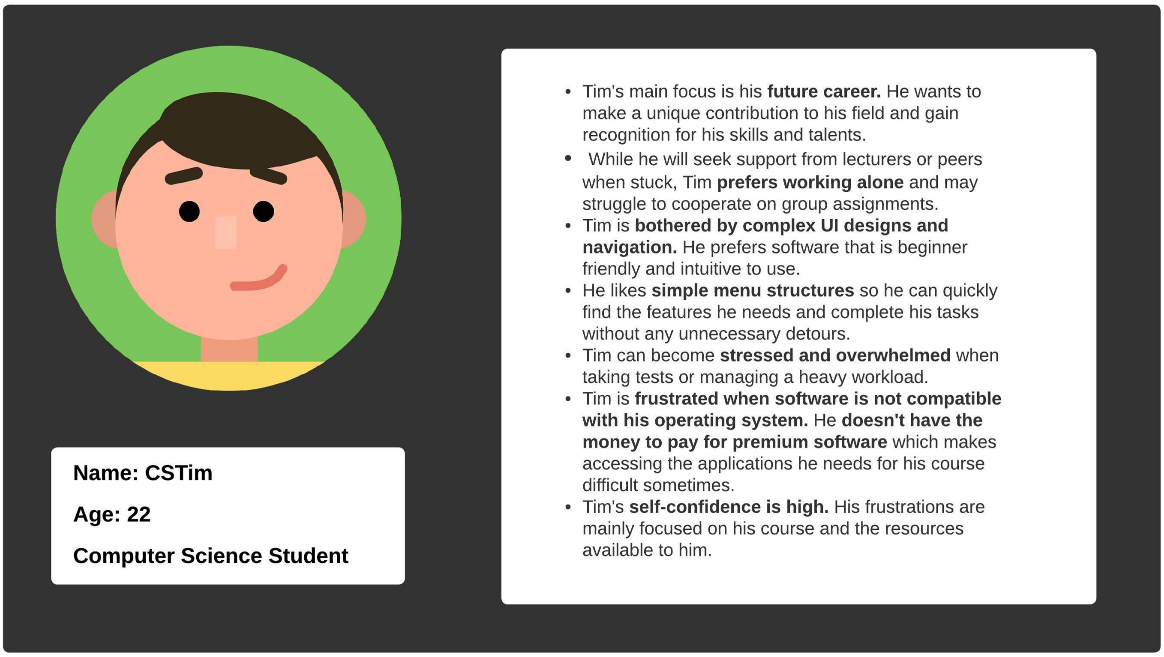
Usability issues were found in all eight cognitive walkthrough tasks. The number of issues detected did not appear to be related to the application's complexity, or whether it is primarily visual or text based. An overview of the results for each persona can be seen in Table 7.

Table 7: Usability Issues Detected (Yes:Y, No:N)

Application	**Task**	**Tim**	**Abi**	**CSTim**	**CSAbi**
MySQL	Create a new database	N	N	N	**Y**
	Create a table using the MySQL workbench	N	N	**Y**	N
BlueJ	Learn to use basic features	**Y**	N	N	N
	Debug a program	N	**Y***	**Y**	**Y***
IntelliJ	Prepare a new project	N	**Y**	N	**Y**
	Learn code completion	N	**Y**	**Y**	N
Trello	Create and share a simple board	N	N	N	N
	Delete an existing Trello Board	N	**Y***	**Y***	**Y***
Canvas	Find recorded sessions	N	Y*	N	Y*
	Joining an existing group (when group is full/not full)	N	**Y***	**Y***	**N**

*Used to denote that the same issue was detected by two or more personas

Most of the issues detected were minor, causing frustration or delays, but not preventing the persona from completing their task. However, there were four notable exceptions. An unintuitive process for deleting project boards in Trello detected by three personas, a hidden link to recorded sessions in Canvas found by two personas, locating student groups for joining in Canvas found by two personas and an "Auto Increment" checkbox in the MySQL Workbench



