Stereotypes of Secondary School Students Towards People in Computer Science

Laura Keil, Fatma Batur, Matthias Kramer, Torsten Brinda

Stereotypes of Secondary School Students Towards People in Computer Science

Laura Keil, Fatma Batur, Matthias Kramer, and Torsten Brinda

University Duisburg-Essen, Computing Education Research Group,
45127 Essen, Germany
https://www.ddi.wiwi.uni-due.de/en/
laura.keil@stud.uni-due.de,
{fatma.batur,matthias.kramer,torsten.brinda}@uni-due.de

Abstract. Though the computer science industry has become more and more important in recent years in Germany and elsewhere, the number of students in IT-related study programs is increasing only slowly. This results in a shortage of skilled IT workforce. Additionally, women are strongly under-represented. A possible cause of these phenomena are stereotypes towards people in computer science. But what image of computer scientists do students at German secondary schools really have? In order to get an overview, 52 upper-secondary school students were surveyed by using an online questionnaire, which included both open-ended and closed questions. The results show that the conceptions are very diverse and individual. However, some characteristics are considered more appropriate than others. For example, many students indicated that people in computer science are intelligent and good at maths and science. Fewer stated that they are team players and have good communication skills. The analysis of subgroups show tendencies that aspects such as gender, previous school experience and interest in computer science, and the personal environment of the students can influence these conceptions.

Keywords: Stereotypes · Computer Science · Conceptions · Upper-Secondary Education · K-12

1 Motivation and Aim

Technologies shape today's society and the basis of these technologies is computer science. The demand for workforce in the IT industry has risen strongly in Germany in recent years and it is to be expected that this process will continue [1, p. 11]. Therefore, the computing field provides attractive future job perspectives for today's students. Additionally, there is a skill shortage in the computer science industry in Germany [1, p. 12]. The proportion of students with study courses related to computer science is rising only slowly [2]. Furthermore, women are strongly under-represented both in the computer science industry [1, p. 15] and in study courses related to computer science [2]. Individual beliefs and stereotypes about computer science are considered to be one

explanation [3], so that targeted action against wrong conceptions of people in this field and the field itself are needed. In this paper, we describe the design and results of an online survey (which took place in North Rhine-Westphalia) that aimed to identify students' stereotypes of people in the computer science field as well as influencing factors in order to provide the basis for developing targeted actions.

2 Theoretical Background and Related Work

The reasons for the slowly growing number of students as well as the lack of specialists and in particular the under-representation of women in the IT field are very diverse. Ashcraft et al. divide the reasons for the under-representation of women in "Formal and Informal Education", "Families, Communities, and Role Models", "Peer Influences", and "Media and Popular Culture" [4, p. 17]. In their last category, they include aspects such as the media portraying the computing field as masculine and "geeky" and people in computer science as "geeks without social skills doing boring and solitary jobs" [4, p. 27-28]. Osunde et al. group the factors into "Biological differences (Essentialist Theory)", "Structural factors", and "Socio-cultural factors" [5, p. 18]. The latter includes conceptions of computer scientists and the ones found in this study are negative and imply that computer science is for "nerds" or men only [5, p. 19]. Another literature review by Sanders grouped the factors amongst other aspects into "Age, Stage and Pipeline Issues" as well as "Societal Influences" [6, p. 2], in which the media are addressed as a factor. An analysis of computer advertising, the Internet, television and movies revealed gender stereotypes of people in technical roles [6, p. 5].

According to Leyens et al., stereotypes are defined as "shared beliefs about person attributes, usually personality traits, but often also behaviours, of a group of people" [7, p. 11]. Stereotypes are thus generalised judgements and ideas, which are assigned to a group [8, p. 315]. The term focuses on the similarities of group members, while differences move into the background. Often, stereotypes are negative shared beliefs. In rarer cases, they could also be positive or have positive components [8, p. 315]. Since stereotypes often played a role in the above phenomena, we now look into studies focusing on stereotypes.

