Mining Gender Bias: A Preliminary Study on Implicit Biases and Gender Identity in the Department of Computer Science at the Technical University of Munich

Ana Petrovska[1]*, Patricia Goldberg[1]*, Anne Brüggemann-Klein[1], and Anne Nyokabi[2]

[1] Department of Informatics, Technical University of Munich, Germany
{petrovsk, figueira, brueggem}@in.tum.de
[2] Siemens AG, Erlangen, Germany
anne.nyokabi@siemens.com

Abstract. The concept of implicit biases is widely seen in many different areas and is regarded as one of the main reasons for the gender disparity between students pursuing degrees in Computer Sciences. Since less than 20% of Computer Science students are female, the information about gender bias in this field is of extreme importance. This research aimed to investigate if and by how much the female students in our department are affected by likely gender bias in their academic life. The data collected in this research was used to evaluate the automatic association that students have towards a specific gender and the computer science field.

Keywords: Gender Bias · Computer Science · University and Academia

1 Introduction and Motivation

Unconscious bias, also known as hidden or implicit bias [1], affects all of us. Very often, each of us, unconsciously and almost instantaneously assess the people around us based on their appearance, gender, and personality traits among others. However, excellence in education, science and research can only be achieved if we select from the broadest range of talents, and that is not possible if unconscious bias is narrowing down the field due to non-scientific reasons. Unconscious bias is a term from the field of psychology which describes the bias an individual has against another person or a situation [12].

Although unconscious bias can be harmless in some present-day situations, it often has damaging, long-lasting adverse effects. Concretely, it is related to labeling and burdening a specific group of people based on their gender, skin color, age, religion, or country of origin. The process of labeling a person is automatic, and people develop these behaviors based on their background, past experiences, and general exposure to cultural attitudes and social stigmas against

---

*these authors contributed equally.

certain groups. The problem with unconscious bias is that when it affects our reasoning, leading us to make fewer fair decisions, especially in the matters that need rational thinking, deluding us from the true facts of the situations, and consequently preventing us from considering the bigger pictures. We cannot move towards narrowing the gender gap issue in the technical fields if we first do not comprehend the reasons that hold us back: the hidden, unconscious, non-rational biases that each individual (regardless the gender) has—on a personal level, which later propagates on an organizational level. Unconscious bias has been identified as the main reason for poor gender-balanced representation in the IT working environments, as well as the disproportion in the numbers of male versus female students pursuing degrees in Computer and Information Sciences worldwide, with around 20% of computer science degrees awarded to women [11]. Concretely, at the university where we conducted our study, the Technical University of Munich (TUM), the overall number of female students makes up to 34% in 2014 [8] and 36% in 2018 [7]. The numbers are somehow expected since we are a technical university with a strong focus on engineering and technical sciences, where the number of women is inherently lower. However, the disproportion of male and female students is more significant at the TUM Department of Computer Science, where only 19% of the students, 15% of the Ph.D. candidates and 11% of the professors are female [7].

Understanding the reasons for the immense gender gap and the disproportion in the numbers of our department has been the core motivation for this work. Mining the gender gap and the biases should potentially lead towards identifying the reasons behind the numbers, which should later enable us to take proactive measures for improving the current situation. In order to understand our biases, first, we need to be aware of them. Concretely, the idea behind this work initially originated from the previous efforts made towards increasing the awareness about the unconscious bias in the TUM Department of Computer Science—the *Unconscious Bias Awareness Training*, further explained in Section 2.1. The work presented in this paper, including the Unconscious Bias Awareness Training, has been conducted in the frame of Informatik-Forum Frauen (IFF) [3], also known as Women in CS @ TUM. IFF is an informal group at the TUM Department of Computer Science, working towards equal participation and support of women and other underrepresented groups at the department.