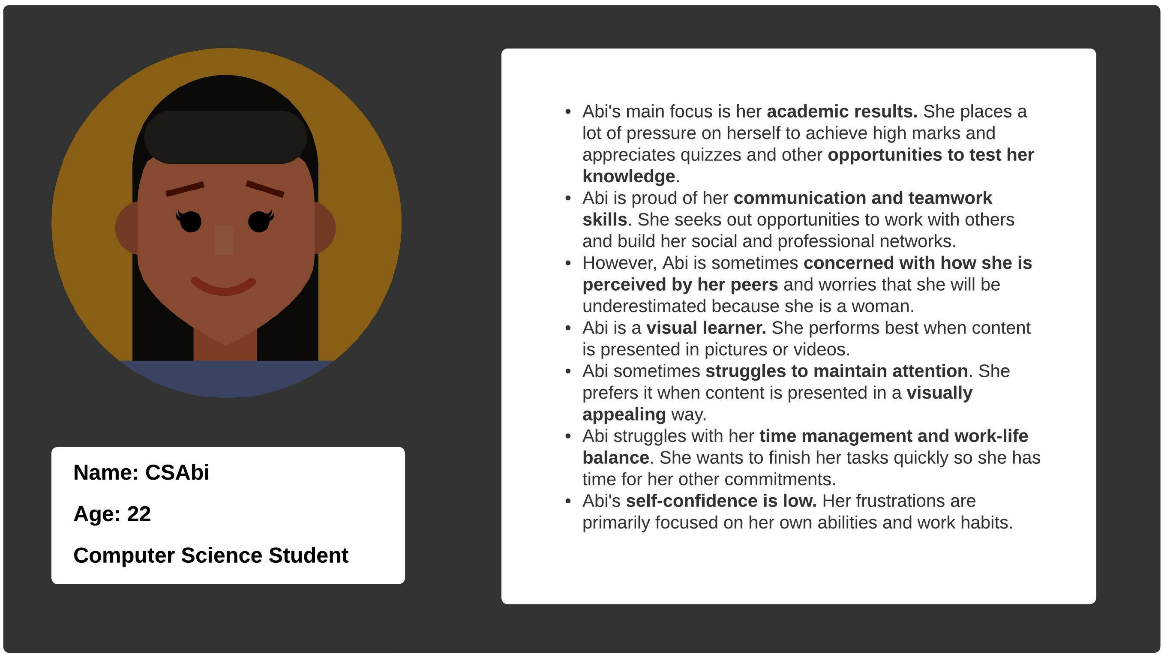


Figure 5: Preliminary CS/SE Student Personas

not directly accessible to users (Figure 6) found with CSTim. The greying out of this checkbox may mislead users that the option is not currently available. They actually need to find and check a separate, less clearly labelled, checkbox to add the AUTO_INCREMENT keyword to their script.

Using GenderMag's Abi revealed the most issues, followed by the two CS/SE student personas. Only two of CSTim's facets were relevant to software usability: frustration with complex UI designs and navigation, and preference for simple menus and limited steps. In contrast five of CSAbi's facets were raised during the walkthroughs. These included preference for visual learning, difficulty maintaining attention when content is not presented in an appealing way, struggles with time management, communication and teamwork skills, and low self-confidence.

However, this persona detected only three unique issues, and both were minor.

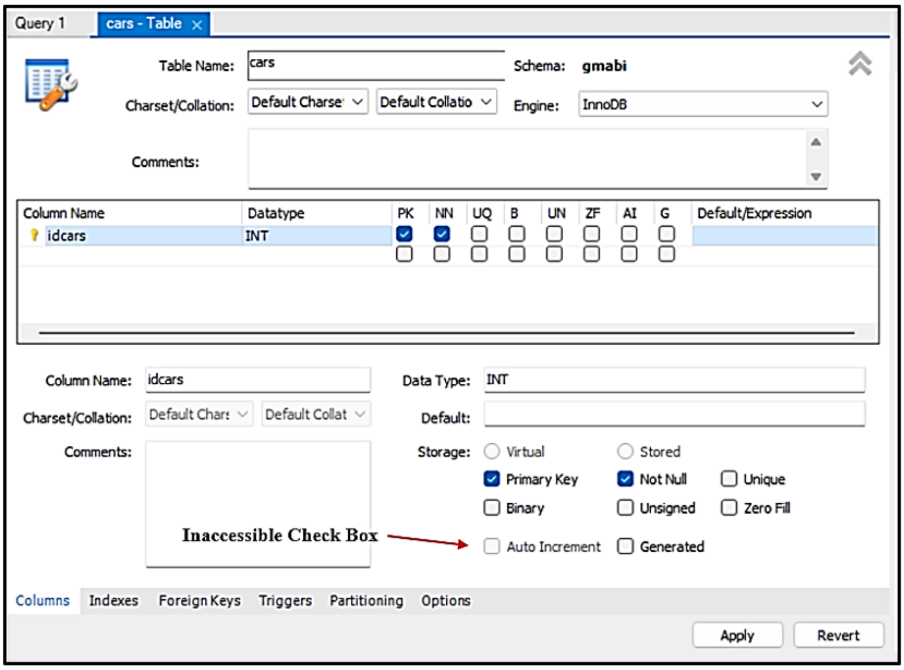


Figure 6: MySQL Auto-Increment

There was overlap between issues raised by CSAbi's low self-confidence and the GenderMag dimension of computer self-efficacy, and all five GenderMag facets were raised when conducting walkthroughs with the Abi persona.

6. Discussion

To understand the facets of CS/SE students that influence their learning experience, we adopted a workshop approach to develop CS/SE student personas. From the personas developed by the CS/SE students themselves, we could identify a wide range of facets. Considering only those that were described in multiple personas, we developed CSTim and CSAbi personas with clearly differentiating facets. Our research findings indicate that workshops can be very useful for developing personas for use in InclusiveMag framework scoping.

The usability evaluation with the GenderMag and newly developed CS/SE personas revealed that, although the GenderMag Abi persona is very effective, CSTim and CSAbi could detect some issues not found by the GenderMag personas. Since GenderMag is developed with problem solving software in mind, all the facets of the GenderMag personas may not be applicable for CS education software. Our findings indicate that some of gender differences of CS/SE students are not captured by the GenderMag personas, and this could influence their experience with CS education software. For example CSTim's facets of "disliking complex UI design and navigation" and "preference for simple menu structures" helped to identify issues such as - a hidden link to recorded sessions in Canvas and potential struggle with deleting a Trello board. Previous research applying GenderMag to review Canvas [23], mostly focused on the content of the courses, however our findings indicate that together with GenderMag's personas, CSTim and CSAbi were also helpful to find issues in the education software.

