



Why do women pursue a Ph.D. in Computer Science? ☆

Erika Ábrahám ^a, Miguel Goulão ^b,* Milena Vujošević Janičić ^c, Sarah Jane Delany ^d,
Amal Mersni ^e, Oleksandra Yeremenko ^f, Özge Büyükdaglı ^e, Karima Boudaoud ^g,
Caroline Oehlhorn ^h, Ute Schmid ^h, Christina Büsing ^a, Helen Bolke-Hermanns ^a,
Kaja Köhnle ^a, Matilde Pato ⁱ, Deniz Sunar Cerci ^j, Larissa Schmid ^k

^a Department of Computer Science, RWTH Aachen University, Aachen, Germany

^b Department of Computer Science, NOVA School of Science and Technology, NOVA LINGS, Lisboa, Portugal

^c Department of Computer Science, Faculty of Mathematics, University of Belgrade, Belgrade, Serbia

^d School of Computer Science, Technological University Dublin, Dublin, Ireland

^e Department of Engineering, Faculty of Engineering and Natural Sciences, International University of Sarajevo, Sarajevo, Bosnia and Herzegovina

^f V.V. Popovskyy Department of Infocommunication Engineering, Faculty of Infocommunications, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

^g Université Côte d'Azur, Laboratoire I3S - CNRS, Sophia Antipolis, France

^h Faculty Information Systems and Applied Computer Sciences, University of Bamberg, Bamberg, Germany

ⁱ Department of Electronical Engineering, Telecommunications and Computers, ISEL, Lisbon School of Engineering, Instituto Politécnico de Lisboa, Lisboa, Portugal

^j Department of Physics, Faculty of Arts and Sciences, Adiyaman, Türkiye

^k Department of Computer Science, Division of Theoretical Computer Science, of the KTH Royal Institute of Technology, Sweden

ARTICLE INFO

Dataset link: <https://doi.org/10.5281/zenodo.16571934>

In loving memory of Barbora Buhnová and Ivica Crnković.

Keywords:

Computer science education
Female under-representation
Gender
Ph.D. career path
Ph.D. challenges and opportunities

ABSTRACT

Context: Computer science, even now, attracts a small number of women, and the proportion of women in the field decreases through advancing career stages. Consequently, few women progress to Ph.D. studies in computer science after completing master's studies. Empowering women at this stage in their careers is essential, not just for equality reasons, but to unlock untapped potential for society, industry and academia.

Objective: This paper aims to identify students' career assumptions and information related to Ph.D. studies focused on gender-based differences. We propose a program to inform female master students about Ph.D. studies that explains the process, clarifies misconceptions, and alleviates concerns.

Method: An extensive survey was conducted to identify factors that encourage and discourage students from undertaking Ph.D. studies. The analysis identified statistically significant differences between those who undertook Ph.D. studies and those who did not, as well as statistically significant gender differences. A catalogue of questions to initiate discussions with potential Ph.D. students which allowed them to explore these factors was developed. These were structured into a *Women's Career Lunch* program where students can explore and discuss the benefits of Ph.D. study.

Results: Encouraging factors towards Ph.D. study include interest and confidence in research arising from a research involvement during earlier studies; enthusiasm for and self-confidence in computer science in addition to an interest in an academic career; encouragement from external sources; and a positive perception towards Ph.D. studies which can involve achieving personal goals. Discouraging factors include uncertainty and lack of knowledge of the Ph.D. process, a perception of lower job flexibility, and the requirement for long-term commitment. Gender differences highlighted that female students who pursue a Ph.D. have less confidence in their technical skills than males but a higher preference for interdisciplinary areas. Female students are less inclined than males to perceive the industry as offering better job opportunities and more flexible career paths than academia.

Conclusions: The insights collected from the survey facilitated the development of a questions catalogue structured into the *Women Career Lunch* program to help students make a more informed decision concerning whether they should pursue a Ph.D. in computer science. Localised versions of this program, in 8 languages, were created to support its adoption in different countries and assist in mitigating the female under-representation challenge.

☆ Editor: Laurence Duchien.

* Corresponding author.

E-mail address: mgoul@fct.unl.pt (M. Goulão).

1. Introduction

Computer science attracts less women than men. In addition, the percentage of women decreases with the advancement of their careers. This poses a severe loss of opportunities, as the significance of women in computer science cannot be overemphasised. Diversity catalyses innovation and progress in the technological sector. Women supply distinctive perspectives, experiences, and problem-solving methods to the industry, making it more diverse and all-encompassing. Additionally, as technology continues to shape our world, women in computer science research play an essential role in addressing crucial societal issues and driving groundbreaking advancements.

Across Europe, numerous activities have been launched to counteract this so-called “leaky pipeline” phenomenon. For example, Informatics Europe created the Working Group *Women in Informatics Research and Education (WIRE)* to support, connect, and structure these activities. One activity rooted in WIRE has been the launch of the annual WIRE Workshop, offering a platform for networking and exchange.

The surprisingly large interest in the WIRE Workshop has shown the need for more additional support. This gave the motivation to extend the WIRE Working Group to a larger European effort, resulting in the European COST Action *European Network for Gender Balance in Informatics (EUGAIN)* (<https://eugain.eu>). The objectives of EUGAIN are, beyond those of WIRE, to create a European platform for networking and exchange and to provide guidelines and recommendations for colleagues working at the forefront of the efforts for gender balance in computer science.

Beyond bachelor and master studies, PhDs in computer science cultivate a more robust, well-rounded, and equitable future for technology. Embracing diversity and empowering women in computer science at the doctoral level is not only an issue of equality but also a way to unlock untapped potential and establish a brighter future for the industry and society.

However, with the transition from master’s to PhD studies, we observe a weakening of the female presence. In Europe, an estimated 25.5% of the master’s students were female, while 24.1% of Ph.D. students were female in 2021/2022 (Di Nitto et al., 2023). This gap varies somewhat depending on the country and institution. For example, at RWTH Aachen University in the year 2022, 20% of the master’s students were female, but only 19% of the Ph.D. students. This gap was significantly wider at the Department for Computer Science, Faculty of Mathematics, University of Belgrade, where there were 41% female students at the master’s level but only 14% at the PhD level. A key goal of groups like Informatics Europe, EUGAIN, and others concerned with equality and inclusion is to increase the number of female Ph.D. students and facilitate them in completing their studies successfully. To achieve this goal, we developed a *program to inform female master’s students about PhD studies*. This program has been constructed to be executable at interested computer science departments without intensive preparation, requiring little time and effort and a small financial budget, but still having a substantial impact. For maximal effect, the program is also transferable to other STEM areas.

This paper reports on this work, explaining our starting objectives, methodology to achieve them, and the developed program itself.

The goal of this program is not to convince female students to start a PhD but to help them make an informed decision that best fits their interests and abilities. To achieve this goal, we needed to identify why fewer women than men choose PhD studies. Once we had explored the reasons, we needed an appropriate way to convey relevant information to the female students. We have to counteract wrong assumptions and uncertainties to equip the involved students with the knowledge they need to identify their optimal career preferences.

Thus, we had the following objectives:

- O1 Identify which aspects students consider when they decide on their career path after their master’s studies, with a focus on gender-based differences. Identify areas of missing knowledge, wrong assumptions, and uncertainties which might hinder master’s students from starting a PhD for inappropriate reasons.
- O2 Identify what needs to be conveyed to female students to improve their decision-making.
- O3 Develop a program to convey this information to female students in an engaging and practical way.

To achieve these goals, we used the following methodology:

- M1 To achieve objective O1, we set up an international survey, targeting primarily European countries but open to respondents worldwide. The survey was designed in such a way that the statistical evaluation of answers to the survey should facilitate answering the following research questions:

RQ1: What main supporting factors encourage enrolment in a PhD program?

RQ2: What main blocking factors discourage enrolment in a PhD program?

RQ3: To what extent do these factors differ by gender?

- M2 To achieve objective O2, the answers to the research questions RQ1–RQ3 were used to design a catalogue of questions to prime a discussion with master students about PhD studies, where this discussion will facilitate conveying the relevant knowledge, strengthening supporting factors, and weakening blocking factors, particularly to female students.

- M3 To achieve objective O3, we designed a program named *Women’s Career (WoCa) Lunch*. The program invites female master’s students whose grades are strong enough to consider a PhD to 8 gatherings in small groups. Each module lasts one hour during a common lunch, which gave the program its name. During lunch, the students can chat with (typically female) guests who have completed their PhDs in computer science and work in companies and universities. This format creates a safe space where prospective female Ph.D. students can be informed about PhD studies, using the questions from the catalogue resulting from M2, where each meeting has a specific topical focus.

The rest of the paper describes the above steps in detail. We start with some background in Section 2, followed by a discussion on related work in Section 3. Section 4 describes how we created, distributed, and evaluated the survey. Section 5 presents the survey results, which are then discussed in Section 6. Section 7 describes the question catalogue, which we derived from the survey’s results. Section 8 describes the Women Career Lunch program and its pilot executions. We conclude the paper with some observations in Section 9.

2. Background

Historically, women have been notably underrepresented in STEM disciplines and continue to be outnumbered by men (Hill et al., 2010). Physics, engineering, and computer science exhibit the most pronounced disparities – with only 20% of bachelor’s degrees earned by women in these fields in 2006 (Hill et al., 2010). The percentage of computer science bachelor’s degrees earned by women in the US has even dropped since 1986 (Hill et al., 2010), emphasising that progress towards gender parity should not be taken for granted. Additionally, there is evidence that even if women are majoring in computer science, only 38% of them go on to work in computer-related jobs (Funk and Parker, 2020). This results in women making up only 26% of computing professionals – around the same percentage as almost 60 years ago (Corbett and Hill, 2015).

This is in line with the metaphor of the *leaky pipeline* in STEM and, more specifically, within the domain of computer science. It describes a pattern where women and underrepresented groups start their education and careers but face challenges at each step leading to fewer of them continuing. These challenges, like gender bias and workplace culture (Corallo et al., 2022; André, 2022; Rowe, 2021; Wagner et al., 2020), cause a gradual drop in their participation (Cronin and Roger, 1999). While many female students start in undergraduate courses worldwide, they do not continue to further computer science post-graduate studies (Miliszewska et al., 2006). The leaky pipeline metaphor is a simple and illustrative way to discuss attrition. However, it is crucial to recognise its limitations regarding, e.g., the normative educational pathway it suggests and the exclusion of interdisciplinary fields (Vitores and Gil-Juárez, 2016; Soe and Yakura, 2008).

To quantify the number of female students at each stage, we use the data collected over the past decade by Informatics Europe, available from their data portal (Di Nitto et al., 2023).

Fig. 1 shows a snapshot of the participation of women in computer science degree programs across all European countries from 2010/11 to 2020/21. Firstly, the overall percentage of women at all stages of academic progression remains consistently low, with a maximum of just over a quarter. Despite the significant strides in promoting gender diversity in recent years, there is still a notable gender gap. More notably, the data illustrates a significant difference in the percentages of women enrolled in bachelor, master, and PhD programs versus those awarded the respective degrees. It is encouraging to observe that a larger proportion of women are awarded the degree compared to those who enrol in these programs, indicating that they tend to excel once women enter these programs. However, the critical observation is the disparity in trends between master and PhD programs. While there has been a steady increase in female students enrolling in computer science master programs, the same does not hold for PhD programs. After reaching a peak in 2018/19, the number of female PhD graduates started to decline. Furthermore, the data reveals a stagnation in the percentage of female Ph.D. students, especially from 2015/16 onward. This stagnation is alarming, considering that there is a growing pool of female master graduates who could potentially transition to PhD programs. In other words, there is an untapped resource of talented women who have completed their master degrees but are not proceeding to PhD studies.

Statistics show that the stagnation in the percentage of female Ph.D. students in computer science is not limited to Europe. For example, the Taulbee survey (Bizot and Zweben, 2020) shows that 23 to 25% of enrolled Ph.D. students were women, while a little over 30% of master degrees were awarded to women in the US. It is worth noting that while this stagnation is prevalent in Western countries (Vitores and Gil-Juárez, 2016), it is less commonplace in other places (Varma, 2010; Lagesen, 2008). However, also in Europe, there are certain countries (e.g. Romania) where the average percentage of PhDs awarded to women since 2014/2015 is 43.3%, with an average of 35.5% of enrolled female students, following an average of 39.3% MScs awarded and 35.7% of enrolled female students, albeit with a relatively low ratio of PhDs awarded per million inhabitants (Di Nitto et al., 2023).

In an era where technology plays an ever-increasing role in shaping our world, it is crucial to ensure diversity in the individuals who contribute to the development and advancement of computer science. A key milestone in the academic journey of computer science students is the transition from a masters degree to a PhD program. This transition not only opens doors to advanced research opportunities but also creates future educators, researchers, and industry leaders. However, little is known about the underlying reasons behind the limited transition of women from masters programs in computer science to enrolment in PhD programs. Assessing the reasons behind their choice holds both academic and practical importance. A comprehensive understanding of the multifaceted factors that influence women's decisions in this context not only contributes to the academic discourse but also informs evidence-based strategies aimed at promoting gender diversity in PhD programs.

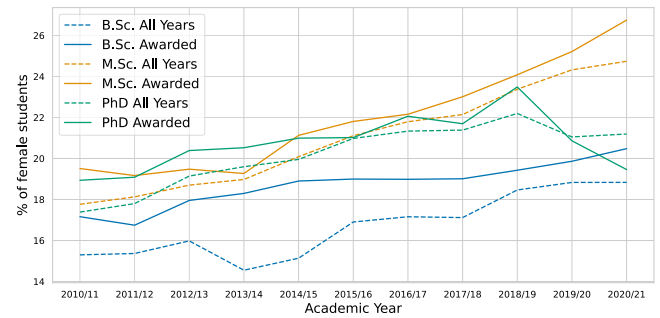


Fig. 1. Percent of female computer science students within Europe (Di Nitto et al., 2023).