Studies dealing exclusively with stereotypes towards people in the field of computer science show that students have stereotypes towards people in this field and that these stereotypes can have an influence on their attitudes towards computer science. Negative attitudes can then prevent students from working towards a future career in this field while positive conceptions could encourage them. The results of the first part of a two-part study by Cheryan et al. (2013) with around 300 participants showed that stereotypes exist for both females and males. The evaluation of the participants' descriptions of people in computer science showed in particular that they are described as having features that are incompatible with the female gender role. Participants often mentioned that these people lack interpersonal skills or are singularly focused on computers [3,

p. 63]. In addition, it has been found that women, who have not taken a computer science class, have more often mentioned one of the stereotypical categories in describing people in the computer science field than those, who have had at least one computer science class [3, p. 64]. The results of the second part with around 50 participants showed that stereotypes can affect attitudes towards computer science (especially for women) and that media can have an impact on stereotypes [3, p. 66-67]. Another study with 100 female students indicated that an interaction with a stereotypical person in the field of computer science has a negative effect on the interest in computer science for women both in the short and in the long term [9, p. 74-75].

The above-mentioned studies give indication that stereotypes can influence both the skill shortage in the computer science industry as well as the under-representation of women by having a negative impact on their interest. There are also indications that perceptions can also be influenced by the media and by prior computer science education. Since the presentation of computer science in the media is also a subject of change, the results found in earlier studies and in other countries might not be fully applicable to the situation in Germany, so that we decided to investigate German school learners' stereotypes about computer science. We decided to survey 10th and 11th graders, as they may have already thought about future career plans. Summing up, the main research question of the study described in this paper is: Which stereotypes towards people in the field of computer science can be identified among German upper-secondary school students? The question is split into the following sub-questions: 1. Which specific ideas of people in the field of computer science do the students have? 2. Which influence do factors such as gender, previous school experience in the field of computer science, interest in computer science, or the personal environment of the students have on these ideas?

3 Research Method

3.1 Data Collection

The study was carried out at four upper secondary schools in the Dusseldorf area (Germany, North Rhine-Westphalia) in May 2019 using an online questionnaire[1].

The questionnaire was structured as follows: The first part contained open-ended questions about students' understanding of computer science as a discipline or a working field (*"Describe your conception of the subject or the industry 'computer science'."*) and their ideas of people in the field of computer science (*"Describe a person who deals with computer science. What do you think are their characteristics?"*). The second part of the questionnaire contained closed questions. On the one hand, the students were asked to evaluate predefined statements about people in the field of computer science (e.g.: *"People in the field of*

[1] You can download the questionnaire in English as a PDF file here: `https://udue.de/questionnairestereotypes`

computer science are intelligent."), and, additionally, they were given a predetermined set of characteristics to choose from which, in their opinion, most likely applied to people in the field of computer science. The third part of the questionnaire was designed to collect students' demographic and sociocultural information in order to build subgroups based on this information. The subgroups were built depending on the gender information, whether the students had received computer science education at school (*"How many years have you had computer science lessons or were in another computer science group until now?"*); their interest in computer science (*"Are you interested in computer science?"*), and if they had people in their lives that worked or studied in the computer science field (*"Do you have family members who deal with computer science? Do you have acquaintances or friends who deal with computer science?"*).

Both the given statements and the given characteristics are based on a theoretically derived category system, which consists of 14 categories. Categories C1 to C6 are based on the categories that are described by Cheryan et al. [3, p. 59-60], categories C7 to C12 were created by analysing computer science related job descriptions and extracting frequently required characteristics. C13 is built on the JIM study 2018. It indicates that digital games "play a central role in the media life of adolescents" [10, p. 79]. Finally, category C14 is based on the results of the paper [11], which shows that many students relate computer science with mathematics and science. Table 1 gives an overview of the described category system. The **bold** descriptions will be used as abbreviation in the rest of the text.

Table 1. Category system

ID	Description
C1	**Intelligent**.
C2	**Male**.
C3	Interested in **new technologies**.
C4	Special **external features**.
C5	**Lack of social skills**, introverted, or not empathetic.
C6	**Focused on computers**.
C7	**Work independently**.
C8	**Team player.**
C9	**Logical thinking**, planning-oriented, structured, and organized.
C10	Good **communication skills**.
C11	**Motivated**, hardworking, and eager to learn.
C12	**Creative.**
C13	Interested in **computer games**.
C14	Good skills in **mathematics and science**.