In previous research, Nosek et al. [9] studied the math-gender stereotype using a mixture of implicit and explicit tests to gather information about the math–gender relationship. Concretely, the authors investigated how the associations between 1) the feeling of belonging to a group for both genders (group membership) and 2) how well they identify themselves as part of that group (group identity), shapes one's individual preferences and performance. Depending on how strong these associations are, the attitudes towards math vary, corresponding in a more negative implicit and explicit math attitude for women but a more positive one for men. Stronger implicit math-male stereotypes were, bigger the negative impact on the women's attitude was. The authors concluded that

---

[3] https://www.in.tum.de/en/current-students/equal-opportunity/

associating the self with females and math with males made it difficult for women, even women who had selected math-intensive majors, to associate math with themselves. However, to the best of our knowledge, a similar study combining implicit and explicit tests, focusing on computer science-gender stereotypes, has not yet been conducted.

Inspired by the above mentioned Unconscious Bias Awareness Training and by the lack of available datasets containing information on the computer science-gender bias, the solution that this work proposes is an online questionnaire that can gather data from both implicit and explicit questions. Therefore, the problem that this work is solving is twofold. The first problem is from the field of psychology—the creation of the questionnaires that could collect information on the person's hidden biases, which could be later analyzed. We measured both the conscious and unconscious prejudice through explicit and implicit questionnaires, respectively. For the conscious measure, we have created explicit (self-reporting) questions on gender identity and questions relevant to the TUM Department of Computer Science. The self-report questions are particularly useful for cases in which we want to evaluate, not only what a person thinks, but also how one would explicitly present their ideas. The self-reporting questions are further explained in Section 3.1. As the unconscious measure, we used the Implicit Association Test (IAT) [4], further explained in Section 2.2. The second is a technical problem, for which we need to create an online surveying form that does not only collect the answers but also 1) collects the time that the respondent took for answering a specific question, and 2) the number of errors that the respondent makes while answering a question, which is of crucial importance for the analysis of the collected data and calculating the IAT results.

2   Previous work

2.1   Unconscious Bias Awareness Training

To lessen the hidden bias consequences, the Informatik Forum Frauen (IFF) conducted Unconscious Bias Awareness Training within TUM Department of Computer Science. The goal of this training is to promote and increase awareness of our hidden biases through real-life experiments and examples. The training has been previously held in different setups inside the university reaching a broad range of people, for example, students, tutors, Ph.D. candidates or other lecturers, who participated in different workshops or pro-seminars. [6].

2.2   Implicit Association Test

Previous work [1] has shown that in principle, the results from explicit or self-reporting questionnaires tend to be misleading, since the respondents are often (subconsciously) dishonest while answering the questions. In psychology, the proposed solution to this problem is a concept named "Implicit Social Cognition" [3], which served as a basis for developing the Implicit Association Test (IAT).

IAT, initially developed and introduced in 1998 by Greenwald et al. [4] within the frame of social psychology, is a tool designed to measure and estimate one's unconscious or implicit bias towards a specific stereotypical group. Since then, it has been researched and used to investigate biases in several groups, most prominently focusing on gender and race. To address the issue of subconscious dishonesty that is present in the explicit questionnaire, IAT relies on the concept of time. Namely, faster response times are expected while pairing concepts that indicate stronger associations, for example male and computer science, compared with linking women and computer science, for which people usually take longer because of a weak mental association. While answering the IAT questionnaire, people need to connect terms that are rapidly shown to them, which makes it difficult to fake answers.

3   Method

3.1   Questionnaire creation

Our questionnaire has two type of questions: explicit (self-reporting) questions further explained in Section 3.2, and implicit, IAT based questions, further explained in Section 3.3.

3.2   Explicit (self-reporting) questions

The first type of questions we developed was the Explicit Test. We created three sub-types of self-reporting question: *Gender Identity*, *Quality Ranking* and *Department Related* questions, classified into three categories: Matrix, Drag and Drop and Single Choice questions.

Gender Identity Questions. In order to have a better insight into who is completing the online questionnaire, we first asked two single choice questions on respondent's gender and their age. The respondents could identify with male or female gender, or could not disclose to which gender they identify with ("Rather not say" option).

The purpose of the gender identity questions was to classify the respondents and to extract their gender identity information. Table 1 shows the full list of questions. Using this type of information, we could better understand the person replying to the questionnaire, what are their personal views about their own gender, and how well are they identifying themselves with their gender. The gender identity questions are structured according to the Matrix category, in which the respondent were asked to answer if they *Strongly Agree*, *Agree*, *Disagree*, and *Strongly Disagree* with a given statement. We created eight gender identity statements in the matrix, with two different tones: "Positive" and "Negative".