3. Related work

Research into the challenges women face in computer science (CS) education (Yates and Plagnol, 2022; Beyer, 2014; Beyer et al., 2003) has investigated the factors influencing their decisions to pursue this field of study (Pantic, 2020; Pantic and Clarke-Midura, 2019; Genut et al., 2019), as well as dropout reasons (Beaubouef and Mason, 2005; Pappas et al., 2016). While these studies have offered strategies to bolster diversity, equity, and inclusivity (DEI) at the undergraduate level (Roberts et al., 2002; Cohoon, 2003; Lagesen, 2007), a distinct research gap remains concerning the transition from master to PhD studies. This transition phase is a critical stage in the academic journey, and understanding its nuances is essential before addressing retention approaches. Understanding the transition phase provides valuable insights and helps address potential misconceptions or barriers from the outset.

While a significant portion of the existing literature on women PhD candidates focuses mainly on retention and attrition dynamics (André, 2022; Alfermann et al., 2021; Wollast et al., 2023), we highlight in Fig. 2 the primary factors and initiatives relating to female enrolment in CS doctoral programmes.

3.1. The main supporting factors encouraging enrolment in a PhD programme

Predominantly, existing studies tend to focus on broader disciplines such as STEM (Diogo et al., 2022), engineering (London et al., 2014; Tarvid, 2014), and computing (Tarvid, 2014; Moreno and Kollanus, 2013) rather than exclusively on computer science. Data from these studies indicate varying levels of female involvement. Some studies have observed noteworthy female representation in their interviews and questionnaires (Diogo et al., 2022; Tarvid, 2014), while others have reported typical levels of participation (London et al., 2014; Moreno and Kollanus, 2013). These findings provide valuable insights into women's perspectives and initial motivations for enrolment. The study by Motogna et al. (2022) is of particular interest, as it explains factors encouraging women to undertake doctoral degrees in computer science, though not confined solely to the enrolment phase.

Self-determination Theory (SDT) and its subsidiary Cognitive Evaluation Theory (CET) suggest that needs for competence, autonomy, and relatedness impact both intrinsic motivations (such as personal fulfilment) and extrinsic motivations including academic career goals, professional development, and career benefits and opportunities (Ryan and Deci, 2000b). SDT suggests that while intrinsic motivation stems from genuine interest, the intensity of extrinsic motivation varies from reward-driven to aligning closely with personal values, often leading to enhanced performance and well-being (Ryan and Deci, 2000b,a).

Women's motivations for embarking on doctoral studies are multifaceted. In pursuit of personal fulfilment, they are often driven by

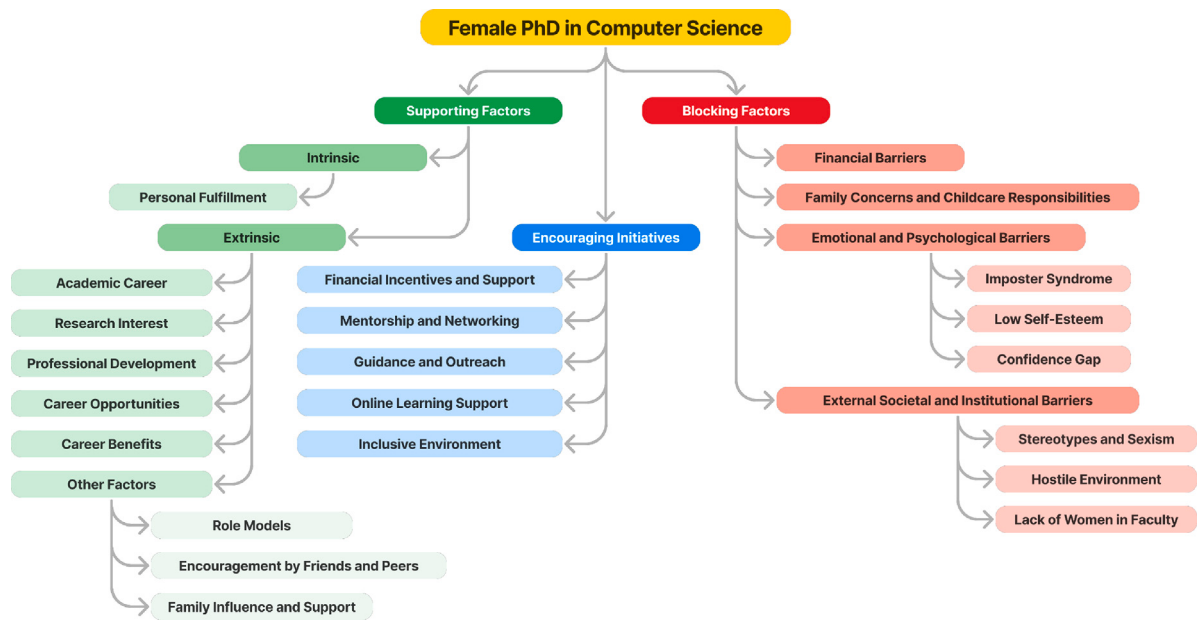


Fig. 2. Primary factors and initiatives related to female enrolment in CS doctoral programmes.

an innate joy of exploration and a deep-seated desire to understand complex subjects. This passion goes hand in hand with aspirations for personal growth and integrating personal interests into one's professional journey via doctoral research (Diogo et al., 2022; London et al., 2014; Tarvid, 2014; Moreno and Kollanus, 2013; Motogna et al., 2022). Alongside these motivations, is an ambition to achieve significant milestones that enhance their career and contribute to global scientific advancements (Diogo et al., 2022; Motogna et al., 2022). Elevating one's social standing and gaining enriching learning experiences also play important roles (Moreno and Kollanus, 2013). Notably, research topic selection is heavily influenced by curiosity and an overarching desire to add value to collective knowledge (Tarvid, 2014; Moreno and Kollanus, 2013).

Students looking for an academic career see a PhD as essential for their future. For some, this commitment to academia has been evident since high school, even when presented with career opportunities in industry. They are interested in teaching and prefer the university environment over corporate settings (Moreno and Kollanus, 2013). For some Ph.D. students the work-life balance of an academic career is compelling, offering part-time roles, flexible schedules, and a conducive work environment (Motogna et al., 2022).

An interest in research is crucial for many Ph.D. students. Doctoral studies deepen students' research capabilities, introduce them to new social networks and refine their writing skills (Tarvid, 2014). Exposure to research, positive experiences, and academic endorsements during master programmes can significantly influence students' decision to advance to a PhD (Moreno and Kollanus, 2013). While the demanding nature of research activities can sometimes diminish enthusiasm (Motogna et al., 2022), these positive factors help sustain the interest in pursuing a PhD (London et al., 2014).

Motivated by professional development, many see a PhD as a strategic investment to enhance career potential, with or without a clear career direction. In contrast, some aim to excel in their chosen fields, viewing doctoral study as a means of refining skills and meeting industry expectations (Moreno and Kollanus, 2013). Pursuing an engineering PhD, in particular, is often associated with the aspiration to align more closely with evolving industry standards and demands (London et al., 2014). Additionally, many advanced roles in scientific research, academia, and industry require a PhD. This advanced qualification equips graduates with deeper knowledge and superior problem-solving skills, setting them apart from peers with lesser qualifications (Diogo

et al., 2022; Tarvid, 2014; Moreno and Kollanus, 2013). For those with significant industrial experience, the PhD is often perceived as a tool for professional exploration, offering the prospect of diverse and potentially more challenging roles within industry (Diogo et al., 2022; London et al., 2014). Moreover, those feeling limited or unfulfilled in their current positions might be attracted to a PhD to further their professional development and open new career pathways (Tarvid, 2014; Moreno and Kollanus, 2013) which can specifically influence female computer science students' decision to undertake doctoral studies (Motogna et al., 2022).

A PhD is a gateway to diverse roles beyond academia, spanning sectors like business, government, and non-profits. Many align their research focus with real-world applications, preparing themselves for entrepreneurial and managerial positions outside universities (Diogo et al., 2022). Doctoral studies offer significant career benefits. In a labour market crowded with bachelor's and master's graduates, a doctoral degree is a distinction that enhances career prospects, improves job security and can promise higher salaries (London et al., 2014; Tarvid, 2014).

Beyond direct career benefits, a significant motivation to pursue a PhD is the influence of successful role models. This motivation becomes stronger when observing the accomplishments of other PhD holders in the labour market, fuelling the desire for similar success and recognition (Tarvid, 2014; Motogna et al., 2022).

In line with CET, encouragement from peers, friends, faculty members, and supervisors is a vital factor driving the decision to embark on a PhD journey (Diogo et al., 2022; London et al., 2014; Tarvid, 2014; Moreno and Kollanus, 2013; Motogna et al., 2022). This choice is usually rooted in familial influences, especially when parents hold degrees. Exposure to academic environments can foster early aspirations, leading some to engage in research during their undergraduate years. Family advice and support further bolster the decision to undertake advanced studies. Furthermore, factors like students' academic achievements, aspirations, and career goals, which their parents' educational background can shape, provide added momentum towards pursuing a PhD (Moreno and Kollanus, 2013; McCallum, 2012; Mullen et al., 2003).

3.2. The main blocking factors discouraging enrolment in a PhD programme

Enrolling on a PhD programme is a significant milestone in educational life, beginning with the decision to pursue an advanced degree.

Yet, female candidates may be reluctant to take this decision due to their gendered experiences. This begins at home, where girls are often encouraged to play with dolls while boys are steered towards video and computer games, facilitating an early familiarity with the technical field (Yates and Plagnol, 2022; Denisco, 2017). Later, these gendered experiences at home and in the school contribute to a reduced pool of women willing and able to enter the CS field at the bachelor and master levels (Yates and Plagnol, 2022; Denisco, 2017). At the PhD enrolment stage, women may face additional limiting factors including financial constraints and emotional and psychological hurdles impacting this decision. Financial barriers include the tendency for females to finance their studies from personal funds, while male candidates are more likely to receive salaried positions (Moskal, 2002). Women also need to provide for their children, among other financial concerns (Motogna et al., 2022). It is often difficult to find affordable daycare, and mothers cannot expect their children to subsist on low-quality food for long periods, usually forcing them to choose between their educational aspirations and raising a family (André, 2022; Rowe, 2021).

Among women in STEM fields, there is a notable presence of impostor syndrome, especially when transitioning into technical areas. Even women actively working in the sector may feel they need additional qualifications to compare favourably with their male colleagues. Such beliefs often stem from deeply ingrained self-doubts, leading them to question their abilities and potential career success (Denisco, 2017; Hyrynsalmi, 2019). From a young age, many women grapple with feelings of low self-esteem (Happe and Buhnova, 2021; Adamecz-Völgyi et al., 2023), making them self-critical and prone to questioning their competence even when they possess the necessary skills. Women tend to doubt their qualifications when faced with challenging or ambiguous job descriptions in technical roles, which can leave them feeling unsuccessful and push them away from their studies or career transitions, even when there is no actual performance deficit (Hyrynsalmi, 2019).

In addition to pre-existing gender biases, societal dynamics further widen the confidence gap among women in technical domains. The observable overconfidence displayed by some men in these areas contrasts starkly with many women's self-doubt (Hyrynsalmi, 2019; Adamecz-Völgyi et al., 2023). In programming courses, for example, men with little to no experience often believe themselves to be equal competitors to women with much higher experience levels (Denisco, 2017). Age biases and competence questions further intensify women's sense of inadequacy. Women often express concerns about being perceived as "too old" for the labour market upon graduation or considering a career change. Some, especially those in their thirties and forties, feel that potential employers might overlook their vast skills and experience in favour of younger candidates (Hyrynsalmi, 2019). Moreover, even in female perception, leadership roles in technical fields are still predominantly associated with males, further contributing to the confidence gap (Hyrynsalmi, 2019; Wang and Redmiles, 2019).

Stereotypes have shaped gender roles, impacting women's experiences in academia and technology. The perception that being a "nerd" or "geek" is unattractive and that computers are mainly for boys negatively impacts young girls' willingness to study computer science (Stross, 2008). From post-secondary education to advanced studies (Cohoon et al., 2009), sexism, real or perceived, also plays a role with boys and men implicitly believed to possess higher competence levels in STEM (Denisco, 2017; Cohoon et al., 2009).

Women in computer science frequently encounter explicit sexism from male peers and professors. From competitions to securing internships at top tech companies, their achievements are often undermined by male counterparts who suggest their success is due to gender-based quotas or a push for workplace diversity. Some male students resort to derogatory comments related to natural female processes (Denisco, 2017). Additionally, male professors might harbour lower expectations for female students, make inappropriate remarks about their personal lives (Denisco, 2017) or ignore questions from female students (Cohoon et al., 2009). Although explicit sexism might be less prevalent in CS

doctoral programmes, its presence remains detrimental, contributing to a higher attrition rate among female doctoral candidates (Cohoon et al., 2009).

Closely related to sexism, hostility in the academic environment can deter women from starting or continuing their doctoral studies. Women tend to perform better in a collaborative and supportive environment, where group success and valued work are more important than individual victories and self-focused learning (Yates and Plagnol, 2022). However, classrooms, especially in CS, often cater to a male-centric perspective (Yates and Plagnol, 2022), making them feel unwelcoming to many women (Cohoon, 2003). Although many institutions have anti-discrimination policies, these measures are often not enough to combat negativity and toxicity in the institutional culture. Such a climate can intensify women's inclinations to abandon their academic pursuits (André, 2022). Dismissive attitudes and a lack of support from supervisors or peers further exacerbate this sentiment (Yates and Plagnol, 2022).

A lack of female faculty members might discourage young women from pursuing advanced careers in male-dominated fields. They might struggle to imagine themselves as potential mentors or professors (Denisco, 2017). Not seeing women at their level and in positions of authority may lead to feelings of isolation, presenting a challenge in terms of motivation due to the need for a more complex integration of social and academic aspects (André, 2022).