A total of 55 students from four different courses and from four upper secondary schools in the above mentioned region were interviewed. The relevant sample for the evaluation was 52 students after excluding three answers due to

incompleteness. The students were 15 to 20 years old and attended the 10th or 11th grade at the time of the survey. 46% (24) of the students were female, 33% (17) male, and 21% (11) made no indication of gender. 38% (20) of the students had previous school experience in the field of computer science, 35% (18) of the students expressed interest in computer science, and 69% (36) of the students had family members or friends, or acquaintances that deal with computer science.

3.2 Evaluation

Overall, the following evaluation focused on the open-ended question about the ideas of people in the field of computer science and the predefined statements about people in the field of computer science. The questions are evaluated on the one hand for the entire sample and on the other hand for the above mentioned subgroups in order to be able to identify any differences depending on the gender, the previous school experience, the interest, or the environment of the students.

For the evaluation of the open-ended questions the deductively derived category system described above was used (see Table 1). After viewing the data, additional categories were added inductively [12, p. 85]. Once all statements were assigned, the categories were sorted according to the frequency of their naming. The evaluation of the predefined statements was done as follows: The arithmetic average i.e. mean was calculated (depending on the distribution and weighting of the answer options per statement). The weighting of the answer options takes into account the formulation of the statement, since some statements were made positive and others negative. A high mean implies that students strongly agree with the category to which the statement belongs, while a low mean implies that students agree little. The lowest possible value is 1 and the highest possible is 4.

4 Results

4.1 Open-ended question

Table 2 gives an overview of the results of the open-ended question about the ideas of people in the field of computer science. It can be seen that in addition to the above-mentioned deductive category system (categories C1-C14), the categories M1-M6 have been added inductively, as some of the students' statements could not get assigned to the existing categories.

The results of the entire sample can be described as follows: The most frequently identified categories are C4 *(Special external feature)* with 31% (16) and the categories C1 *(Intelligent)* and C14 *(Good skills in mathematics and science)* with 29% (15) each. In addition, the categories C5 *(Lack of social skills, introverted, or not empathetic)* and C9 *(Logical thinking, planning-oriented, structured, and organized)* are frequently identified with 23% (12) each and the categories C3 *(Interested in new technologies)* and C6 *(Focused on the computer)* with 19% (10) each. Not identified are the categories C7 *(Work independently)*,

C8 *(Team player)*, and C10 *(Good communication skills)*. The other categories are identified between 2% (1) and 17% (9). It can be stated that in particular, the categories that are known from previous studies (C1-C6) can be identified in the answers very often. With the exception of category C2 *(Male)*, these are identified in at least 19% (10) of the students' answers.

Table 2. Results of the open-ended question

ID	Description	Total	F	M	C	No C	I	No I	E	No E
C4	External features.	31%	**25%**	**35%**	20%	**38%**	**28%**	**32%**	**31%**	31%
C1	Intelligent.	29%	**33%**	**24%**	**30%**	**28%**	11%	**38%**	19%	**50%**
C14	Maths and science.	29%	**38%**	**24%**	15%	**38%**	6%	**41%**	22%	**44%**
C5	Lack of social skills.	23%	13%	**35%**	20%	25%	11%	29%	19%	31%
C9	Logical thinking.	23%	**33%**	6%	20%	25%	**33%**	18%	**25%**	19%
C3	New technologies.	19%	21%	24%	**25%**	16%	17%	21%	**25%**	6%
C6	Focused on computers.	19%	17%	18%	15%	22%	11%	24%	19%	19%
M4	Nerdy.	17%	8%	18%	20%	16%	**22%**	15%	19%	13%
M5	No external features.	15%	25%	6%	20%	13%	17%	15%	22%	0%
C13	Computer games.	13%	13%	6%	0%	22%	6%	18%	14%	13%
M3	Relaxed.	12%	13%	12%	15%	9%	17%	9%	11%	13%
C2	Male.	10%	17%	0%	0%	16%	6%	12%	8%	13%
C12	Creative.	10%	8%	6%	10%	9%	**22%**	3%	11%	6%
M2	Unsporting.	10%	4%	18%	15%	6%	17%	6%	8%	13%
M1	Not generalizable.	4%	0%	12%	5%	3%	6%	3%	3%	6%
M6	Certain nationality.	4%	0%	12%	0%	6%	0%	6%	3%	6%
C11	Motivated.	2%	4%	0%	5%	0%	6%	0%	3%	0%
N	**Total**	**52**	**24**	**17**	**20**	**32**	**18**	**34**	**36**	**16**