Qualities Ranking. In the Qualities Ranking question the respondents were asked to rank ten given terms: *Smart*, *Emotionless*, *Geek*, *Confident*, *Disciplined*, *Independent*, *Principled*, *Opinionated*, *Attractive* and *Strong*, in the order they

Nr.	Tone	Gender	Statement
1	Positive	1-Female	I feel I fit in with other Females in my department
		1-Male	I feel I fit in with other Males in my department
2	Negative	2-Female	I feel annoyed that I am supposed to do some things just because I am a Female
		2-Male	I feel annoyed that I am supposed to do some things just because I am a Male
3	Positive	3-Female	I feel comfortable being a Female in my department
		3-Male	I feel comfortable being a Male in my department
4	Negative	4-Female	I feel people interpret my behavior based on my gender
		4-Male	I feel people interpret my behavior based on my gender
5	Positive	5-Female	I feel that my personality is similar to most Females personalities' in my department
		5-Male	I feel that my personality is similar to most Males personalities' in my department
6	Positive	6-Female	I feel that the things I like to do in my spare time are similar to what most Females in my department like to do in their spare time
		6-Male	I feel that the things I like to do in my spare time are similar to what most Males in my department like to do in their spare time
7	Negative	7-Female	I sometimes think it might be more fun to be of opposite gender
		7-Male	I sometimes think it might be more fun to be of opposite gender
8	Positive	8-Female	I think I am a good example of being a Female
		8-Male	I think I am a good example of being a Male

Table 1: List of Gender Identity questions.

consider them as relevant qualities needed for success in their career or studies (most important at the top). This is a Drag and Drop question, with the most important quality ranked at the top.

Department Related Questions. The last type of self-reporting questions contained explicit questions about the Department of Computer Science at TUM in a Matrix form. This section was mainly designed to comprehend the student's perception of the department gender equality issues. The matrix contained eight statements, regardless of the gender of the respondent. The categories of responses are: *Strongly Agree*, *Agree*, *Strongly Disagree*, *Disagree* and *Not Applicable*. The last category was added since some students may have neither an advisor nor a supervisor, thus some statements were not applicable to those students. The list of statements is shown in Table 2.

3.3 Implicit questions

The implicit questions are based on IAT [4, 10]. In our work, we focus on two categories: Arts and Computer Science; and two targets: Female and Male. For each category/target, five words were chosen as representatives (see Table 3), and the respondents were asked to associate each word to the corresponding category or target. We collect the time and the number of errors a person did

Nr.	Statement
1	Students in my department are treated equally by the staff regardless of their gender
2	I am confident that the staff of my department would address sexism
3	I have witnessed gender discrimination from the staff of my department
5	I feel that my professor treats men and women equally during the lectures
6	I feel that my supervisor/advisor aligns her/his research with mine
7	I feel that my supervisor/advisor helps me identify my training/development needs
8	I feel that my supervisor/advisor shows interest in my progress/success

Table 2: List of Department Related questions.

while associating a word to a target/category. The IAT developed was divided into seven association tasks, split into five steps, as shown in Figure 1. Steps 1, 2 and 4 are practice sessions. The final hidden bias calculation was based only on the results from steps 3 and 5, disregarding the practice tasks.

Arts	Sculpture	Music	Theater	Painting	Melody
Computer Science	Programming	Technology	Code	Mathematics	Electronics
Female	Sister	Mother	Aunt	Grandmother	Daughter
Male	Brother	Father	Uncle	Grandfather	Son

Table 3: Words which represent the categories and the targets of the IAT.

The procedure of the association tasks is the following: the screen is divided into two parts, left and right. In the first step, the respondent needs to associate a term shown in the middle of the screen with Male on the left and Female to the right. In the second step, the term shown in the middle needs to be associated to a category—Computer Science and Arts to the left and right, respectively. The third step includes two consecutive critical association tasks, in which the participants need to associate words related to Male or Computer Science, and words related to Female or Arts. The fourth step was again a practice task, whose aim is to flip the targets, thus Female is located on the left side, while Male on the right. The last step, similarly to step three, had two critical association tasks but in this step Female is associated to Computer Science and Male to Arts.