There was an extensive list of qualities found in the personas created by CS/SE students. Only a subset of those, that were present in multiple personas, were used for CSTim and CSAbi persona development. This extensive list of all qualities needs further investigation if we are to understand CS/SE student characteristics, develop a more gender inclusive learning and teaching environment, and refine the personas to better detect bias within CS education software and/or learning contents, as a whole.

In the following subsections we discuss our findings for each RQs stated in Section 3.

6.1. RQ1: The Facets of CS/SE Students

Thematic analysis of the workshop responses revealed a comprehensive list of personal qualities relevant to the study experience of CS/SE students. Personas were split into gender groups and facets extracted where a quality was present only in one group or expressed in at least half of the personas in a group. Non-binary students were insufficiently represented in the data, so only male and female student personas were reconstructed based on the results.

Within our sample, gender differences were found in preferences, motivations, goals, frustrations, and constraints. With the exception of self-confidence, they do not appear to be captured in the GenderMag personas, and there is some mixed support for them within the literature.

The social skills and goals expressed exclusively in the representations of female CS/SE students, seem to suggest they may be more comfortable and adept in group work situations than male CS/SE students. There is also some support for this in studies of cooperative behaviour [71] and social skills [72]. However, in a large-scale meta-analysis Balliet et al. [73] found that any gender differences in cooperative behaviour may be context dependent. Men were more cooperative while working in a single gender group while women were more cooperative when working with men. It is possible that social aspects were included in the female personas not because they are more skilled in these areas, but because these traits are viewed as being more important in women [74].

Similarly, research in product design suggests that the idea that women place greater importance on visual aesthetics may be a stereotype that is not reflected in their actual behaviour [75]. Nevertheless, in the present study, preferences for visual learning and appealing presentation were indeed exhibited during the persona testing task. While the coding tutorials were not undertaken in a real classroom, these responses may still provide credible insights into the preferences and frustrations of students when being presented with new content. In the cognitive walkthroughs, unappealing UI design was an issue detected in two of the four software applications. If the gender differences found in the sample are reflective of the wider CS/SE student population, this could negatively affect the education experience of female students.

They may also be affected by low self-confidence. The content of the one-line summaries, and nature of the constraints included in the personas, suggest that female students tend to blame themselves when facing difficulties in their studies. This is consistent with the GenderMag computer self-efficacy facet and supported by research investigating gender differences in self-compassion [76], however, may seem to contradict findings that women generally have a more external locus of control than men [77].

Locus of control is a spectrum of thinking between believing that outcomes are caused by one's own characteristics and behaviours (internal), or by outside factors (external) [78]. This is, however, domain dependent [77], and in the education context Ghazvini and Khajehpour [79] found that girls did have a greater tendency towards internal locus of control than boys. They also found gender differences in time management and self-testing behaviours comparable to other facets present in our personas.

At first glance, the male CS/SE student persona's greater concerns with software usability and navigation seem to conflict with the fact that most software is designed in a way that is most suitable for men [17], and appear to be a departure from the GenderMag Tim persona's tinkering behaviour [16]. This finding may be an anomaly stemming from our relatively small sample size and highlight the limitations of trying to draw clear conclusions based on what has been included or excluded from our participant's responses. Conversely, male CS/SE students may have higher expectations for the software they use than the general population and differ from the Tim persona in one or more of their key characteristics.

6.2. RQ2: Consistency with the GenderMag Facets

The survey scores of the students in our sample did not display the gender differences predicted by the GenderMag framework. Of note are the male participant's tendencies towards comprehensive information processing and the female participant's high computer self-efficacy. While these may be explained by the small sample size and expected individual differences in gender expression, it is also possible that these are areas where CS/SE student traits differ from those of the general population.

In their survey of college students Beyer [15] found that computer self-efficacy is a predictor of intention to undertake CS studies irrespective of gender. This suggests that women who choose to undertake CS courses will rate themselves higher on this facet than those who do not. It is interesting, however, that the high computer self-efficacy ratings of the female participants were not reflected in the persona descriptions or our own observations within the workshops. This may suggest that lower self-confidence in female CS/SE students is related to aspects of their coursework other than software use.

Another consideration could be CS students evaluating their self-efficacy differently, resulting from a better idea of "how much to learn" than other people. Although a comparison with other profession or discipline is not available in the literature, self-efficacy has been an important aspect of research within software engineering [80, 81, 82], indicating a potential difference with other disciplines.

6.3. RQ3: Applicability of CS/SE personas in usability evaluation

CSTim and CSAbi were useful in cognitive walkthrough tasks and could reveal usability issues similar to GenderMag Tim and Abi personas. CSTim identified one usability issue that was not detected by any other personas. Several facets were identified from the workshops that were used to develop CSTim and CSAbi peersonas. In the ten cognitive walkthrough tasks we carried out, two facets of CSTim and five facets of CSAbi were useful. These outcome are promising and indicate that, depending on the nature of the cognitive walkthrough tasks, different facets of CSTim and CSAbi personas can be helpful in detecting different usability issues.