3.3. Initiatives to encourage female enrolment in PhD Computer Science programmes

The various initiatives to support female students in Computer Science PhD programmes are multifaceted. Table 1 offers a concise summary of these initiatives, which we have categorised under five groupings (i) financial incentives and support, (ii) mentorship and networking, (iii) guidance and outreach, (iv) online learning support, and (v) inclusive environment.

Financial challenges may discourage female students from enrolling in PhD programmes in various fields, including computer science (Lindner, 2020; Horta et al., 2018). Many **financial incentives and support** such as scholarships, research grants, and, in some cases, stipends have been introduced to cover tuition fees and other costs explicitly dedicated to female students (Lockett, 2023). This financial assistance can increase motivation and significantly improve women's PhD opportunities.

In addition to financial aid, institutions have actively sought to increase female enrolment in PhD programmes through various other campaigns and support programmes (André, 2022; Europe, 2023). Among these, **mentorship programmes and networking opportunities** (Europe, 2023; Anon, 2023) are essential tools to address students' doubts or concerns before starting their PhD journey. By organising events that bring together prospective PhD candidates with current female doctoral students or professors who might become potential mentors, institutions provide a platform where experiences can be shared. Future challenges and opportunities can be openly discussed. Such interactions empower attendees to ask questions, fully understand the PhD process (André, 2022; Fisher et al., 2019), and make informed decisions. Additionally, celebrating female PhD graduates' successes, as done by Stony Brook's Computer Science Department (Department of Computer Science, 2023), offers attendees an opportunity to network and be inspired by the achievements of their peers.

Furthermore, institutions often host **guidance and outreach** initiatives, such as specific PhD orientation activities. For instance, Stanford University's Computer Science Department (Stanford Engineering, 2023) organises events designed to answer students' questions, explain the details of the programme, and highlight post-PhD career opportunities (Anon, 2023). Similarly, Brown University's Computer Science De-

Table 1
Overview of Enrolment Initiatives for Females in PhD Computer Sciences programmes.

Initiative category	Ref	Description	Examples
Financial Incentives and Support	Lockett (2023)	Grants , scholarships, and fellowships for PhD women in CS	Google PhD fellowship programme, Microsoft Research Women's Fellowship
Mentorship and Networking	Europe (2023), Anon (2023)	Networking opportunities with successful peers and role models, mentoring programs	CRA-WP, Women in ML, ACM-W,ACM-W Europe
Guidance and Outreach	André (2022), Europe (2023), at Brown University (2023)	programmes educating potential applicants about the PhD process, by providing feedback and guidance	PhD orientation at Stanford's Computer Science Department, Application Feedback programme for underrepresented applicants at Brown University
Online Learning Support	computerscience.org (2023), Munoz-Hernandez and Osman (2016)	Online doctoral programmes, remote meetings with supervisors	Online or on campus Computer Sciences PhD at the University of South Carolina, College of Engineering and Computing
Inclusive Environment	Rubegni et al. (2023), Jaccheri et al. (2020)	Anti-discrimination policies and support to make the academic environment more welcoming for females	NTNU's gender-sensitive policies, Chalmers University of Technology Genie – Gender Initiative for Excellence

partment has a specialised PhD application programme (at Brown University, 2023) dedicated to underrepresented groups, including women, ensuring a smooth application process. These tailored initiatives are crucial in helping female students reach their academic goals.

Following the 2019 pandemic, there has been an accelerated interest in **online learning supports** that might aid in recruiting female doctoral students. Online doctoral programmes in information technology (computerscience.org, 2023) highlight this increased visibility. Such programme designs offer advantages for women juggling family responsibilities. A study (Munoz-Hernandez and Osman, 2016) presented the benefits of online education for women with limited resources, such as remote meetings with supervisors, suggesting its potential to support greater female participation in PhDs.

European universities have developed projects, policies, and strategies to promote an **inclusive environment** and increase women's participation in CS (Rubegni et al., 2023; Jaccheri et al., 2020). The Norwegian University of Science and Technology (NTNU) stands out for its gender-sensitive policies and anti-discrimination awareness programmes (of Science and Technology, 2023). Beyond policy adjustments, some institutions, like Koç University, have adopted gender-responsive pedagogical practices and incorporated gender-equitable design principles into their teaching materials to cultivate a lasting positive impact (Koc University, 2023).

Retaining female students in computing is as important as recruiting them. Robust retention initiatives ensure that women continue their doctoral journey and position them as role models for potential entrants. Witnessing the successes and perseverance of their peers can inspire and motivate more women to consider and pursue a PhD in the field. Ongoing mentorship and training (in Informatics, 2022, 2023), early exposure to computing topics through workshops and hackathons (de Aguiar Gomes et al., 2018), supportive networks for women at all academic stages (Rubegni et al., 2023), and skills-enhancing training are essential. It bears reiterating that the continued presence of female role models in computing encourages others to follow similar paths.

4. Survey & approach

Our goal was to identify the factors that students consider when they decide on their career path after their master's studies, with a particular focus on gender-based differences. This section describes the instrument design, the data collection approach, the data cleaning performed on the responses received and the demographics of the valid responses received.

4.1. Instrument design

Our survey instrument includes 58 questions (see Table 2) used to collect data on:

- **Demographics (Q1–Q7).** These were aimed at finding out and collecting basic demographic data (enrolment country, nationality, degree, progress, gender).
- **Respondents' background as students (Q8–Q21).** These questions were aimed at profiling our respondents concerning their academic trajectory, providing insights about their experiences and preferences.
- **Intentions concerning the possibility of pursuing a PhD (Q22–Q46).** These questions include not only the intentions, but also respondents' perceptions on the factors that are in favour of, or detrimental to, enrolling in a PhD. Those factors include personal factors (e.g. family related), as well as more structural factors (e.g. related to career prospects, or financial security).
- **Respondents' Self-perception (Q47–Q58).** The last group is mostly focused on respondents' perceptions of their strengths and weaknesses compared to their peers.

We used a mixture of open and closed questions, with a clear predominance for the latter (49 out of 58):

- **Likert.** We have 39 5-point *Likert-scale* questions, where the options are *definitely yes*, *rather yes*, *neutral*, *rather not*, and *not at all*. We also included a *does not apply* option in the Likert-scale questions so respondents would not randomly choose one of the options if for some reason the question did not apply to them. The variables collected via these questions are **ordinal**.
- **SC.** We have 5 *Single Choice* questions that require a *yes* or *no* answer. The variables collected via these questions are **nominal**.
- **SMC.** We have 4 *Simple Multiple Choice* questions. These questions allow the respondent to choose only one of the provided answer alternatives. They include a special *Other* alternative, to allow the respondent to answer in open free text if none of the available alternatives is a good match for the respondent's answer. The variables collected via these questions are **nominal**.
- **MMC.** We have 2 *Multi-select Multiple Choice* questions. In one of them, we added the *Other* option, to provide a free text alternative to our respondents, if necessary. The variables collected via these questions are **nominal**.
- **Open.** We have 8 *Open* questions, devoted to the collection of additional comments or insights the respondents might want to share with us. The variables collected via these questions are **open text**.

Table 2
Survey questions - condensed excerpt.

Part	Id	Question	Type
I. Personal Info	Q1	Please name the country in which the university you are/were enrolled is located.	SMC+Other
	Q2	Please name your nationality.	SMC+Other
	Q3	Please indicate your highest educational rank:	SMC+Other
	Q4	Please name the study program in which you are/were enrolled:	Open
	Q5	Please indicate your current occupation:	MMC+Other
	Q6	What is your gender?	SMC+Other
	Q7	Do you have any further comments?	Open
II. Questions on Own Studies	Q8	Did you have any computer science courses during high school?	SC
	Q9	Are/Have you been personally supervised by a researcher, e.g. by a PhD student or a teaching assistant for a seminar or project work?	SC
	Q10	Are/Have you been involved in research projects during your studies, e.g. in terms of being involved in research at a department/chair or in a larger scientific project?	SC
	Q11	Are/Have you been personally supervised by a research leader, e.g. by a professor or lecturer for a seminar or project work?	SC
		Please rate each of the following statements on the scale displayed.	
	Q12	Do you feel enthusiasm for computer science?	Likert
	Q13	Do you enjoy your studies (without considering the pandemic)?	Likert
	Q14	Would you say that you have good grades?	Likert
	Q15	Would you say that you are successful in your studies?	Likert
	Q16	Do you feel self-confident in interaction with other students?	Likert
	Q17	Did/Do you regularly participate in study groups during your studies?	Likert
	Q18	Do you like theoretical computer science subjects (esp. logic and maths)?	Likert
	Q19	Do you like practical computer science subjects (esp. programming)?	Likert
	Q20	Do you like interdisciplinary areas of computer science?	Likert
	Q21	Do you have any further comments?	Open
III. Questions on PhD (intentions)		Please evaluate each of the following statements concerning your current situation on the scale displayed.	
	Q22	When being a Bachelor student, I plan(ned) to enrol in master studies.	Likert
	Q23	I have thought about doing a PhD.	Likert
	Q24	I plan to do a PhD.	Likert
	Q25	I am aware of the aims and contents of PhD studies.	Likert
	Q26	I am aware of the requirements for starting a PhD.	Likert
	Q27	I have been encouraged to do a PhD.	Likert
	Q28	After completing my current studies, I am considering moving abroad for a while.	Likert
	Q29	My family encourages me to do a PhD.	Likert
	Q30	I was motivated by industrial contacts to do a PhD.	Likert
		Please evaluate each of the following arguments in favour of enrolling in a PhD on the scale displayed.	
	Q31	The aims and contents of PhD studies suit me.	Likert
	Q32	I believe I meet the requirements for getting a PhD position.	Likert
	Q33	A PhD would improve my future career.	Likert
	Q34	Pursuing a PhD would be fun.	Likert
	Q35	Better gender balance in computer science would increase my interest in PhD studies.	Likert
		Please evaluate each of the following arguments against enrolling in a PhD on the scale displayed.	
	Q36	I am worried that PhD studies would negatively impact my future family plan (getting married, having kids, ...).	Likert
	Q37	I have concerns that during my PhD studies, I would not have enough money to support my lifestyle.	Likert
	Q38	I have better job opportunities in the industry than at university.	Likert
	Q39	I am hesitant to start a PhD because of many factors and circumstances (e.g. career opportunities, flexibility, working environment, ...) are unclear to me.	Likert
	Q40	I think the industry offers more flexible career opportunities compared to the structured path of PhD studies.	Likert
	Q41	I am worried about the impact on my future career if I start but do not complete my PhD studies.	Likert
	Q42	I have reservations about making long-term job commitments for the whole PhD duration.	Likert
	Q43	I am afraid of being one of only a few females in a PhD group.	Likert
	Q44	Do/Did you participate in any programs or activities that informed you about PhD studies?	SC
	Q45	Please describe these programs or activities shortly.	Open
	Q46	Do you have any further comments concerning PhD intentions?	Open
IV. Questions on self-reflection		Please evaluate your strengths as compared to your peers:	
	Q47	I have the competencies to succeed in computer science.	Likert
	Q48	I have strong programming skills.	Likert
	Q49	I possess theoretical understanding.	Likert
	Q50	I have self-confidence.	Likert
	Q51	I embrace problem-oriented learning.	Likert
	Q52	I am driven by curiosity to learn.	Likert
	Q53	I am successful in collaborating with others.	Likert
	Q54	I love computer science.	Likert
	Q55	Do you have any insights on what encourages students to start a PhD in computer science?	Open
	Q56	Do you have any insights on what discourages students from starting a PhD in computer science?	Open
	Q57	Do you have any further comments?	Open
	Q58	How did you find out about this survey?	MMC

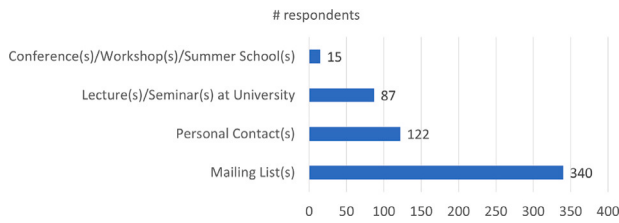


Fig. 3. How participants found about this study.

We conducted a pilot session with colleagues (all academics) to solicit their feedback, during one of our COST Action meetings, before launching the survey. We reviewed the whole questionnaire to detect clarity issues, problematic questions (e.g. overloaded questions), and any other structural problems. We incorporated this feedback. We also asked participants to fill in the questionnaire so we could assess how long it took them to do so and incorporate that information in the questionnaire invitation letter.

4.2. Data collection

The target population for this study was computer science students, researchers, and practitioners in Europe. In the case of participants who already have a PhD, they were encouraged to reflect on their experience. Although this study was open to respondents from countries outside of Europe, we focused our respondents' recruitment efforts on Europe, with all its cultural diversity.

We used an *open invitation strategy* (Wagner et al., 2020). We advertised our survey through several channels, including mailing lists, personal contacts, classes and seminars at Universities, conferences, workshops, and summer schools. We leveraged the contacts of the COST Action members and those of Informatics Europe to reach a broad audience of potential respondents. As a key focus of our research is gender inclusion in Informatics, we were particularly interested in gaining insights from female participants. We specifically targeted potential female respondents in our contacts, as several of our research questions depended on having many female respondents. We also wanted to have answers from respondents of other genders, so we could find out which candidate factors are gender-related.

The responses include answers from people in different life stages, ranging from undergraduate and graduate students considering pursuing a PhD to researchers and practitioners who may have completed a PhD or decided against doing one.

Fig. 3 summarises how our respondents found out about the survey. Several respondents learned about the survey via more than one channel.

We targeted European countries, as this survey was primarily aimed at understanding the phenomenon in the European context. That said, the survey link was public and it did reach people from other continents (e.g. we had several respondents from universities in China). We decided to keep answers from other origins, given the mobility of Ph.D. students. It is common to find non-European Ph.D. students in European Universities.