4   Data collection

With the permission of the TUM Data Protection Officer and the support from a few professors at the department, our questionnaire was filled in eight introductory and advanced lectures and practical courses, on bachelor and master level. We collected the data within a period of a month, and the students answered the questionnaire in-class supervised by at least one of the authors of this work and the lecturer of the course. In total, 457 students completed the questionnaire
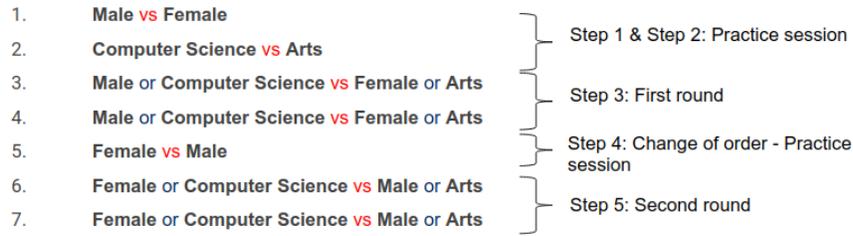
Fig. 1: Association tasks

from which 184 (41%) were females, 267 (58%) were males, and 6 students (1%) did not identify their gender.

5 Implementation of the Questionnaire

During our research we identified a lack of available—and ideally open-source— online survey tools that not only collect the respondents' answers but also the time each respondent takes to complete a section of the questionnaire. In our work, keeping track of the time was an essential factor, since calculating the IAT scores depends on the time that the respondents took to answer the question and the amount of errors that they made while doing so. The solution was to develop our own, in-house online questionnaire tool, which fulfills all the previously identified and elicited requirements that we had for our tool. The implementation of our questionnaire is open-source: `https://gitlab.com/patygold3254/hiddenbias`.

6 Results

In this paper, we explore the following two hypotheses:
Hypothesis 1: In the explicit questions, female and male students in our department give similar answers to the same question.
Hypothesis 2: In the implicit questions, each gender associates computer science easier with their own gender.

6.1 Data exploration of the explicit (self-reporting) questions

Gender Identity Questions. Figure 2 represents the percentage of replies for each statement from Table 1. The graph is sub-divided into Positive and Negative tone statements, and it is colored by the categories—from Strongly Disagree to Strongly Agree. The negative tone statements had the most accumulative percentage of Strongly Disagree and Disagree answers. Similarly, statements 1, 3 and 5 in the positive tone statements had the most accumulative percentage of Strongly Agree and Agree answers. Statements 2, 3 and 5 did not show a significant change in the answers among females and males respondents (less

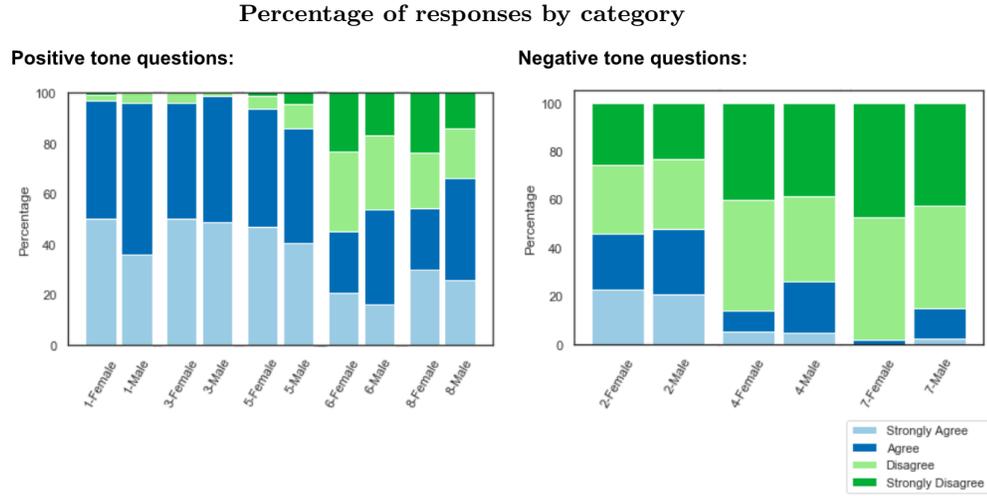**Percentage of responses by category**

Fig. 2: Percentage of replies for each statement in the Gender Identity questions.

than 5% difference in each agreement level). In this question, the data collected did not represent a significant difference in responses among females and males.