6.4. RQ4: Performance of the CS/SE Student Personas compared to GenderMag

While reasonable facets could be extracted from the workshop responses, the cognitive walkthroughs revealed that only a subset of these are valuable in software usability testing. CSAbi and CSTim both found issues not detectable with the GenderMag facets, however, the GenderMag Abi persona still appears to be more effective at detecting gender bias in software, even when applying a walkthrough adapted for use in education research.

However, the preliminary personas we introduced are broad in scope, and may be able to supplement the GenderMag personas. The student qualities revealed in this study are diverse, wide-ranging, and highlight the personal qualities our student participants considered most relevant to their study experience. There was also very little overlap between these qualities and the GenderMag facets.

7. Threats to Validity

7.1. Construct Validity

Examples were needed in the workshop presentation to ensure that all participants understood what was required of them. Despite efforts made to ensure these examples were general and not directly related to education experiences, we cannot be certain that they did not influence participant responses. For example, during the pilot workshop colour was used as an example of UI design that may be perceived differently by people of different genders. This may have influenced one of the participants who said certain colours make them feel uncomfortable. This example was removed from the presentation in subsequent workshops to avoid further biasing participant responses.

7.2. External Validity

Generalisability is a major concern of this type of research. We recruited participants from one large Australian University due to our approved human research ethics protocol. The study was advertised with social media messages, via different university clubs, an on-campus leaflet distribution, and snowballing. However due to the nature of participation requirements we could only find a small number of students with the time and interest to participate. Although a rich set of male and female CS student attributes were identified in this study, the findings need to be verified with larger sample size recruited from more organizations.

Another key threat to external validity is the sample bias. As a requirement of the approved human research ethics protocol, all our advertisements clearly stated our motivation to address the issue of "gender bias" in the CS education environment. As such those who are interested in the topic, have awareness, or have any lived experience may have opted for participation. This is also evident from the fact that an almost equal number of male and female students participated in the study, as this does not reflect the gender distribution of the population. To reduce the impact of any "sample bias", we did not collect any information on their views on the topic, and instead the participants were requested to develop personas containing their or the peers' attributes as they see them. The instructions of the workshop were crafted in such a way, that no "gender" related concepts were used, so that the participants can think about the attributes of students of any gender without relating those to "gender bias".

7.3. Internal Validity

A major motivation of this research was to better understand the attributes and characteristics of male and female CS students. We used persona development workshops with CS students themselves for two reasons - (1) personas are a powerful tool for understanding end users, and (2) our research is inspired by the success of the GenderMag framework which uses persona as one of its essential elements. We found several characteristics from the persona development workshops and some significant differences between males and females. The findings were compiled in CSTim and CSAbi personas which were able to find some bias related issues in sample CS/SE software. Although these workshops were very helpful in identifying facets of CS students, we can not overlook the fact that the personas were developed by CS students themselves. The accuracy/consistency of these facets are highly dependent on the ability of the participants to engage in "self-assessment" and "role play". As such the findings need to be confirmed with a replication study.

While most participants provided detailed and personal responses to the tasks, study participants were offered a gift card to thank them for their time, and one participant indicated this was their primary motivation for participating. Conducting workshops in groups may have also limited the openness of participant responses, particularly when thinking about personal limitations and negative experiences. Furthermore, while almost all participants created a single persona based on themselves, secondary personas developed based on friends or classmates may be affected by the participant's own assumptions and biases. Any of these factors may have influenced the validity of individual answers. As such qualities and gender differences highlighted in our results were only considered when present in the responses of multiple students.

8. Conclusion

This paper investigated tools for identifying gender bias within CS/SE education software. A series of workshops were held to identify the personal qualities of students that influence their experience in CS courses and gender differences across these qualities were used to construct personas for use in usability testing. The performance of these personas was then compared to GenderMag's Tim and Abi. Although the effectiveness of the GenderMag personas is well supported by existing research, their facets are based on the general population. The participants in our small sample were not well represented by the GenderMag facets, and more research is needed to investigate the applicability of that framework to CS/SE students as a unique group. In the future, we plan to replicate the research study, using different methods for creating CS/SE student personas, to confirm or refute the facets founds.

The CS/SE student personas did not perform as well as Abi, however, they were both able to identify usability issues not detectable with the GenderMag facets. Additionally, the research produced a detailed list of student qualities that may assist with the identification of bias within the broader CS/SE education environment. Further research investigating gender differences across these qualities in a larger and more diverse sample may allow for the development of personas and methods that can be applied to identify bias in other learning materials, not just software. Supporting the consideration of gender inclusivity in this way could improve women's course completion rates and foster greater diversity within the CS field.

9. Acknowledgement

We thank all the students who participated in the workshops. Kanij and Grundy are supported by ARC Laureate Fellowship FL190100035.

References

[1] E. F. Keller, *Reflections on gender and science*. Yale University Press, 1995.

[2] L. SAAD, "Gallup vault: A sea change in support for working women." https://news.gallup.com/vault/214328/gallup-vault-sea-change-support-working-women.aspx, 2017.