F: Female; **M:** Male; **C:** CS Education Received; **No C:** No CS Education Received; **I:** Interested in CS; **No I:** Not Interested in CS; **E:** Environment with CS; **No E:** No Environment with CS; *Highest values per subgroup in* ***bold print*** *and values referred to the text highlighted in* *grey*

When comparing the results of the two subgroups of self-identified male and female students, strong differences in the responses can be seen for the categories C9 *(Logical thinking, planning-oriented, structured, and organized)*, C5 *(Lack of social skills, introverted, or not empathetic)*, and M5 *(No special external features)*. While 33% (8) of the female students describe people in the field of computer science as logical thinking, planning-oriented, structured, and organized only 6% (1) of the male students do so (C9). 35% (6) of the male students state that people in the field of computer science are lacking social skills, are introverted, or not empathetic. Only 13% (3) of the female students mention aspects in this category (C5). 25% (6) of the female students state that people in the field of computer science do not have any special external features while only 6% (1) of the male students do so (M5).

When comparing the results of the two subgroups of students with and without prior educational background in computer science, the strongest differences can be seen for the categories C14 *(Good skills in mathematics and science)*, C13 *(Interested in computer games)*, and C4 *(Special external features)*. While 38% (12) of the students without school experience in the field of computer science describe people in this field as with good skills in mathematics and science, only 15% (3) of the students with school experience do so (C14). A similar difference can be seen for the category C13. 22% (7) of those students without school experience describe people in the field of computer science as interested in computer games, while none of the students with school experience do so. In addition, 38% (12) of those without school experience describe people in the field of computer science with special external features, whereas only 20% (4) of those with school experience mention such aspects (C4).

The results for the subgroups depending on the interest and the environment of the students also show interesting differences and similarities and can be seen in Table 2.

4.2 Closed questions

All calculations have been done with the R environment [13], including the packages *psych* [14] and *pastecs* [15]. Table 3 gives an overview of the results of the predefined statements. The results of the entire sample can be described as follows: The category with the highest mean of 3.4 ($SD = 0.69$) is C7 *(Work independently)*. Similar high means of 3.38 ($SD = 0.66$) and 3.34 ($SD = 0.58$) are to be found in the categories C3 *(Interested in new technologies)* and C6 *(Focused on the computer)*. Students overwhelmingly tend to agree that computer scientists fulfill these stereotypes. The lowest mean is assigned to category C10 *(Good communication skills)* with a mean of 2.31 ($SD = 0.83$). All other categories have a mean between 2.5 and 3.31 ($0.54 < SD < 1.02$). Looking only at the categories C1-C6 resulting from the mentioned study, it can be seen that categories with a positive connotation such as C3 *(Interested in new technologies)* or C1 *(Intelligent)* are in the upper half of the descending order, while categories addressing the often portrayed computer scientist such as C5 *(Lack of social skills, introverted, or not empathetic)*, C2 *(Male)*, and C4 *(Special external features)* are in the lower half. Of particular interest might be the correlations between the categories. The highest positive correlation of $r = 0.58$ can be found between categories C8 *(Team player)* and C10 *(Good communication skills)*. Interestingly, the same category C8 is moderately negative correlated to C4 *(Special external features)* with $r = -0.55$, C5 *(Lack of social skills, introverted, or not empathetic)* with $r = -0.58$ and C6 *(Focused on the computer)* with $r = -0.63$. Due to the presence of a large number of underlying categories, the items are not expected to contribute in the same way to the result. Therefore, coefficients for neither τ-equivalent nor τ-congeneric reliability were calculated.