Qualities Ranking. Figure 3 shows the percentage of the words ranked in the *first place* by all of the respondents, divided by gender. Since the total number of respondents who did not identify with any gender was low (1%), the focus of this section will be only on the answers given by female and male respondents. The word with the highest percentage, among both female and male participants, was "Smart", which was ranked on a first place by 20% of the female and 30% of male respondents. On the contrary, the least amount of people ranked the word "Attractive" on the first place in their rankings. It is worth noting that the words "Emotionless", "Geek", "Opinionated" and "Strong" had twice the number of female respondents ranking it first place than male respondents.

Although the data collected from the *Gender Identity* questions do not show a significant difference between male and female respondents' answers, the data collected in *Qualities Ranking* does. This question shows a significant difference in adjectives which normally are used to describe the opposite gender. For example, words such as "Emotionless" and "Strong" that are normally associated as a male characteristic, were better positioned by female respondents as an important quality to have in order to succeed in their career or studies, rather than by the male respondents. This result shows a weaker gender identity of female students compared to male students, raising a theory that female students and professionals in the field of computer science, may be often adapting their behavior in order to be better accepted, and succeed in a male-dominated environment.

Department Related Questions. Figure 4 represents the percentage of replies for each statement related to the Department of Computer Science, colored by

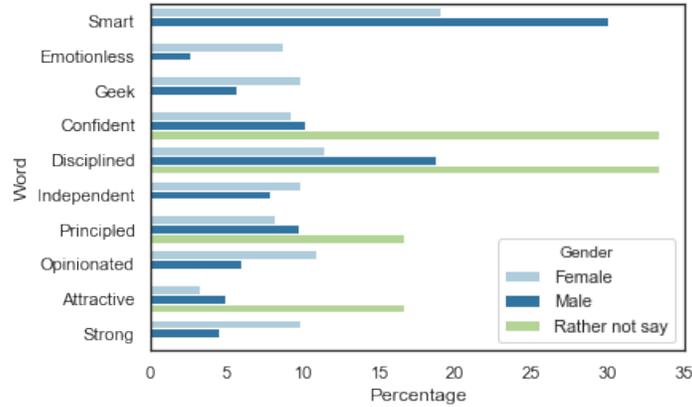**Percentage of word in first place by gender**

Fig. 3: Percentage of Qualities Ranking words ranked in first place.

the category of the answer and sub-divided by the gender of the respondent. Statements 3, 6, 7 and 8 had almost 50% of the respondents answering as "Not Applicable". We could explain this by assuming that there are students that do not have a supervisor or advisor. However, statement 3 cannot be interpreted in the same way, as it was a question about gender discrimination. Furthermore, this sentence had the highest percentage of "Disagree" answers, leading, initially, to no conclusion from the answers. However, the conclusion that we could draw is that this question was either not well formulated or did not contain enough information to be answered, since more than a half of the respondents could not give a clear agree or disagree answer to it. Importantly, in the overall matrix, no great difference could be spotted while comparing the answers of different genders, therefore the perception of the department is similar to all the students.

6.2   Data exploration of the implicit questions

The results of the implicit questions are calculated by counting the errors and response time of step 5 minus step 3 (see Figure 1). If the result is negative, the winner is female, meaning that the respondent has an automatic association between Female and Computer Science. If the result is positive, the winner is male. The automatic association term refers to the hidden bias a person has of the target and the category. This association is subdivided into 4 classifications groups: "little to no", "slight", "moderate" and "strong". "Little to no" means that the respondent demonstrates hardly any automatic association to that gender and Computer Science, while "strong" means a strong association [1].