[3] Department of Industry, Science and Resources, *University enrolment and completion in STEM and other fields*. Commonwealth Government of Australia, July. 20, 2023. [Online]. Available: https://www.industry.gov.au/publications/stem-equity-monitor/higher-education-data/university-enrolment-and-completion-stem-and-other-fields.

[4] Office of the Chief Scientist, *Australia's STEM workforce: Science, technology, engineering and mathematics*. Commonwealth Government of Australia, July, 24 2020. [Online]. Available: https://www.chiefscientist.gov.au/news-and-media/2020-australias-stem-workforce-report.

[5] A. Yatsko and W. Suslow, "Areas of computer science," in *Insight into Theoretical and Applied Informatics*, Germany: Walter de Gruyter GmbH, 2016.

[6] Department of Industry, Science and Resources, *STEM Equity Monitor Data report 2022*. Commonwealth Government of Australia, Sept. 2022. [Online]. Available: https://www.industry.gov.au/sites/default/files/2022-09/stem-equity-monitor-data-report-2022.pdf.

[7] L. J. Barker, C. McDowell, and K. Kalahar, "Exploring factors that influence computer science introductory course students to persist in the major," *ACM Sigcse Bulletin*, vol. 41, no. 1, pp. 153–157, Mar. 2009. https://doi.org/10.1145/1539024.1508923.

[8] E. Höhne and L. Zander, "Belonging uncertainty as predictor of dropout intentions among first-semester students of the computer sciences," *Z Erziehungswiss*, vol. 22, no. 5, pp. 1099–1119, Nov. 2019. https://doi.org/10.1007/s11618-019-00907-y.

[9] P. Medel and V. Pournaghshband, "Eliminating gender bias in computer science education materials," in *Proc 2017 ACM SIGCSE Tech.*, pp. 411–416, Mar. 2017. https://doi.org/10.1145/3017680.3017794.

[10] F. Henwood, "Exceptional women? gender and technology in u.k. higher education," *IEEE Technology and Society Magazine*, vol. 18, no. 4, pp. 21–27, 1999.

[11] S. Cheryan, S. A. Ziegler, A. K. Montoya, and L. Jiang, "Why are some stem fields more gender balanced than others?," *Psychological Bulletin*, vol. 143, no. 1, pp. 1–35, Jan. 2017. https://doi.org/10.1037/bul0000052.

[12] J. M. Cohoon, "Gendered experiences of computing graduate programs," *SIGCSE Bull.*, vol. 39, no. 1, p. 546–550, Mar. 2007. https://doi.org/10.1145/1227504.1227494.

[13] C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham, and J. Handelsman, "Science faculty's subtle gender biases favor male students," *Proc Nat. Acad. Sciences*, vol. 109, no. 41, pp. 16474–16479, Sept. 2012. https://doi.org/10.1073/pnas.1211286109.

[14] D. Metaxa-Kakavouli, K. Wang, J. A. Landay, and J. Hancock, "Gender-inclusive design: Sense of belonging and bias in web interfaces," in *Proc. 2018 CHI Conf. on Human Factors in Comput. Syst.*, p. 1–6, Apr. 2018. https://doi.org/10.1145/3173574.3174188.

[15] S. Beyer, "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," *Comp. Sci. Educ.*, vol. 24, no. 2-3, pp. 153–192, Oct. 2014. https://doi.org/10.1080/08993408.2014.963363.

[16] M. Burnett, S. Stumpf, J. Macbeth, S. Makri, L. Beckwith, I. Kwan, A. Peters, and W. Jernigan, "Gendermag: A method for evaluating software's gender inclusiveness," *Interacting with Comput.*, vol. 28, no. 6, pp. 760–787, Nov. 2016. https://doi.org/10.1093/iwc/iwv046.

[17] I. Nunes, A. Moreira, and J. Araujo, "Conceptual modeling of gender-inclusive requirements," in *International Conference on Conceptual Modeling*, pp. 395–409, Springer, 2021. https://doi.org/10.1007/978-3-030-89022-3_31.

[18] M. Vorvoreanu, L. Zhang, Y.-H. Huang, C. Hilderbrand, Z. Steine-Hanson, and M. Burnett, "From gender biases to gender-inclusive design: An empirical investigation," in *Proc. 2019 CHI Conf. on Human Factors in Comput. Syst.*, pp. 1–14, May 2019. https://doi.org/10.1145/3290605.3300283.

[19] T. Adlin and J. Pruitt, *The essential persona lifecycle: Your guide to building and using personas.* Burlington, Massachusetts, USA: Morgan Kaufmann, 2010. https://doi.org/10.1016/C2009-0-62475-2.

[20] T. Miaskiewicz and K. A. Kozar, "Personas and user-centered design: How can personas benefit product design processes?," *Design Studies*, vol. 32, no. 5, pp. 417–430, Sept. 2011. https://doi.org/10.1016/j.destud.2011.03.003.

[21] C. Mendez, L. Letaw, M. Burnett, S. Stumpf, A. Sarma, and C. Hilderbrand, "From gendermag to inclusivemag: An inclusive design meta-method," in *2019 IEEE Symp. VL/HCC*, pp. 97–106, Oct. 2019. https://doi.org/10.1109/VLHCC.2019.8818889.