Concerning sample size, we used Yamane's equation (Yamane, 1973), as suggested in Wagner et al. (2020), to compute a suitable sample size n :

$$n = \frac{N}{1 + N \cdot e^2} \quad (1)$$

With an estimated 28.7 million developers worldwide by 2024 (QBIT, 2023), this would mean around 400 participants, for a precision set to 0.05. This is a conservative estimate, as the number of potential PhD candidates is much lower, especially if we focus on Europe.

$$n = \frac{28,700,000}{1 + 28,700,000 \cdot 0.05^2} = 400 \quad (2)$$

We did not include any localisation for our survey. The same survey version was available for all respondents, regardless of their geographic location.

We did not offer any financial or personalised incentive for participation. The lack of gender balance in Informatics is noticeable and this was enough to attract a significant number of participants.

The data collection process was prolonged in time, due to difficulties in convincing enough people to offer their valuable time for research. We started disseminating the survey in October 2021, mostly via contacts made by the authors, leveraging their networks. By June 2022 we had 129 respondents, mostly from Germany and Portugal. We then increased our efforts to bolster the number of respondents and the diversity of their locations, by encouraging other EUGAIN members to disseminate the questionnaire including Informatics Europe in that process. By February 2023, we had 506 responses. While we were doing some preliminary data analysis, we reiterated our requests for help from EUGAIN members and intensified our dissemination campaign. We finished our data collection on May 30, 2023, with 867 responses from 53 countries.

4.3. Data cleaning and transformation

We filtered out all responses that did not meet our inclusion criteria. In particular, we excluded responses that did not answer our research questions (e.g. respondents only answered basic demographic questions). Of the total 867 responses to our survey, 280 answers contained no relevant information concerning our research questions, leaving us with 587 valid responses.

We sanitised some answers (e.g. we transformed empty answers to the question "What is your gender?" into "I prefer not to disclose". and did the same for a respondent who answered "Other", but then did not fill in the form with any value for "Other" - in practice, the respondent preferred not to disclose the gender, as well).

We computed an extra variable to encode whether the respondent has a PhD, is in the process of studying for a PhD, or plans to do a PhD, or if the person has no PhD and no plans to obtain one. We refer to these values in our data analysis tables and figures as "PhD" vs. "No PhD", respectively.

Finally, we moved free-text responses to a separate file, to preserve anonymity.

4.4. Survey response demographics

Of the 587 valid responses, 309 respondents identified as *female*, 267 identified as *male*, and 4 respondents reported their gender as *other*: 1 *female-to-male transgender* and 3 *non-binary* respondents, one of which also identified as *queer* and *agender*. 7 respondents *did not disclose* their gender. We received responses from people from 53 nationalities. Although our target population were people working in Europe, the questionnaire was available to anyone with the link and publicly announced at an international level, so we also got some answers from other continents (e.g. with 10 participants from India and 8 from China). The countries with the most responses were Serbia (98), Germany (87) and Denmark (86). Fig. 4 summarises respondents' nationalities.

Concerning their current occupation illustrated in Fig. 5, most of our respondents were master (185) and Bachelor (181) students. Some of the respondents hold more than one position (e.g. some of the MSc students also work in industry).

We also asked our participants about their highest educational ranking achieved. We recorded the 7 answers in the *Other* category: 6 of them implied holding a PhD, while the remaining one implied holding an MSc. Overall, most of our respondents declared holding a BSc or an MSc (see Fig. 6).

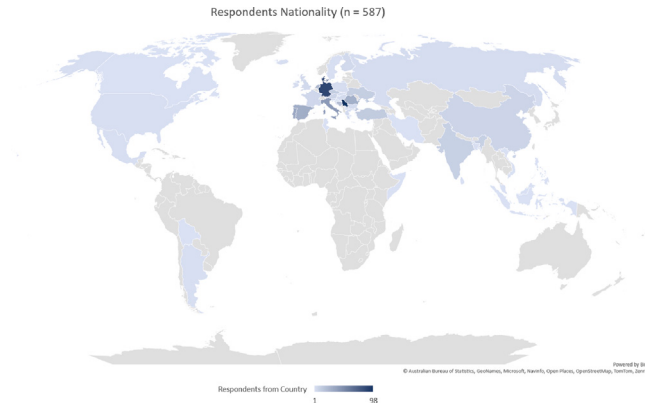


Fig. 4. Respondents nationality.

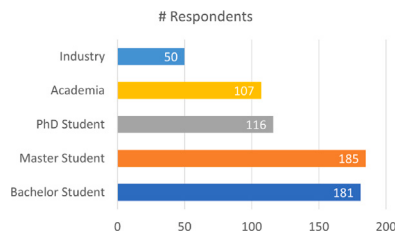


Fig. 5. Respondents current occupation. Some respondents hold more than one position (e.g. Master Student and Industry).

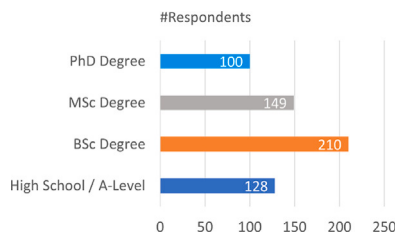


Fig. 6. Respondents highest achieved degree.

5. Analysis and results

We describe the approach used for both the quantitative and qualitative analysis and the results of both types of analysis on each research question as outlined in Section 1.

5.1. Approach used for quantitative analysis

Our questionnaire included several closed questions for which we wanted to perform quantitative analyses, based on the frequencies of the options chosen by our respondents. As discussed in Section 4.1, some of the variables used in our data analysis are nominal, while others are ordinal. We started by collecting descriptive statistics on these variables. We then formulated hypotheses and performed the appropriate hypothesis testing to identify whether differences were statistically significant. Our statistical analysis uses only **nominal** and **ordinal** data. As such, we only use **non-parametric** procedures in our data analysis which **do not require normality**.

We used the Chi-Square test, for single-choice questions (i.e. nominal data). The Chi-square examines the difference between categorical variables from a random sample in order to determine whether the expected and observed results are well-fitting. In particular, we use it to test if there is a difference in proportions between several groups identified by nominal variables. For example, we use the Chi-Square test to contrast the answers from people who decided to do a PhD with those who decided against doing a PhD, concerning whether they were involved in research projects during their studies.

We used the Mann–Whitney U test, for questions where the participant could choose their answer using a Likert-scale value, i.e. ordinal data. The Mann–Whitney U test checks data distribution for a significant difference between two independent groups. The test applies ordinal (which is the case here) or continuous variables. It requires independent observations and the shapes of the distributions of the two groups under comparison to be roughly the same.

Both for Chi-Square tests and for Mann–Whitney U we used $p < .05$ for the level of significance.

As we analyse several independent variables in the context of finding relations with the same dependent variable (whether the participant is “PhD material” — has done, is doing, or plans to do a PhD — or not), it can be argued that this implies a form of multiple comparisons so we also included post-hoc multiple comparison corrections, including the Bonferroni and the Holm–Bonferroni. The Bonferroni correction (Bonferroni, 1936) is a conservative post-hoc method that controls the Type I error rate, but significantly increases the probability of a Type II error (Perneger, 1998). The Bonferroni correction compensates for the risk of making a Type I error by dividing the significance level by the number of hypotheses. For example, for our desired $p < .05$, if we test 5 hypotheses, the corrected value is $\frac{.05}{5} = .01$, so results will be statistically significant using the Bonferroni correction if $p < .01$.

To achieve a more balanced control of Type I and Type II errors, we also tested whether our results are still significant after performing the Holm–Bonferroni method (Holm, 1979). Compared to Bonferroni, the Holm–Bonferroni method increases the Type I error probability and decreases the Type II error probability. The Holm–Bonferroni method sorts the p-values from lowest to highest and compares them to nominal alpha levels of $\frac{\alpha}{m}$, where m represents the number of tested hypotheses. For example, with our desired $p < 0.05$, for 5 hypotheses, the adjusted values are $\frac{.05}{5} = .01$, $\frac{.05}{4} = .0125$, $\frac{.05}{3} = .1(6)$, $\frac{.05}{2} = .025$, and $\frac{.05}{1} = .05$.

For all our hypothesis tests, the significance level is highlighted with different font styles, depending on whether the differences are “just” statistically significant, or remain significant when using a Holm–Bonferroni correction or even when using a Bonferroni correction. and using Bonferroni correction.

We computed the effect size using Cramér’s V (φ_c) (Cramér, 1946) for the Chi-Square tests. With one degree of freedom, the common interpretation thresholds for the effect size computed with φ_c are: .10 — small effect size; .30 — medium effect size; .50 — large effect size. The Mann–Whitney U tests require a different effect sizes measure. We used the eta-square (η^2) effect size measure, which provides the percentage of the variance of the dependent variable (the decision to do a PhD, or not) that can be explained by the independent variable tested in the corresponding Mann–Whitney U test (e.g. the self-confidence in one’s programming skills). We use the following recommended thresholds for interpretation of the effect size values computed with η^2 : .01 — small effect size; .06 — medium effect size; .14 or higher — large effect size.

To make our findings clearer we adopted a visual notation in our tables and text with the following symbols:

Key for the icons used in this section**Participants:**

- Participant who chose to do a PhD
- Participant who chose not to do a PhD
- Female participant
- Male participant

Descriptive statistics:

- # Number of participants
- % Percentage of participants
- $\bar{x}$ mean score
- $\tilde{x}$ median score
- Mode mode score

Participants and descriptive statistics icons can be combined, e.g.:

- # Number of female participants who chose to do a PhD
- % Percentage of male participants who chose not to do a PhD

Statistics tests:

- $\chi^2(1)$ Chi square statistic
- MW-U Mann-Whitney U

Effect sizes:

- φ_c Effect size (for Chi square test)
- η^2 Effect size (for Mann-Whitney U test)

p-values formatting:

- .ddd Not statistically significant
- .ddd Statistically significant
- .ddd Statistically significant with Holm-Bonferroni Correction
- .ddd Statistically significant with Bonferroni correction

Results summary:

- ^ Statistically significant difference, with a low effect size
- ^^ Statistically significant difference, with a medium effect size
- ^^^ Statistically significant difference (higher), with a high effect size
- (similar for lower v, vv, vvv)
- ⊙ Bonferroni-corrected statistically significant difference (higher), with a low effect size
- ⊙⊙ Bonferroni-corrected statistically significant difference (higher), with a medium effect size
- ⊙⊙⊙ Bonferroni-corrected statistically significant difference (higher), with a high effect size
- (similar for lower ⊙, ⊙⊙, ⊙⊙⊙)

The summary of each of the results presented is preceded by a reference group identification.

For example:

- ⊙⊙⊙ Participants who decided to do a PhD had a statistically significantly higher value of the variable under comparison than those who decided not to do a PhD. The difference is significant even when using the Bonferroni correction. The effect size is medium.

We complemented the statistics tests with visualisations (divergent stacked bar charts) to help data interpretation.

5.2. Approach used for qualitative analysis

In addition to the closed questions, there are also several open questions in the questionnaire. It is also relevant to analyse their answers since data from free-text answers to open questions can provide deeper insights into the participants' personal experiences, concerns, or motivations which the other questions in the survey may not have detected. Among 8 open questions: question Q7 is a basic demographics question; questions Q21, Q46 and Q57 are introduced to provide space for participants to write their further comments related to previous questions; question Q45 aims to collect detailed information about the

informative programs and activities the participants attended related to pursuing a PhD (only for the participants who answered “yes” for question Q44). All the answers to these questions were carefully read by the authors and were subsequently excluded from analysis as their contents ended up being irrelevant. Only one comment from question Q46 was reported to highlight an interesting perspective of one participant where participants were asked if they have any further comments concerning PhD intentions in general. We selected the remaining two open questions (Q55 and Q56) for further analysis to gain insights into the perception of the participants on the encouraging and discouraging factors for pursuing PhD studies. These two questions had a much larger number of answers, compared to other open questions, and are the most relevant open questions for this research considering the research questions. These two questions are as follows:

- Q55 — Do you have any insights on what encourages students to start a PhD in computer science?
- Q56 — Do you have any insights on what discourages students from starting a PhD in computer science?

Neither of these open questions was compulsory, so the participants were free to write their comments or not. Out of 587 valid survey responses, 165 and 192 of the participants answered question Q55 and question Q56 respectively, after filtering out answers indicating that the respondent has no response to give. Note that these respondents could be considered as the group who need more information about the advantages or disadvantages of pursuing a PhD. However, for this analysis, they were removed from the data set.

We conducted a qualitative analysis employing thematic analysis (Braun and Clarke, 2006), (Cunningham and Wells, 2017) on these free-text comments. The process involved one researcher reading all the comments and becoming familiar with both data sets: first and second question answers. Considering the recurring keywords and meanings, we identified different themes for two data sets. We assigned answers to these themes, considering their content. One answer can be assigned to multiple themes to cover the whole content of the answer. For example, we assigned the answer “*I suppose being interested in the subject matter would be a start. Or maybe having fun teaching others*”. to the themes “*interest and confidence in research*” and “*interest in academic career and teaching*”. After assigning the themes for the answers, a second researcher checked the coding decisions randomly for 10% of the answers for each data set. We discussed the disagreements and when necessary, made adjustments and recurring issues emerged.

5.3. Analysis of RQ1: What main supporting factors encourage enrolment in a PhD program?

We organised our data analysis for RQ1 around six groups of questions: (i) participants' education experience through single choice questions (questions Q8–Q11), as well as whether the participant had attended any programs or activities that informed participants about PhD studies (Q44, a single choice question); (ii) participants' preferences (questions Q12–Q20); (iii) participants' intentions (questions Q22–Q23; Q25–Q30), (iv) arguments in favour of enrolling on a PhD in Computer Science (questions Q31–Q35); (v) self-reflection questions on participants' perceptions of their strengths (questions Q47–Q54); and (vi) qualitative results from the open questions identifying supporting factors for enrolling on a PhD in Computer Science.