Furthermore, Table 3 shows that the five most agreeable categories in the whole group of assessed students (C7, C3, C6, C14, and C1) are the same for the subgroups of self-identified female and male students. They merely differ

Table 3. Results of the closed questions

ID	Description	Total	F	M	C	No C	I	No I	E	No E
C7	Work independently.	3.40	3.42	**3.41**	**3.60**	3.28	**3.67**	**3.26**	**3.39**	**3.44**
C3	New technologies.	3.38	**3.38**	**3.41**	**3.35**	**3.41**	**3.33**	**3.41**	**3.39**	**3.34**
C6	Focused on computers.	3.34	3.19	**3.44**	3.30	3.36	3.31	3.35	**3.34**	3.25
C14	Maths and science.	3.31	**3.5**	3.29	3.10	**3.44**	3.06	3.44	3.33	3.25
C1	Intelligent.	3.29	**3.38**	3.29	3.15	**3.38**	3.11	**3.38**	3.19	**3.5**
C13	Computer games.	3.15	3.17	3.12	3.05	3.22	2.89	3.29	3.17	3.13
C12	Creative.	3.10	3.13	3.12	3.25	3.00	**3.44**	2.91	3.06	3.19
C9	Logical thinking.	3.00	2.92	3.00	3.13	2.91	3.22	2.87	2.93	3.13
C11	Motivated.	2.92	3.04	2.82	2.85	2.97	2.67	3.06	2.86	3.06
C5	Lack of social skills.	2.83	2.75	2.84	2.72	2.9	2.61	2.94	2.78	2.94
C2	Male.	2.69	2.54	2.65	2.50	2.81	2.61	2.74	2.69	2.69
C8	Team player.	2.64	2.60	2.63	2.69	2.57	2.82	2.51	2.55	2.77
C4	External features.	2.50	2.38	2.65	2.45	2.53	2.39	2.56	2.56	2.38
C10	Communication skills.	2.31	2.29	2.29	2.40	2.25	2.56	2.18	2.25	2.44
N	**Total**	**52**	**24**	**17**	**20**	**32**	**18**	**34**	**36**	**16**

F: Female; **M:** Male; **C:** CS Education Received; **No C:** No CS Education Received; **I:** Interested in CS; **No I:** Not Interested in CS; **E:** Environment with CS; **No E:** No Environment with CS; *Highest values per subgroup in* ***bold print*** *and values referred to the text highlighted in* *grey*

in their order. Female students mostly agree with categories C14 *(Good skills in mathematics and science)*, C7 *(Work independently)*, C3 *(Interested in new technologies)*, and C1 *(Intelligent)*. Whereas male students mostly agree with categories C6 *(Focused on the computer)*, C7 *(Work independently)*, and C3 *(Interested in new technologies)*. As can be seen from Table 3, the means are all very close with a range between 3.19 and 3.5. Interestingly, the most agreeable aspect of one group scores only fifth place in the other group and vice versa. As tempting as it might be to infer anything from those numbers, a Wilcoxon rank sum test provided neither differences in the distributions for category C14 ($W = 174, p = 0.38$) nor for C6 ($W = 250, p = 0.21$).

Comparing the overall differences in the means of the categories for the female and the male students, large differences can also be observed in the categories C4 *(Special external features)* and C11 *(Motivated, hardworking, and eager to learn)*. In neither case did the Wilcoxon rank sum test provide a significant difference in the distributions. This could be a result of a significant part (about 21%) of participants not indicating any gender. Therefore, the already small sample of 52 individuals gets further reduced to 41 and then split into groups.

Table 3 also displays the differences of students who have already received computer science education or have participated in an extracurricular lab class. For the students with school experience in the field of computer science, the category with the highest mean is category C7 *(Work independently)* with a mean of 3.6 ($SD = 0.6$). The categories with the second and the third highest mean are the categories C3 *(Interested in new technologies)* with a mean of 3.35

($SD = 0.67$) and C6 *(Focused on the computer)* with a mean of 3.3 ($SD = 0.64$). For the students without any school experience in the field of computer science, the category C14 *(Good skills in mathematics and science)* has the highest mean of 3.44 ($SD = 0.62$). The categories C3 *(Interested in new technologies)* and C1 *(Intelligent)* rank second and third with means of 3.41 ($SD = 0.67$) and 3.38 ($SD = 0.55$). Again, categories can be found in which the means of the two subgroups differ strongly. Examples are the categories C14 *(Good skills in mathematics and science)* (3.44 vs. 3.1) and C7 *(Work independently)* (3.6 vs. 3.28). Yet, as in the former subgroup analysis regarding gender, the Wilcoxon rank sum test provided no significant differences between the groups, although the result for category C14 ($W = 231, p = 0.06$) could be indicative and should be investigated in further studies.