[22] A. Shekhar and N. Marsden, "Cognitive walkthrough of a learning management system with gendered personas," in *GenderIT '18*, p. 191–198, May. 2018. https://doi.org/10.1145/3196839.3196869.

[23] A. Chatterjee, L. Letaw, R. Garcia, D. U. Reddy, R. Choudhuri, S. S. Kumar, P. Morreale, A. Sarma, and M. Burnett, "Inclusivity bugs in online courseware: A field study," in *ICER '22 -Volume 1*, pp. 356–372, Aug. 2022. https://doi.org/10.1145/3501385.3543973.

[24] A. Murphy, B. Kelly, K. Bergmann, K. Khaletskyy, R. V. O'Connor, and P. M. Clarke, "Examining unequal gender distribution in software engineering," in *EuroSPI 2019*, pp. 659–671, Sept. 2019. urlxhttps://doi.org/10.1007/978-3-030-28005-5_51.

[25] Department of Industry, Science and Resources, *STEM Equity Monitor: Data highlights 2021.* Commonwealth Government of Australia, 2021. [Online]. Available: https://www.industry.gov.au/sites/default/files/May%202021/document/stem-equity-monitor-highlights-report-2021_0.pdf.

[26] Y. Dai, G. Byun, and F. Ding, "The direct and indirect impact of gender diversity in new venture teams on innovation performance," *Entrepreneurship Theory and Practice*, vol. 43, no. 3, pp. 505–528, Oct. 2019. https://doi.org/10.1177/1042258718807696.

[27] C. R. Østergaard, B. Timmermans, and K. Kristinsson, "Does a different view create something new? the effect of employee diversity on innovation," *Research Policy*, vol. 40, no. 3, pp. 500–509, 2011. https://doi.org/10.1016/j.respol.2010.11.004.

[28] K. Kohl and R. Prikladnicki, "Benefits and difficulties of gender diversity on software development teams: A qualitative study," in *Proc. SBES '22*, p. 21–30, Oct. 2022. https://doi.org/10.1145/3555228.3555253.

[29] G. Rodríguez-Pérez, R. Nadri, and M. Nagappan, "Perceived diversity in software engineering: a systematic literature review," *Empirical Softw. Engg.*, vol. 26, sep 2021.

[30] K. Silveira and R. Prikladnicki, "A systematic mapping study of diversity in software engineering: A perspective from the agile methodologies," pp. 7–10, 05 2019.

[31] A. Murphy, B. Kelly, K. Bergmann, K. Khaletskyy, R. O'Connor, and P. Clarke, *Examining Unequal Gender Distribution in Software Engineering*, pp. 659–671. 08 2019.

[32] K. Kohl and R. Prikladnicki, "Benefits and difficulties of gender diversity on software development teams: A qualitative study," in *Proceedings of the XXXVI Brazilian Symposium on Software Engineering*, SBES '22, (New York, NY, USA), p. 21–30, Association for Computing Machinery, 2022.

[33] Y. Kovaleva, J. Kasurinen, E. Kindsiko, and A. Happonen, “State-of-the-art review on current approaches to female inclusiveness in software engineering and computer science in higher education,” *IEEE Access*, vol. 12, pp. 1360–1373, 2024.

[34] S. Cheryan, V. C. Plaut, P. G. Davies, and C. M. Steele, “Ambient belonging: How stereotypical cues impact gender participation in computer science.,” *Journal of Personality and Social Psychology*, vol. 97, no. 6, p. 1045–1060, 2009.

[35] Y. Wang and D. Redmiles, “Implicit gender biases in professional software development: An empirical study,” in *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, pp. 1–10, 2019.

[36] A. Master, A. N. Meltzoff, and S. Cheryan, “Gender stereotypes about interests start early and cause gender disparities in computer science and engineering,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 48, p. e2100030118, 2021.

[37] A. Nguyen-Duc and L. Jaccheri, “Gender equality in software engineering education – a study of female participation in customer-driven projects,” in *Sustainability in Software Engineering and Business Information Management* (V. Gupta, L. Rubalcaba, C. Gupta, and T. Hanne, eds.), (Cham), pp. 39–49, Springer International Publishing, 2023.

[38] I. Al-Taharwa, “Female student participation in software engineering projects: Opportunities to model project evaluation and to improve early prediction of teamwork failure,” *Journal of Education and Practice*, vol. 11, no. 35, pp. 30–36, 2020.

[39] L. Heels and M. Devlin, “Investigating the role choice of female students in a software engineering team project,” in *Proceedings of the 3rd Conference on Computing Education Practice*, CEP ’19, (New York, NY, USA), Association for Computing Machinery, 2019.

[40] K. S. Choi, “A comparative analysis of different gender pair combinations in pair programming,” *Behaviour & Information Technology*, vol. 34, pp. 825 – 837, 2015.

[41] O. S. Gómez, M. Solari, C. Pardo, and A. Ledezma-Carrizalez, *A Controlled Experiment on Productivity of Pair Programming Gender Combinations: Preliminary Results*, pp. 197–210. 05 2017.

[42] L. Jarratt, N. Bowman, K. Culver, and A. Segre, “A large-scale experimental study of gender and pair composition in pair programming,” ITiCSE ’19, pp. 176 – 181, ACM, 2019.