We consider the themes underlying all these questions, as stated in the questionnaire statements, to be plausible candidates for factors encouraging enrolment in a PhD program in Computer Science. That said, an overall disagreement concerning one of these factors may imply that something that should be an incentive (e.g. strong encouragement from the industry to study for a PhD) is, in practice, a detrimental factor (e.g. the lack of strong encouragement from the industry to study for a PhD may hinder the drive for enrolling into a PhD).

(i) *Education experience.* Questions Q8 through Q11, and Q44, assess the participants' previous studies experience and are single choice questions, where respondents answer either Yes or No.

We start by defining our hypotheses. For questions Q8 through Q11, and Q44, these are as follows:

H_{i-0} : There is no statistically significant difference in *factor x*, when comparing participants who choose to do a PhD (👤) and those who do not (👤).

H_{i-1} : There is a statistically significant difference in *factor x*, when comparing participants who choose to do a PhD (👤) and those who do not (👤).

For example, for Q10, the corresponding hypotheses were:

H_{10-0} : There is no statistically significant difference in *having been involved in research projects during the participants' studies* when comparing participants who choose to do a PhD (👤) and those who do not (👤).

H_{10-1} : There is a statistically significant difference in *having been involved in research projects during the participants' studies* when comparing participants who choose to do a PhD (👤) and those who do not (👤).

The data collected with these questions meets the following pre-conditions: (i) the variables are categorical, (ii) all observations are independent, (iii) the cells in the contingency table (see Table 3) are mutually exclusive, and (iv) the expected value of cells should be 5 or greater in at least 80% of cells. We performed a chi-square test of independence to examine the relation between the answer to each of those questions and the choice of our respondents to do a PhD, or not.

Table 3 summarises the results of the Chi-Square test for these 5 questions. The columns represent the question identification, the answer, the number of respondents who decided not to do a PhD (#👤), the percentage of those respondents who provided that answer (%👤), the number (#👤) and percentage (%👤) of participants who decided to do a PhD that provided each answer, the Chi-Square test value ($\chi^2(1)$), the p -value for that test highlighting the statistically significant results ($p < .05$) and their corresponding effect size ϕ_c , a pictorial representation of the statistical significance of any identified differences, their direction and effect size (👤 vs 👤) and total number of respondents. In the column 👤 vs 👤 we represent these differences from the perspective of the PhD respondent (👤). With 5 tests, the Bonferroni-corrected significance level is $p < 0.01$, while the Holm-Bonferroni-corrected significance levels progress from $p < 0.01$ to $p < 0.05$, as explained in Section 5.1. In the case of Q10, the p -value of .000 (the actual value is 1.15624939375694E-16) is significant even with the **Bonferroni correction** (and the less strict **Holm-Bonferroni correction**).

For question Q8, the relation between these variables was not significant. The proportion of subjects who decided to do a PhD did not differ by having any computer science courses during high school (Q8). For questions Q9, Q10, Q11, and Q44 on the other hand, the relation between these variables was statistically significant (Q10 and Q11 even when using the Holm-Bonferroni, or the stricter Bonferroni correction, while Q44 remained statistically significant with the Holm-Bonferroni correction, but not the Bonferroni correction). Participants who were involved in research projects during their undergraduate studies, and/or personally supervised by a research leader were more likely to enrol in a PhD, in both cases with a medium effect size (Q10 and Q11, respectively). This effect is small for participants who were personally supervised by a more junior researcher, such as a Ph.D. student, or a teaching assistant (Q9). Participants who attended programs or activities that informed about PhDs were more likely to

Table 3

How does the student education experience relate to the decision of pursuing (👤) or not (👤) a PhD?

Id	Answer #	#👤	%👤	#👤	%👤	$\chi^2(1)$	p-value	ϕ_c	👤 vs 👤	N
Q8	No	154	41.5	78	38.4	.519	.471	.030		574
	Yes	217	58.5	125	61.6					
Q9	No	164	44.2	72	35.3	4.319	.038	.087	👤^	575
	Yes	207	55.8	132	64.7					
Q10	No	255	68.9	67	33.0	68.683	.000	.346	👤👤👤	573
	Yes	115	31.1	136	67.0					
Q11	No	191	51.8	42	20.7	52.372	.000	.303	👤👤👤	572
	Yes	178	48.2	161	79.3					
Q44	No	298	87.4	147	79.0	6.399	.011	.110	👤^	527
	Yes	43	12.6	39	21.0					

Note: Chi-Square test. The used Bonferroni-corrected significance level is $p < .01$.

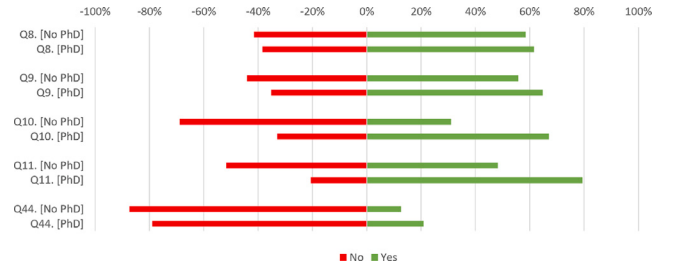


Fig. 7. Supporting factors for pursuing a PhD from single choice questions.

enrol in a PhD, again with a small effect size (Q44). Fig. 7 summarises these results.

Key insights 1: (Research involvement and PhD life information) Participants who were involved in a research project during their undergraduate, or MSc studies, particularly when directly supervised by a senior researcher, were more likely to enrol on a PhD. Participants who attended activities aimed at informing of life as a PhD were also more likely to enrol in a PhD.

(ii) *Participants' preferences and self-confidence.* Questions Q12 through Q20 focus mostly on the preferences expressed by our respondents concerning Computer Science as an area. The overarching rationale behind these questions is that Ph.D. students often express their enthusiasm for the research topic they are working on. As such, we hypothesised participants who choose to do a PhD would like Computer Science topics more than their peers not pursuing a PhD, and feel more confident about their skills. For example, for Q14, the hypotheses would be:

H_{14-0} : There is no statistically significant difference in the *perception of having good grades*, when comparing participants who choose to do a PhD (👤) and those who do not (👤).

H_{14-1} : There is a statistically significant difference in the *perception of having good grades*, when comparing participants who choose to do a PhD (👤) and those who do not (👤).

We created similar hypotheses for the remaining questions. Table 4 summarises the collected data for all these Likert-scale questions. The data collected with these questions meets the following criteria: (i) the variables are ordinal, (ii) the observations are independent, and (iii) the shapes of the distributions are roughly the same. With these pre-conditions met, we used a Mann-Whitney U test to assess these hypotheses. The same set of preconditions for the Mann-Whitney U test were checked throughout the paper for all these tests. In particular,

Table 4

How do participants' enthusiasm and self-confidence relate to the decision of pursuing () or not () a PhD?

Id	p	η^2	 vs 	$\bar{x}$	$\bar{x}(\text{No PhD})$	$\bar{x}(\text{PhD})$	$\bar{x}$	$\bar{x}(\text{No PhD})$	$\bar{x}(\text{PhD})$	Mode	N
Q12	.000	.056	 	4.486	4.704	4.367	5	5	5	5	587
Q13	.000	.043	 	4.255	4.487	4.134	4	5	4	5	577
Q14	.000	.072	 	3.953	4.316	3.766	4	4	4	4	576
Q15	.000	.043	 	4.038	4.289	3.905	4	4	4	4	580
Q16	.001	.021	 	3.762	4.000	3.637	4	4	4	4	580
Q17	.799	.000	 	3.289	3.289	3.290	3	3	3	4	570
Q18	.000	.039	 	3.732	4.030	3.573	4	4	4	5	582
Q19	.895	.000	 	4.227	4.216	4.233	4	4	4	5	582
Q20	.002	.016	 	4.161	4.291	4.090	4	5	4	5	583

Note: Mann-Whitney U test. The used Bonferroni-corrected significance level is $p < .0013157895$.

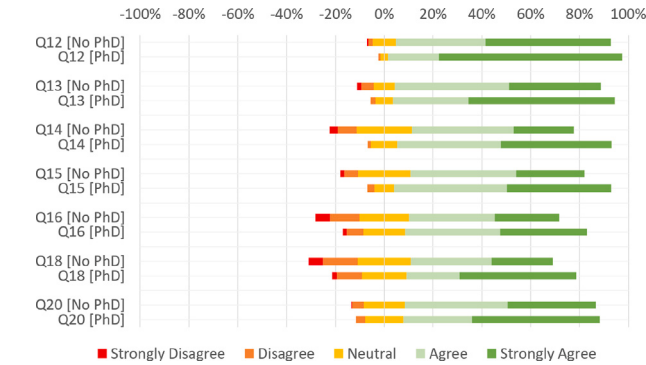


Fig. 8. The distributions of responses on enthusiasm and self-confidence, contrasting participants who decided not to do a PhD [No PhD] () with those who decided to do one [PhD] (). Questions for which there were no statistically significant differences were filtered out.

these preconditions were checked for the tests summarised in Tables 4, 5, 6, 7. While we present these tables separately, we computed the Holm-Bonferroni and Bonferroni corrections to the tests summarised in these tables altogether. These 38 questions lead to a Bonferroni corrected significance level of $p < .0013157895$, while the Holm-Bonferroni corrected significance level ranges from $p < .0013157895$ to $p < 0.05$.

Table 4 presents the results including the question number, the Asymptotic 2-tailed significance level for the Mann-Whitney U tests, the effect size (η^2), a visual representation of the difference between the distributions in participants who decided to do a PhD and those who did not ( vs ), the mean ($\bar{x}$), the mean for those who plan to do a PhD, are doing a PhD, or have done a PhD ($\bar{x}(\text{No PhD})$), and for those who do not ($\bar{x}(\text{PhD})$), median ($\bar{x}$), median for those who plan to do a PhD ($\bar{x}(\text{No PhD})$) and for those who do not ($\bar{x}(\text{PhD})$), mode and the number of cases. Note that, as participants had the opportunity to select the option “does not apply” for any of these questions, the number of valid responses varies from one question to the next.

We found statistically significant differences for most questions, although effect sizes tend to be small. In all cases where we found statistically significant differences, participants who had decided to do a PhD had a higher level of agreement with the factors than those who decided not to do a PhD. Fig. 8 presents these distributions.

The only medium effect size observed in this set of questions was on question Q14. A Mann-Whitney test indicated that the perception of having good grades was greater for participants who decided to perform a PhD than for those who did not, with a medium effect size. There were several other statistically significant differences, all with a small effect size, and all with higher levels of agreement among participants who decided to do a PhD. These include feeling enthusiasm for Computer Science (Q12), enjoying one's studies (Q13), feeling successful in those

Table 5

How do participants' intentions and other context variables relate to the decision of pursuing () or not () a PhD?

Id	p	η^2	 vs 	$\bar{x}$	$\bar{x}(\text{No PhD})$	$\bar{x}(\text{PhD})$	$\bar{x}$	$\bar{x}(\text{No PhD})$	$\bar{x}(\text{PhD})$	Mode	N
Q22	.234	.003	 	4.366	4.404	4.545	5	5	5	5	530
Q23	.000	.057	 	3.543	3.995	3.301	4	5	4	5	529
Q25	.000	.065	 	3.417	3.849	3.177	4	4	3	3	518
Q26	.000	.100	 	3.340	3.930	3.018	4	4	3	4	523
Q27	.000	.179	 	3.050	3.877	2.591	3	4	3	1	524
Q28	.544	.001	 	3.052	3.000	3.080	3	3	3	3	498
Q29	.000	.101	 	2.834	3.431	2.502	3	4	3	3	506
Q30	.016	.012	 	1.726	1.900	1.628	1	1	1	1	503

Note: Mann-Whitney U test. The used Bonferroni-corrected significance level is $p < .0013157895$.

studies (Q15), feeling self-confident in one's interactions with other students (Q16), and liking theoretical computer science (Q20) and interdisciplinary areas of computer science (Q20). In contrast, We did not find statistically significant differences in questions Q17 (whether participants were involved in study groups) and Q19 (whether participants enjoyed practical computer science subjects, such as programming).

Key insights 2: (Enthusiasm and self-confidence) Participants who decided to do a PhD in Computer Science feel more enthusiasm for Computer Science in general, and theoretical computer science and interdisciplinary studies in particular. They also enjoy their studies more than other participants. In addition, they express a higher self-confidence in the success of their studies, and their interaction with their peers.

(iii) *Participants' intentions and context.* Questions Q22, Q23 and Q25 through Q30 focused on the participant's current situation, particularly concerning the participants' views on the possibility of pursuing a PhD. We followed a similar data analysis approach to the one above (Section 5.3(ii)). Our hypotheses follow the same structure as previously.

Most of these questions had statistically significant results for the Mann-Whitney U test we performed, with effect sizes ranging from small to large, suggesting differences between participants who decided to do a PhD and those who did not. Only the responses to Q22 and Q28 were not statistically significant. Table 5 summarises these results.

Participants who decided to do a PhD had a higher perception of having been encouraged to do so than the others, and this difference had a large effect size (Q27). A notable feature of these answers is that the mode in the group of participants that did not decide to do a PhD is 1, the lowest possible value on this scale. We observed a similar relation concerning participants' perception of having been encouraged to do a PhD by their family, albeit with a medium effect size (Q29). The awareness of requirements for starting a PhD, as well as of the aims and contents of PhD studies, were also higher for those who decided to start a PhD than for the other participants, with medium effect sizes (Q26 and Q25, respectively). Questions Q23 (having thought about doing a PhD) and Q30 (having been encouraged by industrial contacts to start a PhD) also had statistically significant results (the latter only when considered in isolation), albeit with small effect sizes. A distinguishing factor about Q30 is that it had the overall lowest agreement scores in the whole questionnaire. Fig. 9 illustrates this distribution.