The results for the subgroups depending on the interest and the environment of the students show interesting differences and similarities, can be seen in Table 3.

5 Discussion and Conclusions

Results of the first and second part are very different for some categories, while they are very similar for other categories. For example, in the whole group category C4 *(Special external features)* is the most frequently mentioned in the first part of the survey, but simultaneously one of the least frequently mentioned in the second part. Notably, some students chose different open-ended answers which are in contrast to their selected statements of the closed questions. To obtain more insights into this opposing result, further investigation is needed.

Overall, the results show that students' perceptions of people in the field of computer science are very diverse and individual. This study provides an overview and basis for further investigations. First tendencies regarding rarer and more frequent ideas can be recognised. Compared to the literature (see section 2) this study shows similar results on students' perceptions. In particular, the results of the open-ended questions make it clear that many features have been named that fit the image of the "computer nerd". Additionally, first tendencies for the fact that factors such as gender or computer science education, but also the interest in computer science or the availability of people in the social environment who work in the field of computer science, have an impact on the students' ideas towards people in this field. To make a more general statement in this regard, the sample of subgroups is insufficient. Nevertheless, the results show that female students seemingly do have negative connoted images of people who deal with computer science.

To get a closer look, further studies could focus on factors separately at a time to study its influence more closely. Especially, studies on female students' conceptions on a larger group with different ages should be considered. Concerning targeted actions to deliver a positive and realistic picture of computer scientists among youth, there is still a long way to go. Mandatory and positively connoted informatics education for all might help to create a more realistic pic-

ture of the computer science industry in order to reduce problems such as the skill shortage in the industry or the under-representation of women in this field.

References

1. Statistik der Bundesagentur für Arbeit: IT-Fachleute., Nürnberg. (2019)
2. Destatis (Federal Office for Statistics): Students: Germany, semester, nationality, sex, subject (computer science, informatics) (2019)
3. Cheryan, S., Plaut, V.C., Handron, C., Hudson, L.: The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. Sex roles **69**(1-2) (2013) 58–71
4. Ashcraft, C., Eger, E., Friend, M.: Girls in IT: The facts. Boulder, CO: National Centre for Women & Information Technology (NCWIT). (2012)
5. Osunde, J., Windall, G., Bacon, L., Mackinnon, L.: Female under-representation in computing education and industry-A survey of issues and interventions. International Journal of Advanced Computer Science and Applications **5**(10) (2014)
6. Sanders, J.: Gender and technology in education: A research review (2005)
7. Leyens, J.P., Yzerbyt, V., Schadron, G.: Stereotypes and social cognition. Sage Publications, Inc, London (1994)
8. Batinic, B., Appel, M., et al.: Medienpsychologie. Springer Medizin Verlag Heidelberg, Berlin, Heidelberg (2008)
9. Cheryan, S., Drury, B.J., Vichayapai, M.: Enduring influence of stereotypical computer science role models on womens academic aspirations. Psychology of Women Quarterly **37**(1) (2013) 72–79
10. Medienpädagogischer Forschungsverbund Südwest: JIM-Studie 2018. Jugend, Information, Medien, Stuttgart. (2018)
11. Brinda, T., Napierala, S., Behler, G.A.: What do secondary school students associate with the digital world? In: Proceedings of the 13th Workshop in Primary and Secondary Computing Education. WiPSCE 18, New York, NY, USA, Association for Computing Machinery (2018)
12. Mayring, P.: Qualitative content analysis. A companion to qualitative research 1. Weinheim (2004)
13. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. (2018)
14. Revelle, W.: psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. (2018) R package version 1.8.10.
15. Grosjean, P., Ibanez, F.: pastecs: Package for Analysis of Space-Time Ecological Series. (2018) R package version 1.3.21.