Figure 5 represents the distribution of the IAT results' variables collected from the respondents of our questionnaire. The x-axis represents the classification categories, while the y-axis—the number of responses in that category. Each
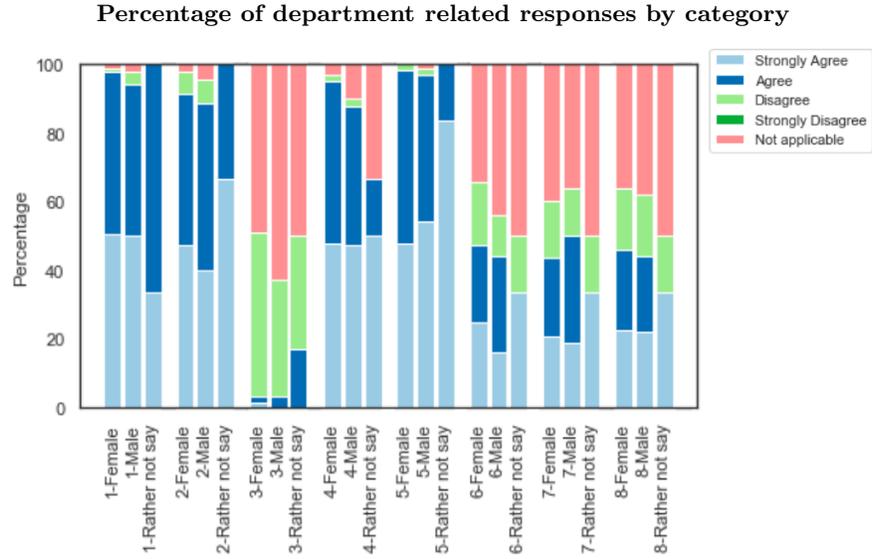
Percentage of department related responses by category

Fig. 4: Percentage of department related replies for each statement.

bar plot represents the gender of the respondent, divided by "Female" (41% of the respondents), "Male" (58%) and "Rather not say" (1%). And the variable "Winner" represents the winner of the Implicit Association Test, explained above. It is worth noticing that 60% of female respondents showed an automatic association with Female and Computer Science, and this association is spread into the classification categories, varying from "little to no" until "strong". Meanwhile, 67% of male respondents showed a stronger automatic association with Male and Computer Science, having roughly the same intensity of classification categories. The data result, shown in figure 5, show that the majority of male respondents

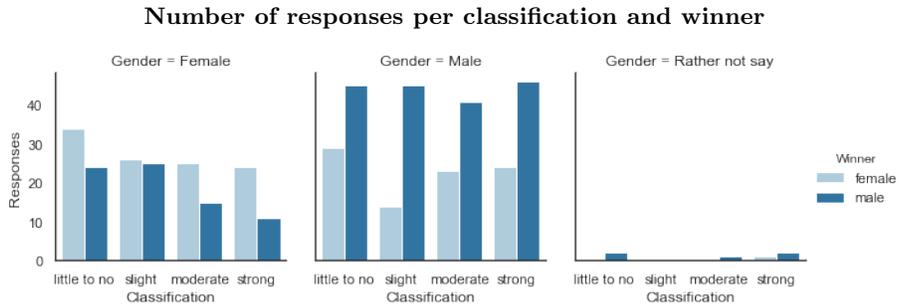**Number of responses per classification and winner**

Fig. 5: Bar plot representing the total number of responses in IAT.

demonstrated a strong automatic association between Male and Computer Science. This result leads to the conclusion that men in the department of Computer Science at TUM have a stronger implicit association between Male and Computer Science. Meanwhile, women have a less strong association between Male and Computer Science, tending to demonstrate actually a strong association between Female and Computer Science.

6.3   Data analysis of the implicit questions

Furthermore, we did data analysis on statistical evaluations of the IAT measurements. Following the IAT best practises [5], the measurements' outliers were deleted using the z-score given by $z = \frac{x-\mu}{\sigma}$, in which μ represents the mean and σ the variance [2].