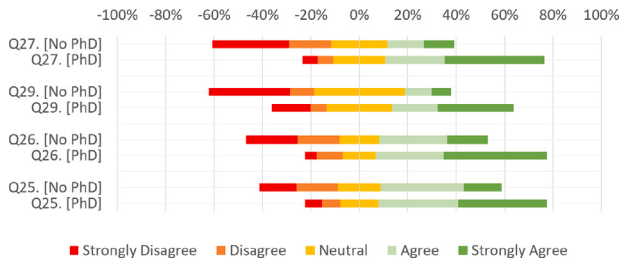


Fig. 9. The distribution of responses to questions related to participants' intent and context factors for those questions with statistically significant differences, with medium effect sizes, from those who chose to do a PhD and those who did not.

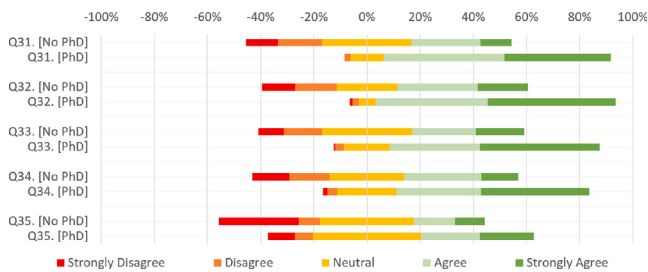


Fig. 10. The distribution of responses to questions related to participants' perceptions of supporting factors in favour of PhD studies from those who choose to do a PhD and those who didn't.

Key insights 3: (Encouragement and requirements) Participants encouraged to do a PhD were much more likely to enrol in one. Having such an incentive from one's family, or by an industrial contact is also relevant, but the latter is very uncommon. Our data also shows how many participants were unaware of the requirements for starting a PhD, and how being aware of such requirements can make a difference.

(iv) *Arguments in favour of enrolling in a PhD.* Questions Q31 through Q35 aim to assess our participants' perceptions concerning supporting factors that might encourage students to enrol in a PhD program. We followed the same analysis strategy as in the previous questions set. The hypotheses associated with each of these questions are defined similarly.

Our test results are summarised in Table 6. Participants who decided to do a PhD perceive the aims and contents suit them better (Q31), are more convinced they meet the requirements (Q32) and have a higher perception that a PhD will improve their future career (Q33) than those who decided not to start a PhD. All these differences have a large effect size.

Participants who decided to do a PhD are more likely to perceive it as fun than those who did not with a medium effect size (Q34). Finally, participants who decided to do a PhD generally agree that better gender balance in Computer Science would increase their intentions of doing a PhD more than those who did not, with a small effect size. Fig. 10 illustrates these differences.

Table 6

How do arguments in favour of starting a PhD relate to the decision of pursuing () or not () a PhD?

Id	<i>p</i> -value	η^2	 vs 	$\bar{x}$	$\bar{x}$ ()	$\bar{x}$ ()	$\bar{x}$ ()	$\bar{x}$ ()	Mode	N	
Q31	<u>.000</u>	.234	   	3.509	4.232	3.085	4	4	3	4	501
Q32	<u>.000</u>	.181	   	3.663	4.341	3.272	4	4	3	4	505
Q33	<u>.000</u>	.146	   	3.599	4.196	3.267	4	4	3	3	514
Q34	<u>.000</u>	.130	   	3.469	4.059	3.139	4	4	3	4	516
Q35	<u>.000</u>	.053	   	2.930	3.360	2.697	3	3	3	3	516

Note: Mann-Whitney U test. The used Bonferroni-corrected significance level is $p < .0013157895$.

Table 7

How do participant's plans and other context variables relate to the decision of pursuing () or not () a PhD?

Id	p-value	η^2	vs	$\bar{x}$	$\bar{x}$ (👤)	$\bar{x}$ (👤)	$\bar{x}$ (👤)	$\bar{x}$ (👤)	Mode	N
Q47	.152	.004		4.249	4.322	4.209	4	4	4	522
Q48	.080	.006		3.594	3.679	3.548	4	4	4	527
Q49	<u>.000</u>	.024		4.030	4.212	3.933	4	4	4	527
Q50	.602	.001		3.640	3.701	3.607	4	4	4	525
Q51	.468	.001		4.249	4.271	4.237	4	4	4	518
Q52	<u>.000</u>	.032		4.378	4.582	4.269	5	5	4	526
Q53	<i>.019</i>	.011		4.088	4.202	4.026	4	4	4	525
Q54	.068	.006		4.224	4.326	4.68	4	4	5	523

Note: Mann-Whitney U test. The used Bonferroni-corrected significance level is $p < .0013157895$.

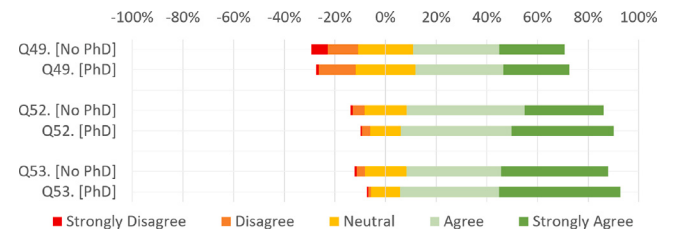


Fig. 11. The distribution of responses to questions related to participants' reflection on their strengths from those who chose to do a PhD and those who did not.

Key insights 4: (Perception of the PhD process) Participants who chose to do a PhD were more inclined to perceive it as something that would fit their needs, improve their future careers, and be fun to do. They also felt more prepared to do so and are more inclined to think that more gender balance would increase their intentions of doing a PhD.

(v) *Self-reflection on own strengths.* Questions Q47 through Q54 assess the participants' reflections on their strengths when compared to their peers. We follow the same analysis approach adopted in the previous sections. Table 7 summarises the results of the Mann-Whitney U tests run to test our hypotheses.

Participants who decided to do a PhD feel more driven by curiosity to learn (Q52), to have a better understanding of theoretical computer science (Q49), and more successful in collaborating with others than their peers who decided not to do a PhD (Q53). The results from Q53 are only statistically different when considering this question in isolation, but not when applying the Holm-Bonferroni, or the Bonferroni corrections. For the remaining questions, we found no evidence supporting an association with the decision to enrol in a PhD. Fig. 11 highlights the differences in these distributions.

Table 8

Themes for Q55 — Do you have any insights on what encourages students to start a PhD in computer science?

Themes	Explanation	Sample comment(s)	Nr. of comments
1	Interest and confidence in research	Strong interest in a particular research topic or field of study was mentioned by many respondents as a motivating factor for pursuing a PhD. This category also includes curiosity, passion for learning, and desire to solve complex problems.	104
2	Personal goals	Factors such as independence, expertise, recognition, dignity and personal development were mentioned as motivations for pursuing a PhD. Some more detailed personal reasons such as living abroad are also categorised under this theme.	30
3	Interest in academic career and teaching	Becoming a university professor or having a career in academia, and enjoying teaching mentioned to be some of the motivations for some individuals to pursue a PhD.	30
4	Career benefits and opportunities	Some respondents mention that a PhD can have lots of benefits for their career, providing more independence and variety in their choices and potentially leading to better job opportunities.	28
5	Encouragement from mentors or colleagues	Some respondents mention that encouragement from mentors or their colleagues would play a significant role in the decision to pursue a PhD.	28
6	Lack of interest in industry jobs	Some respondents mentioned that the fast-paced and competitive nature of industry work would influence their decision to pursue a PhD.	17
7	Seeing role models	Some respondents state that seeing or working with successful role models during their studies is a big motivation for them on their decisions to pursue a PhD.	14

Key insights 5: (Personal strengths) Albeit with small effect size, a greater curiosity to learn, a better understanding of theoretical computer science, and the ability to collaborate with peers are perceived as more developed in people who have decided to do a PhD. We found no significant differences in the competencies to succeed in Computer Science, having strong programming skills, embracing problem-oriented learning, or loving computer science, as these are similarly likely to be found in participants who decided not to do a PhD.

(vi) *Arguments for starting a PhD (qualitative analysis).* The open question Q55 investigates the encouraging factors for pursuing a PhD. Table 8 presents the themes determined from the answers to this question, their explanations, a sample comment selected among responses assigned to that theme, and the total number of responses assigned to that theme. Note that, the unique survey participant ID is also reported in the sample comment column. Fig. 12 illustrates this data in a bar chart to better visualise the distribution of the answers on this theme. Since a single response may have been assigned to more than one theme, the percentages are calculated by dividing the number of responses assigned for that theme by the total number of responses for that question.

The most popular theme that was mentioned in 55% of the responses is “Interest and confidence in research” as shown in Fig. 12. In many of these responses, this theme was mentioned as excitement, curiosity, and passion about learning and discovering new things and having a special interest in a particular research topic. Twelve respondents highlighted the positive impact of participating in research projects during bachelor or master studies in increasing interest and confidence in research. Some of these comments are:

- P320 - “Being part of real research projects in their courses and projects play a big role I think”.
- P682 - “... A research project during bachelor or master studies could show students what it looks like and feels like to be a researcher”.
- P835 - “... Educate them about what research and PhD are and that might help them to consider enrolling. This can be by including them in

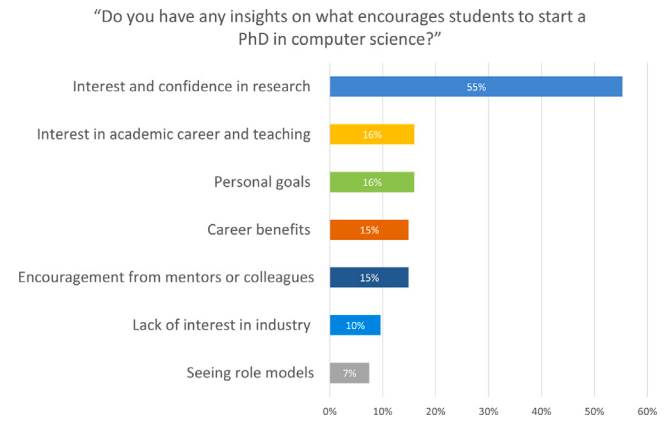


Fig. 12. Distribution of answers over themes identified in a qualitative analysis of the insights on encouraging factors for starting a PhD.

research projects, papers, providing research methodology as a course for them”.

For some themes, the gender distributions of the respondents are noticeably unbalanced. For example, the theme “Seeing role models” was mentioned by only 2 male respondents, whereas all the remaining answers were given by female respondents or the respondents who identified themselves as “other” and also the ones who did not want to disclose their gender. Similarly, 71% of the respondents mentioned the theme “Encouragement from professor or colleagues”, and 78% of the respondents of the theme “Lack of interest in the industry” were females or other genders different than males.

The responses to this open question provide valuable insights into increasing the enrolment of students pursuing their PhDs. For the encouraging factors, motivating and supporting bachelor and master students with activities that help them to feel excited and confident about doing research may have the most significant impact on encouraging students to pursue a PhD. It is also essential to show them the career benefits of having a PhD and its potential to help in personal and professional development.

Table 9

How do arguments against starting a PhD relate to the decision of pursuing () or not () a PhD?

Id	p-value	η^2	 vs 	$\bar{x}$	$\bar{x}$ ()	$\bar{x}$ ()	$\bar{x}$ ()	$\bar{x}$ ()	Mode	N
Q36	.612	.001	 vs 	2.606	2.628	2.594	3 3	2	1	510
Q37	.076	.006		2.930	2.790	3.006	3 3	3	4	512
Q38	.063	.007		3.947	3.820	4.018	4 4	4	5	511
Q39	.000	.075		3.120	2.663	3.365	3 3	3	3	493
Q40	.000	.052		3.637	3.262	3.847	4 3	4	4	509
Q41	.786	.000	 vs 	3.000	3.022	2.987	3 3	4	4	497
Q42	.014	.013		2.926	2.744	3.034	3 3	3	3	473
Q43	.626	.001		2.241	2.118	2.311	2 2	2	1	257

Note: Mann–Whitney U test. The used Bonferroni-corrected significance level is $p < .0013157895$

5.4. Analysis of RQ2: What main blocking factors discourage enrolment in a PhD program?

For the analysis of RQ2, we included (i) the quantitative analysis of a set of closed questions (Q36 through Q43) that implied candidate arguments against starting a PhD in Computer Science, and (ii) the qualitative analysis of one open question (Q56) for participants to identify arguments against doing a PhD.

(i) *Arguments against doing a PhD (quantitative analysis).* We performed a similar analysis to that done in RQ1 but with a focus on participants' perceptions of blocking factors that can hinder the likelihood of prospective PhD candidates doing a PhD. The structure of the hypotheses remains the same as for RQ1. For instance, the hypothesis associated with Q39 is as follows:

H_{39-0} : There is no statistically significant difference in the participant being hesitant to start a PhD due to perceiving some factors and circumstances are unclear when comparing participants who choose to do a PhD () and those who do not ().

H_{39-1} : There is a statistically significant difference in the participant being hesitant to start a PhD due to perceiving some factors and circumstances are unclear when comparing participants who choose to do a PhD () and those who do not ().

We performed Mann–Whitney U tests on the distribution of responses to each of these questions given by those who chose a PhD and those who did not (see Table 9).

Participants who decided to do a PhD feel there are less unclear factors and circumstances associated with a PhD (Q39), with a medium effect size. They also have a lower agreement with the perception that there are more flexible career opportunities in the industry (Q40), and are less afraid of making long-term job commitments for the whole duration of their PhD studies (Q42) than their peers who decided not to do a PhD, both with a small effect size. The results from Q42 are only statistically significant when considering this question in isolation, but not when applying the Holm–Bonferroni, or the Bonferroni corrections. We did not find evidence supporting an association with the decision to pursue a PhD for the remaining questions in this subset. Fig. 13 illustrates this distribution.

Key insights 6: (Uncertainty) Uncertainty about what it is like to do a PhD was the strongest discouraging factor for doing it, closely followed by a perception of lower job flexibility by participants who decided against starting a PhD. The long-term job commitment implied by starting a PhD may be a deterrent factor for some.