[43] M. G. S. Anna Lindqvist and E. A. Renström, “What is gender, anyway: a review of the options for operationalising gender,” *Psychology & Sexuality*, vol. 12, no. 4, pp. 332–344, 2021.

[44] A. Bittner and E. Goodyear-Grant, “Sex isn’t gender: Reforming concepts and measurements in the study of public opinion,” *Political Behavior*, vol. 39, no. 4, p. 1019–1041, 2017.

[45] A. Lishinski, A. Yadav, J. Good, and R. Enbody, “Learning to program: Gender differences and interactive effects of students’ motivation, goals, and self-efficacy on performance,” in *Proc. ACM ICER ’16*, p. 211–220, Aug 2016. https://doi.org/10.1145/2960310.2960329.

[46] M. W. Jaspers, “A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence,” *International Journal of Medical Informatics*, vol. 78, no. 5, pp. 340–353, 2009.

[47] J. Nielsen, *Usability engineering.* Interactive Technologies, Cambridge, England: AP Professional, 1st edition ed., 1993. https://doi.org/10.1016/C2009-0-21512-1.

[48] T. Mahatody, M. Sagar, and C. Kolski, "State of the art on the cognitive walkthrough method, its variants and evolutions," *Intl. Journal of Human–Computer Interaction*, vol. 26, pp. 741–785, Jul. 2010. https://doi.org/10.1080/10447311003781409.

[49] M. Farzandipour, E. Nabovati, and M. Sadeqi Jabali, "Comparison of usability evaluation methods for a health information system: heuristic evaluation versus cognitive walkthrough method," *BMC Medical Informatics and Decision Making*, vol. 22, pp. 1–11, Jun. 2022. https://doi.org/10.1186/s12911-022-01905-7.

[50] R. Khajouei, M. Zahiri Esfahani, and Y. Jahani, "Comparison of heuristic and cognitive walkthrough usability evaluation methods for evaluating health information systems," *Journal of the American Medical Informatics Association*, vol. 24, no. e1, pp. e55–e60, Aug. 2016. https://doi.org/10.1093/jamia/ocw100.

[51] R. Oliwa, "The process of designing the functionalities of an online learning platform–a case study," *Teaching English with Technology*, vol. 21, pp. 101–120, Jul. 2021. Available: https://eric.ed.gov/?id=EJ1314776.

[52] H. Şahin, Muhittin and Yurdugül, "Learners' needs in online learning environments and third generation learning management systems (lms 3.0)," *Technology, Knowledge and Learning*, vol. 27, p. 33–48, Oct. 2022. https://doi.org/10.1007/s10758-020-09479-x.

[53] E. Ors, F. Palomino, and E. Peyrache, "Performance gender gap: Does competition matter?," *Journal of Labor Economics*, vol. 31, no. 3, pp. 443–499, 2013.

[54] K. Lim, Y. O. Nam, S. Eom, Y. Jang, D. Kim, and M. H. Kim, "Structural gender differences in lms use patterns among college students," *Sustainability*, vol. 12, no. 11, 2020.

[55] S. V. Kolekar, R. M. Pai, and M. Pai M.M., "Adaptive user interface for moodle based e-learning system using learning styles," *Procedia Computer Science*, vol. 135, pp. 606–615, 2018. The 3rd International Conference on Computer Science and Computational Intelligence (ICCSCI 2018) : Empowering Smart Technology in Digital Era for a Better Life.

[56] P. Yalamu, C. Chua, and W. Doube, "Does indigenous culture affect one's view of an lms interface: A png and pacific islands students' perspective," in *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, OzCHI '19, (New York, NY, USA), p. 302–306, Association for Computing Machinery, 2020.

[57] K. R. Parker, "Selecting software tools for is/it curricula," *Education and Information Technologies*, vol. 15, pp. 255–275, 2010. https://doi.org/10.1007/s10639-010-9126-8.

[58] A. Schlesinger, W. K. Edwards, and R. E. Grinter, "Intersectional hci: Engaging identity through gender, race, and class," in *Proc. CHI '17*, p. 5412–5427, May. 2017. https://doi.org/10.1145/3025453.3025766.

[59] A. Minichiello, J. R. Hood, and D. S. Harkness, "Bringing user experience design to bear on stem education: A narrative literature review," *Journal for STEM Education Research*, vol. 1, pp. 7–33, Oct. 2018. https://doi.org/10.1007/s41979-018-0005-3.

[60] I. Benest, J. Carter, and J. Chandler, "Identifying the dominant personality and attitudinal traits in computer science students," in *Proc. Int. Conf. Eng. Educ. 2003*, Jul. 2003. Available: https://api.semanticscholar.org/CorpusID:14144701.

[61] L. M. Larson, T. F. Wu, D. C. Bailey, C. E. Gasser, V. S. Bonitz, and F. H. Borgen, "The role of personality in the selection of a major: With and without vocational self-efficacy and interests," *Journal of Vocational Behavior*, vol. 76, no. 2, pp. 211–222, 2010.

[62] R. R. McCrae, P. T. Costa, and I. Psychological Assessment Resources, *NEO inventories for the NEO Personality Inventory-3 (NEO-PI-3), NEO Five-Factor Inventory-3 (NEO-FFI-3), NEO Personality Inventory-Revised (NEO PI-R) : professional manual.* PAR, 2010.