Figure 6(a) represents the duration of the respondents (in seconds), divided by gender, per task. Having in mind that step 3 associates Male and Computer Science in two association tasks, and step 5 associates Female and Computer Science in two tasks also, the boxplot shows a significant wider interquartile range (IQR) of the male respondents in step 5 comparing to female respondents. Therefore, the time range of male respondents when associating Female and Computer Science was bigger than the time range of female respondents. The median of the male respondents in step 5 is significantly higher, comparing to their median in step 3. Figure 6(b) represents the errors per step, per respondent

(a) Duration of the respondents   (b) Errors of the respondents
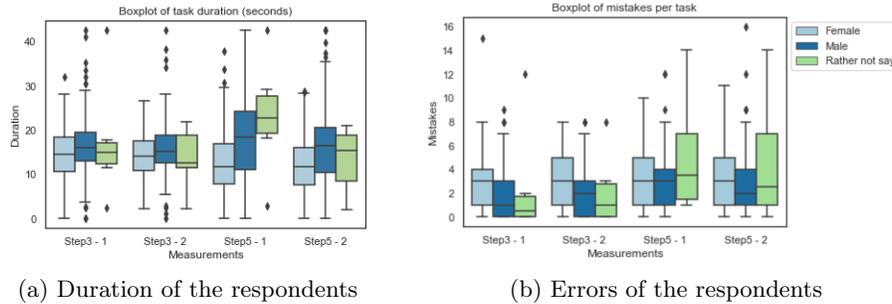
Fig. 6: Boxplot divided by gender, per task.

gender. While the boxplot for Females remains almost constant across tasks, the boxplot for the Males shows a higher IQR for step 5.

There are significant differences in the median and in the interquartile range of the duration and the number of errors when grouping the respondents per gender. This result leads to the conclusion that men take longer and make more errors than women while associating Female and Computer Science.

7   Conclusion

Since hidden biases are not trivial to measure, calculate and analyze, we propose explicit and implicit questions, based on which we collect data from our department to understand better if our students are affected by gender bias in their academic life. We have set up two hypotheses, which our results supported. Concretely, from the explicit self-reporting) questionnaire, our results showed that there is no significant difference in responses among female and male respondents on the *Gender Identity* and *Department Related* questions. Nevertheless, this differed in *Qualities Ranking*, where female students ranked characteristic usually linked to the opposite gender, as "Emotionless" and "Strong", as more important for success in their studies and careers.

Although society, media, and women's upbringing tend to picture men closer to computer science, or STEM in general, the results from our IAT questionnaire show that the majority of the of female students (60%) in our department associate computer science with female easier than with male, and 67% of the male respondents have a stronger automatic association with male and computer science. The results show that each gender associate computer science with their own gender more easily, but males do so more than females.

In summary, although we see a correlation of the respondent's gender with their hidden bias result, this study does not have sufficient data to conclude causality. As next steps we continue collecting and analyzing data from the department, that should enable us to do better statistics and draw better conclusions.

Acknowledgement

References

1. Banaji, M.R.: Blindspot : hidden biases of good people. New York : Delacorte Press, [2013] ©2013 ([2013]), `https://search.library.wisc.edu/catalog/9910203165702121`
2. Ghosh, D., Vogt, A.: Outliers: An evaluation of methodologies. In: Joint statistical meetings. pp. 3455–3460. American Statistical Association San Diego, CA (2012)
3. Greenwald, A.G., Banaji, M.R.: Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychological review **102**(1), 4 (1995)
4. Greenwald, A.G., McGhee, D.E., Schwartz, J.L.: Measuring individual differences in implicit cognition: the implicit association test. Journal of personality and social psychology **74**(6), 1464 (1998)
5. Greenwald, A.G., Nosek, B.A.: Health of the implicit association test at age 3 (2001)

6. IFF: Unconscious Bias Training (2018), `https://www.in.tum.de/en/current-students/equal-opportunity/projects/unconscious-bias-training/`
7. of Informatics, T.D.: THE DEPARTMENT OF INFORMATICS IN FACTS AND FIGURES 2018 (2018)
8. of Munich, T.U.: Diversity at tum, focus: Gender and family (2015)
9. Nosek, B.A., Banaji, M.R., Greenwald, A.G.: Math= male, me= female, therefore math$\neq$ me. Journal of personality and social psychology **83**(1), 44 (2002)
10. Nosek, B.A., Greenwald, A.G., Banaji, M.R.: Understanding and using the implicit association test: Ii. method variables and construct validity. Personality and Social Psychology Bulletin **31**(2), 166–180 (2005)
11. for Science, N.C., Statistics, E.: Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019. Tech. rep., National Science Foundation (2019)
12. Steele, C.M.: Whistling Vivaldi: And other clues to how stereotypes affect us (issues of our time). WW Norton & Company (2011)