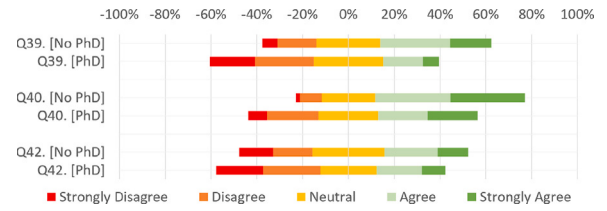


Fig. 13. The distribution of responses to questions related to participants' perceptions of discouraging factors for pursuing a PhD with statistically significant differences between those who chose to do a PhD and those who did not.

(ii) *Arguments against starting a PhD (qualitative analysis).* The open question Q56, seeks the discouraging factors for pursuing a PhD and has more responses than the open question about encouraging factors. There is also more variety in the factors mentioned for this question. Nine different themes were identified to categorise these responses. Table 10 presents these themes and Fig. 14 shows the distribution of the responses across these themes.

More than half of the participants (Fig. 14) who answered this question mentioned the uncertainty in the benefits of pursuing a PhD in their future careers and the low salaries during and after PhD studies compared to alternative jobs in industry. This is followed by the answers stating the negative perception of PhD studies. Most comments under this theme mention the long study duration and seeing this time as a “waste of time” for this fast-paced sector. The third most mentioned theme is the preference for practical experience where most respondents, instead of stating their excitement for working in the industry, mentioned their financial concerns about working in academia and the benefits of industry positions. 65% of the respondents who mentioned their preference for practical experience also mentioned the higher salaries in industry compared to academia.

An interesting observation for this question is that the theme “Perceived sexism and hostile environment” was mentioned by only the female respondents or the respondents who identify themselves as “other” gender or who do not want to disclose their gender.

There is a comment worth mentioning regarding sexism in academia being a discouraging factor or not. Participant P47 stated their ideas in the open question Q46 as follows:

- P47 - “I don't think someone with a master's in computer science would be averse to getting a PhD just because of gender imbalance. They already have a master's, they've experienced it before”.

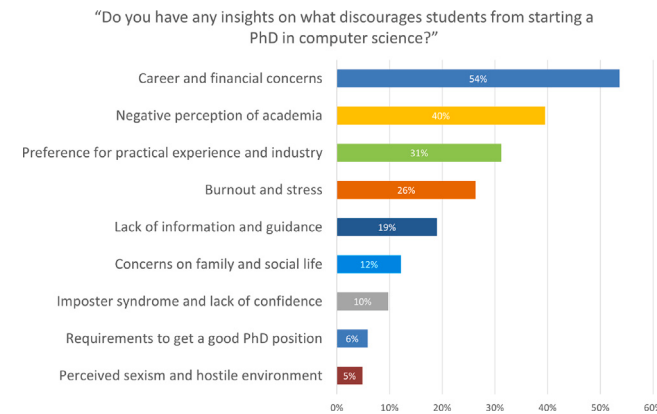
This argument may be one of the reasons for the low ratio of answers in the “Perceived sexism and hostile environment” theme.

These themes and specific comments can help us better understand the factors discouraging students from continuing their studies at the PhD level and work on clarifying or providing more information to the potential Ph.D. students to minimise the impact of these discouraging factors. According to the conclusions of this qualitative study, it is essential to clarify the career concerns and doubts of the students. Many respondents mentioned the uncertain job opportunities after finishing their PhD. These concerns can be decreased by providing more information to these students on the wide range of options and providing advice and mentorship on how to reach them. Meeting with successful role models would also help these students change their negative perception of pursuing a PhD and may help them find answers to their doubts about PhD life, family and social life balance, and the work environment.

Table 10

Themes for Q56 — Do you have any insights on what discourages students from starting a PhD in computer science?

Themes	Explanation	Sample Comment(s)	Nr. of comments
1	Career and financial concerns	Many respondents feel that pursuing a PhD may delay their entry into the job market and result in lower salaries compared to industry positions. They mentioned the attractive and better job opportunities and higher income available in the industry.	110
2	Negative perception of academia and research	Many respondents state a negative perception of academia, considering it slow, theoretical, and disconnected from the real world. They also express doubts about the value and relevance of research conducted during a PhD. Long study duration and lack of funding during the PhD studies are also categorised under this theme.	81
3	Preference for practical experience	Some respondents mention their preference for gaining practical experience in the industry rather than pursuing a PhD, as they believe they can learn and develop skills faster in a professional environment.	64
4	Burnout and stress	The demanding nature of a PhD, including long working hours, stressful deadlines, and high workloads can lead to burnout and anxiety. Respondents stated the high possibility of feeling overwhelmed by the pressure and struggling to maintain a healthy work-life balance.	54
5	Lack of information and guidance	Respondents mention a lack of clear information and guidance about the PhD process, including the application process, expectations, and career paths. They may feel uncertain about what they would do after completing their PhD and how it would impact their job opportunities outside of academia.	39
6	Concerns on family and social life	Some of the respondents mention the difficulties of starting a family or spending free time on weekends without being worried about the PhD work.	25
7	Imposter syndrome and lack of confidence	Some respondents mention Imposter syndrome and low self-confidence as factors that discourage them from pursuing a PhD. They may feel intimidated by the academic environment, especially if they are the only female or if they perceive a lack of respect or support.	20
8	Requirements to get a good PhD position	Some respondents mentioned that to obtain a good PhD position and grant, one should have good grades and “excellence” in their previous studies.	12
9	Perceived sexism and hostile environment	A few respondents expressed concerns about facing sexism and unequal treatment in academia. They may have witnessed or experienced discriminatory comments or disrespectful behaviour, which can discourage them from pursuing a PhD.	10

**Fig. 14.** Distribution of answers over themes identified in a qualitative analysis of the insights on discouraging factors for starting a PhD.

5.5. Analysis of RQ3: To what extent do these factors differ by gender?

Looking into the overall statistics of factors favouring or discouraging the enrolment in a Computer Science PhD can obfuscate gender-related differences within our sample of participants. On the other hand, simply contrasting the answers from participants of different genders without considering whether or not they decided to start a PhD can also obfuscate key differences, as said differences may manifest only in particular subgroups. For instance, females who choose not to do a PhD may have significant differences from males who choose not to do a PhD, while those differences are not present when considering females and males who choose to do a PhD. As such, in this section, we consider four groups: females who choose not to do a PhD (♀), males who choose not to do a PhD (♂), females who decide to do a PhD (♀), and males who decide to do a PhD (♂). We then compare females vs males who decide not to do a PhD, as well as females vs males who decide to do a PhD, to identify gender differences when the decision to do a PhD, or not, is considered. We define hypotheses

Table 11

Supporting factors for a PhD, segregated by doing a PhD (♂), or not (♀). The Mann-Whitney U test compares the results of females and males in each subset of participants (♀ vs ♂, and ♀ vs ♂, respectively). For brevity, we only present questions with statistically significant differences in at least one of the groups (♀, or ♂). We sort results in descending order of the maximum effect size for each pair of questions. The differences for all remaining questions concerning supporting factors for starting a PhD were not statistically significant, suggesting no relevant gender differences. The used Bonferroni-corrected significance level is $p < .0006756757$.

Id	♀ vs ♂	p	η^2	♀ vs ♂	$\bar{x}$	$\bar{x}(\text{♀})$	$\bar{x}(\text{♂})$	$\bar{x}(\text{♀})$	$\bar{x}(\text{♂})$	Mode	N
Q48	♂	.000	.093	♀	3.548	3.249	3.885	4	3	4	334
Q48	♀	.000	.088	♀	3.676	3.344	4.047	4	3	4	182
Q19	♂	.000	.055	♀	4.236	4.021	4.474	4	4	5	369
Q19	♀	.062	.017	♀	4.218	4.109	4.348	4	4	5	202
Q50	♂	.000	.053	♀	3.620	3.364	3.910	4	4	4	332
Q50	♀	.145	.012	♀	3.709	3.604	3.910	4	4	4	182
Q20	♂	.245	.004	♀	4.101	4.149	4.046	4	4	4	368
Q20	♀	.000	.067	♀	4.289	4.464	4.076	5	5	4	204
Q16	♂	.000	.034	♀	3.655	3.469	3.863	4	4	4	371
Q16	♀	.415	.003	♀	4.010	3.954	4.078	4	4	4	198
Q35	♂	.002	.031	♀	2.667	2.891	2.420	3	3	3	315
Q35	♀	.002	.053	♀	3.354	3.611	3.050	3	3	3	175
Q12	♂	.003	.024	♀	4.368	4.274	4.474	5	4	5	372
Q12	♀	.526	.002	♀	4.324	4.723	4.685	4	5	5	182
Q54	♂	.526	.002	♀	4.324	4.051	4.297	4	4	5	182
Q54	♀	.008	.021	♀	4.167	4.292	4.360	4	4	4/5	330
Q47	♂	.017	.017	♀	4.221	4.121	4.333	4	4	4	330
Q47	♀	.010	.037	♀	4.320	4.198	4.459	4	4	5	181
Q52	♂	.010	.020	♀	4.270	4.170	4.382	4	4	5	333
Q52	♀	.077	.017	♀	4.577	4.646	4.500	5	5	5	182
Q29	♂	.212	.005	♀	2.516	2.613	2.412	3	3	3	316
Q29	♀	.015	.033	♀	3.439	3.360	3.181	4	4	3	180
Q15	♂	.035	.012	♀	3.914	3.832	4.001	4	4	4	370
Q15	♀	.295	.006	♀	4.292	4.236	4.360	4	4	4	199
Q18	♂	.040	.011	♀	3.568	3.687	3.434	4	4	4	370
Q18	♀	.950	.000	♀	4.030	4.036	4.022	4	4	5	201

similar to those in the previous research questions, but now with these more refined criteria. For instance, for Q48, the hypotheses are:

$H_{48-0-N_{\text{PhD}}}$: There is no statistically significant difference in the perception of having strong programming skills when comparing females (♀) and males (♂) who decide not to do a PhD.

$H_{48-1-N_{\text{PhD}}}$: There is a statistically significant difference in the perception of having strong programming skills when comparing females (♀) and males (♂) who decide not to do a PhD.

$H_{48-0-P_{\text{PhD}}}$: There is no statistically significant difference in the perception of having strong programming skills when comparing females (♀) and males (♂) who decide to do a PhD.

$H_{48-1-P_{\text{PhD}}}$: There is a statistically significant difference in the perception of having strong programming skills when comparing females (♀) and males (♂) who decide to do a PhD.

Gender differences in supporting factors for pursuing a PhD. Here, we consider the same survey questions as in RQ1. This time we segregate responses using the gender of the respondents. We compare the answers provided by females with those provided by males. Of those questions analysed with the Chi-Square test, only one of them (Q44) had a statistically significant difference ($p = .039$), without considering a

Holm-Bonferroni or a Bonferroni correction, when contrasting females and males who did not do a PhD, with a small effect size ($\varphi_c = .112$). A slightly higher percentage of females than males had participated in a program or activities informing them about PhD studies. Table 11 summarises the results for the supporting factors in which there were statistically significant differences. We present in more detail those for which this difference was significant even when considering the Holm-Bonferroni and the Bonferroni correction, as these are also the ones with larger effect sizes, and briefly mention the remaining ones. The comparison between females and males (♀ vs ♂) is always presented from the female (♀) perspective.

While the factors addressed by the questions in Table 11 are intrinsically encouraging factors, a lower level of agreement expressed by a particular subgroup of participants with an encouraging factor suggests that this encouraging factor is less prevalent for that particular subgroup. For instance, female participants who choose not to start a PhD have a worse opinion of their programming skills than males with a medium effect size (Q48 ♂). The same happens for those who choose to do a PhD, with females' opinions being worse than males' and a medium effect size (Q48 ♀). Female participants who choose not to start a PhD like practical Computer Science subjects, such as programming less than males with a small effect size (Q19 ♂). This difference is not significant when considering participants who choose to do a PhD (Q19 ♀). While there is no difference between participants of both genders who choose not to start a PhD (Q20 ♂), for those who decide to start it, female participants like interdisciplinary areas of Computer Science more than male participants, with a medium effect size (Q20 ♀). Female participants who choose not to start a PhD have lower self-confidence than males, with a small effect size (Q50 ♂). Again, this difference is not significant for those who decide to start a PhD (Q50 ♀). Female participants who choose not to start a PhD have a lower self-confidence in their interaction with other students than male participants, with a small effect size (Q16 ♂). The difference is not significant for those who decide to start a PhD (Q16 ♀). (see Fig. 15).

All the statistically significant differences in the previous factors remain significant even when using the Holm-Bonferroni, or the Bonferroni corrections. In addition, we also found several statistically significant differences, when considered in isolation, all with a small effect size. Some of these differences led to higher levels of agreement from females than males.

Regardless of their decision, females were more likely to consider that better gender balance would increase their possibilities of starting a PhD than males (Q35 ♂). Females who decided to do a PhD considered family encouragement as a more relevant factor than males (Q29 ♀). This difference is not significant when considering participants who decided not to do a PhD.

Females who decide not to start a PhD have nevertheless a stronger preference for theoretical Computer Science than males (Q18 ♂). Still, this difference is no longer statistically significant when considering participants who decided to do a PhD.

In contrast, females who decided not to start a PhD are less enthusiastic than males about computer science (Q12 ♂). They also love computer science less than males (Q54 ♂) and feel less successful in their studies (Q15 ♂). None of these differences are statistically significant for those who decided to start a PhD. Regardless of their decision about starting a PhD, females express lower confidence in their competencies to succeed in computer science than males (Q47 ♂).