[63] W. W. F. Lau and A. H. K. Yuen, "Gender differences in learning styles: Nurturing a gender and style sensitive computer science classroom," *Australasian Journal of Educational Technology*, vol. 26, Dec. 2010. https://doi.org/10.14742/ajet.1036.

[64] M. Burnett, A. Peters, C. Hill, and N. Elarief, "Finding gender-inclusiveness software issues with gendermag: A field investigation," in *Proceedings 2016 CHI Conf. on Human Factors in Comput. Syst.*, CHI '16, p. 2586–2598, May. 2016. https://doi.org/10.1145/2858036.2858274.

[65] M. Denscombe, *The Good Research Guide For Small-Scale Social Research Projects.* Maidenhead, England: McGraw-Hill International (UK) Ltd, 4th ed. ed., 2010.

[66] I. Etikan, S. A. Musa, and R. S. Alkassim, "Comparison of convenience sampling and purposive sampling," *American Journal of Theoretical and Applied Statistics*, vol. 5, pp. 1–4, Jan. 2016. https://doi.org/10.11648/j.ajtas.20160501.11.

[67] D. Karolita, J. Grundy, T. Kanij, H. Obie, and J. McIntosh, "What's in a persona? a preliminary taxonomy from persona use in requirements engineering," in *Intern. Confer. ENASE 2023*, pp. 39–51, Apr. 2023. https://doi.org/10.5220/0011708500003464.

[68] Open Educational Resources (OER) Commons, "Cognitive styles," 2021. [Online.] Available: https://oercommons.org/courses/interactive-tool-cognitive-styles-slider-3/view.

[69] V. Clarke and V. Braun, "Thematic analysis," *The Journal of Positive Psychology*, vol. 12, no. 3, pp. 297–298, Dec 2017. https://doi.org/10.1080/17439760.2016.1262613.

[70] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, pp. 77–101, Jul. 2006. https://doi.org/10.1191/1478088706qp063oa.

[71] M. V. Vugt, D. d. Cremer, and D. P. Janssen, "Gender differences in cooperation and competition: The male-warrior hypothesis," *Psychological Science*, vol. 18, no. 1, pp. 19–23, Jan 2007. https://doi.org/10.1111/j.1467-9280.2007.01842.x.

[72] K. S. Groves, "Gender differences in social and emotional skills and charismatic leadership," *Journal of Leadership & Organizational Studies*, vol. 11, no. 3, pp. 30–46, 2005. https://doi.org/10.1177/107179190501100303.

[73] D. Balliet, N. P. Li, S. J. Macfarlan, and M. Van Vugt, "Sex differences in cooperation: a meta-analytic review of social dilemmas.," *Psychological Bulletin*, vol. 137, no. 6, p. 881, 2011. https://doi.org/10.1037/a0025354.

[74] L. MacNell, A. Driscoll, and A. N. Hunt, “What’s in a name: Exposing gender bias in student ratings of teaching,” *Innovative Higher Education*, vol. 40, pp. 291–303, Aug. 2015. https://doi.org/10.1007/s10755-014-9313-4.

[75] X. Dai, Y. Lin, J. Liang, and C. Yang, “Appearance for females, functionality for males? the false lay belief about gender difference in product preference,” *Journal of Consumer Research*, vol. ucad054, Sept. 2023. https://doi.org/10.1093/jcr/ucad054.

[76] L. M. Yarnell, K. D. Neff, O. A. Davidson, and M. Mullarkey, “Gender differences in self-compassion: Examining the role of gender role orientation,” *Mindfulness*, vol. 10, pp. 1136–1152, Jun 2019. https://doi.org/10.1007/s12671-018-1066-1.

[77] A. C. Sherman, G. E. Higgs, and R. L. Williams, “Gender differences in the locus of control construct,” *Psychology and Health*, vol. 12, no. 2, pp. 239–248, 1997. https://doi.org/10.1080/08870449708407402.

[78] J. B. Rotter, “Generalized expectancies for internal versus external control of reinforcement.,” *Psychological monographs: General and applied*, vol. 80, no. 1, pp. 1–28, 1966. https://doi.org/10.1037/h0092976.

[79] S. D. Ghazvini and M. Khajehpour, “Gender differences in factors affecting academic performance of high school students,” *Procedia-Social and Behavioral Sciences*, vol. 15, pp. 1040–1045, 2011. https://doi.org/10.1016/j.sbspro.2011.03.236.

[80] D. Ribeiro, R. Lima, C. França, A. Souza, I. Silva, and G. Pinto, “Understanding self-efficacy in software engineering industry: An interview study,” in *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering*, EASE ’23, (New York, NY, USA), p. 101–110, Association for Computing Machinery, 2023.

[81] Hazzan and T. Seger Guttmann, “Recruiting software practitioners: The importance of self-efficacy,” *CROSSTALK The Journal of Defense Software Engineering*, vol. 23, pp. 8–12, 06 2010.

[82] M.-T. Tsai and N.-C. Cheng, “Programmer perceptions of knowledge-sharing behavior under social cognitive theory,” *Expert Systems with Applications*, vol. 37, no. 12, pp. 8479–8485, 2010.