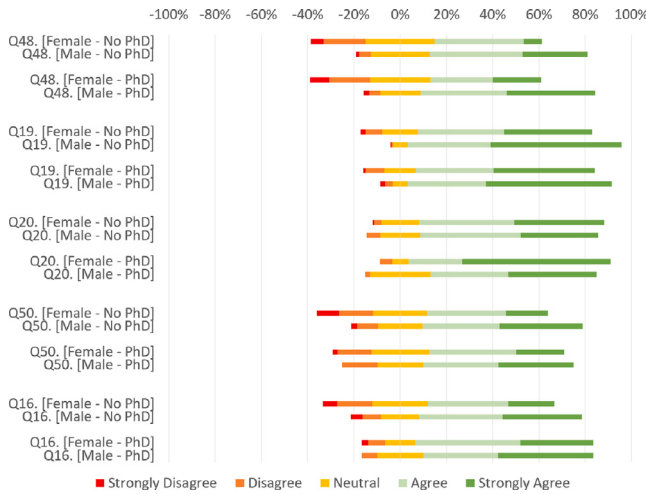


Fig. 15. The distributions of encouraging factors for pursuing a PhD grouped by gender, filtered by cases where at least one of the groups has statistically significant differences concerning another group by gender (e.g. Q48. has statistically significant differences between male and female participants who did not pursue a PhD, and another one, between male and female participants who decided to enrol in a PhD).

Key insights 7: (Enthusiasm and self-confidence (continued)) Two recurring themes in our findings are a consistently slightly lower enthusiasm for computer science and a slightly lower self-confidence in their computer science abilities expressed by females who choose not to start a PhD compared to males. These differences disappear when considering participants who decided to start a PhD, suggesting love for and self-confidence in doing computer science are less prevalent in females than in males who decide not to start a PhD. This lack of self-confidence when considering programming skills is stronger, even for females who decide to start a PhD. On the other hand, females who decide to start a PhD expressed a higher preference for interdisciplinary areas of computer science and valued family encouragement for starting a PhD more than males. Regardless of their decision about starting a PhD, females have a stronger belief than males that a better gender balance would increase their likelihood of starting a PhD in computer science.

Gender differences in factors against pursuing a PhD. Here, we consider the same survey questions as in RQ2. This time, we segregate responses using the gender of the respondents. We compare the answers provided by females with those provided by males. Table 12 summarises the results for the discouraging factors in which there were statistically significant differences when considered in isolation with a small effect size. None of these differences are significant when considering the Holmes-Bonferroni or the Bonferroni correction. The comparison between females and males (♀ vs ♂) is always presented from the female (♀) perspective. Females who decided not to do a PhD expressed lower agreement with the idea that the industry offers better job opportunities than the University than males (Q38 ♀). This was also the case for those who decided to do a PhD, with females also expressing a lower agreement than males (Q38 ♂). Females who decided not to do a PhD also expressed a lower agreement than males with the idea of industry offering more flexible career paths (Q40 ♀). This was also the case for those who decided to start a PhD, with females having a lower agreement than males with the perception of industry offering more flexible career opportunities, (Q40 ♂).

Fig. 16 illustrates the distribution of responses for discouraging factors with statistically significant differences across males and females.

Table 12

Detrimental factors for doing a PhD, segregated by doing a PhD, or not. The Mann-Whitney U test compares the results of females and males in each subset of participants (♀ vs ♂, and ♀ vs ♂, respectively). For the sake of brevity, we only present questions with statistically significant differences in at least one of the groups (♀, or ♂). We sort results in descending order of the maximum effect size for each pair of questions. The differences for all remaining questions concerning supporting factors for starting a PhD were not statistically significant, suggesting no relevant gender differences in their results. The used Bonferroni-corrected significance level is $p < .0006756757$.

Id	♀ vs ♂	p	η^2	♀ vs ♂	$\bar{x}(\text{♀})$	$\bar{x}(\text{♂})$	$\bar{x}(\text{♀})$	$\bar{x}(\text{♂})$	Mode	N
Q38 ♀	♀	.002	.029	♀	4.041	3.899	4.199	4 4	5	319
Q38 ♂	♀	.050	.021	♀	3.818	3.684	3.965	4 4	4	181
Q40 ♀	♀	.015	.019	♀	3.868	3.750	3.993	4 4	4	318
Q40 ♂	♀	.023	.029	♀	3.260	3.063	3.477	3 3	4	181

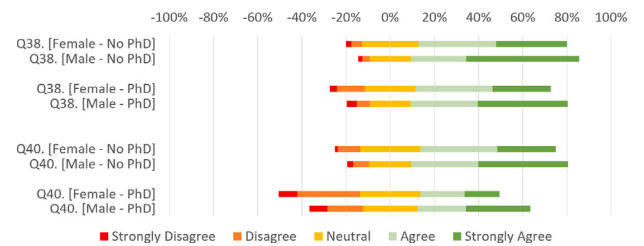


Fig. 16. The distribution of responses for questions about discouraging factors for pursuing a PhD grouped by gender. Only the questions with statistically significant differences are included.

Key insights 8: (Industry offers) Females expressed a slightly lower agreement than males concerning the perception that the industry offers better job opportunities and more flexibility than academia.

6. Discussion

This section discusses the insights obtained from the data and how they may inform specific recommendations. We then discuss the limitations of our study and threats to validity in the context of sampling and response bias, construct, measurement, conclusion and internal validity. This section finishes with some implications of this research and some lessons learned.

6.1. Insights

We start with the insights obtained from the data about encouraging factors. Then, we address discouraging factors and discuss the gender-based differences found in our data. We suggest some recommendations for practice as strategies to mitigate noticed issues.

Encouraging factors. Involving students in research projects and having them supervised by senior researchers is a good way of making them aware of the possibility of enrolling in a PhD in computer science. Several participants also suggested involving those students in writing and submitting research papers as a beneficial strategy to encourage them to do a PhD. Indeed, the most cited reason for pursuing a PhD in computer science was an interest and confidence in the ability to conduct research work in this area, which significantly outnumbers any other reason expressed by participants in this way. Self-confidence and interest differences between female and male students are often pointed out as key factors for the underrepresentation of women in computer science (Beyer, 2014).

The second most mentioned motives were personal goals such as independence, expertise, recognition, or personal development. Several of these goals are commonly listed as desirable outcomes of a

PhD (Lovitts, 2005; Waaijer, 2017). The willingness to pursue an academic career received similar attention from our participants.

This research interest can be complemented with an awareness program designed to provide information to potential PhD candidates. Such programs are not common, but they can raise the awareness of potential Ph.D. students about the career opportunities a PhD may enable, as well as the inherent challenges in doing a PhD and how to navigate through them. In this paper, we present one such program, in Section 8.

Furthermore, encouraging students with a high potential to do a PhD can make a difference. This encouragement often comes from academic staff and peers, but it can also come from the student's family or industry, in line with the findings of Yates and Plagnol (Yates and Plagnol, 2022), albeit it is the male family members or industry contacts. Unfortunately, the latter is relatively uncommon. While this is problematic, it implies a clear growth opportunity, as well, by strengthening the synergies between the industry and academia, possibly through the proposal of more applied research topics.

Unsurprisingly, people who decide to start a PhD in Computer Science tend to be more enthusiastic about it and more self-confident in their abilities than their peers. In particular, students who prefer theoretical computer science or interdisciplinary computer science topics tend to pursue a PhD more often than their peers. PhD candidates tend to perceive a PhD as something that can fulfil their needs (including a greater curiosity than their peers), improve their future career options, and, ultimately, be fun. They also feel a better gender balance would increase their chances of starting a PhD.

Discouraging factors. The key factors against doing a PhD seem to be the uncertainty of what it is like to do a PhD, and a perception of lower job flexibility when compared to what industry offers. One of the most cited factors against doing a PhD is a negative perception of academia and the [lack of] value of doing a PhD for the student in the long run. These perceptions are far from exclusive to computer science students but may be, at least to some extent, counterbalanced by personal goals (Waaijer, 2017). The program proposed in this paper can help mitigate these concerns, as obtaining a PhD may open additional opportunities for those who do it to find more intellectually appealing jobs (Waaijer, 2017).

The long-term commitment implied by starting a PhD is also perceived as problematic for many. This seems to be tied to some of the feedback we got from our qualitative data. Indeed, the most cited discouragement for starting a PhD is a combination of career and financial concerns, with a PhD viewed as delaying an increased quality of life improvement by at least four years. The generally low value of PhD scholarships, when compared to salaries offered in the industry significantly harms the option of doing a PhD in Computer Science, at least during the years the student is working on the PhD. Apart from structural changes leading to more competitive scholarships to bridge the salary gap, another strategy would be to provide evidence on the salary evolution for Computer Science professionals holding a PhD (not only in Academia, but particularly for those working in the Industry), as a way to support the argument against the perception that the return on the investment in doing a PhD is insufficient. For instance, in the US, professionals earn a substantially higher salary with a PhD in computer science rather than only a bachelor's or master's degree (Steele, 2023).

Differences between females and males. When contrasting the differences between responses from female and male participants, concerning encouraging factors, a noticeable feature is that, in several cases, these differences are statistically significant when one segregates participants between those who decided to do a PhD and those who did not. This was the case in 10 out of the 13 questions for which we found statistically significant differences. In 7 of those 10 questions, the differences exist for those who choose not to do a PhD, which suggests these questions provide insights on existing biases against starting a PhD. For instance, several answers highlight the “imposter syndrome”

as more prevalent in female participants who decide not to do a PhD, although, as in other disciplines, men are also affected by it (Bravata et al., 2020).

In contrast, we also find that factors like having multidisciplinary computer science subjects appeal more to females. This is in line with the observation that hybrid computer science courses including modules from other disciplines are more attractive to female undergraduate students (Alvarado et al., 2012). Family support is also more relevant for female students, which seems consistent, for instance with the findings in Webber and Dismore (2021).

Another factor we also noticed is a lower self-perception of programming skills (possibly another symptom of the already mentioned imposter syndrome), in line with a lower love for computer science in general (but not theoretical computer science in particular) expressed by female participants. This is consistent with findings of lower confidence in technological abilities and lower motivation (Yates and Plagnol, 2022).

Recommendations for practice. Considering the observations on encouraging and discouraging factors and how they vary when considering gender differences, we offer a few practical, albeit potentially challenging recommendations:

- Promote programs clarifying the process of doing a PhD, as a way of raising awareness among students about the opportunities they may have and the underlying challenges of doing a PhD. Our results suggest that many participants who decided not to do a PhD were significantly less aware of this information than those who decided to do one. This is not surprising, but it suggests that we may be missing out on good PhD candidates who never considered that possibility, among other factors, because they might not have been so aware of that possibility or what life looks like for a Ph.D. student. While this is problematic for all genders, our data suggests it is more common problem among females. This gender gap aligns with the generally lower risk-taking nature of females compared with males. As the unknowns resulting from lack of information entail risks, providing valuable information to potential candidates mitigates some risks and, presumably, leads to better-informed decisions about starting a PhD. This entails a double advantage: more students may be aware and consider doing a PhD, and fewer students will enrol in a PhD “by mistake”, potentially leading to a more effective PhD program. We propose a “Women Career Lunch” program that embodies this recommendation (see Section 8).
- Systematically address the imposter syndrome by recurrently providing evidence against it. Our data suggests the gender gap associated with this syndrome is noticeably mitigated among those who choose to start a PhD, although not eliminated (e.g. concerning the participants' perceptions of their programming skills). If we are to bolster the applications of females for PhD studies, this is one of the factors we need to address. The official data from Informatics Europe provides no evidence of a generally lower performance of female students in their studies. Indeed, overall, females tend to represent a higher percentage of graduates than the percentage of females who enrol in a course, suggesting their success rate is higher than the one achieved by males at the bachelor, master, and PhD levels.
- Promote multi-disciplinary topics for PhDs. As females show considerable interest in multi-disciplinary research, there seems to be a yet-to-explore potential for growth. To be clear, we do not advocate promoting these topics *instead* of more “traditional” computer science topics, but rather *in addition* to those topics.

6.2. Limitations and threats to validity

We discuss limitations and validity threats to this survey, and our mitigation strategies to handle them.

6.2.1. Sampling bias

Having a random sample of people taking or already having a degree in Computer Science is impractical, as there are no publicly available contact lists covering our target population. This challenge is common to Software Engineering studies (Amir and Ralph, 2018; Baltes and Ralph, 2022). We used a combination of convenience and snowball sampling to mitigate this risk. We leveraged the contacts of the EUGAIN COST action and Informatics Europe to reach potential respondents well beyond the contacts available to the authors of this article. Naturally, convenience and snowball sampling may introduce biases in the obtained sample. We tried to mitigate this bias by using a broad and diverse channels to reach potential respondents. The diversity of origins of our respondents suggests that, at least to some extent, we succeeded in this strategy. That said, the percentage of respondents from the different countries is not aligned with the actual percentage of potential respondents in the population, with countries like Germany or Serbia being over-represented compared to other countries. This may introduce cultural-based biases in our sample. Also, the female gender is over-represented in our sample compared to the estimated population. In this case, we were particularly interested in having a large set of responses from females, to help us answer RQ3, so we oversampled females by targeting our survey advertisement efforts more to potential female respondents. Finally, our sample is Europe-centric by design although we did not exclude respondents from other origins, as it is not uncommon for students and academics to have mobility among continents.

6.2.2. Response bias

The questionnaire was anonymous and freely available through an online link, so we cannot completely rule out the possibility of someone answering the survey more than once. We cannot also rule out the possibility of having a higher response rate from participants who were particularly interested in our research, as perceived from the invitation. There is also a potential cultural effect resulting from the responses' geographical distribution. While our analysis did not consider geography as a direct factor countries from which we got more answers may have their participants' cultural values oversampled. All that stated, we are contrasting people who decided to do a PhD with people who decided against it (RQ1 and RQ2) and females with males (RQ3), using normalised data (e.g. the percentage of people in each cohort who provided a particular answer), the effect of the different number of members in each cohort was mitigated. Also, we had a relatively high number of respondents (587), well over our target sample size (400), so these threats were mitigated.

6.2.3. Construct validity

When analysing the responses to our questionnaires, we emphasised that the answers reflect the perceptions of participants rather than facts. For instance, their perception of their programming skills may not match their actual programming skills. However, we should note that we were trying to establish relationships between participants' perceptions and their decision to start a PhD, which matches what happens in the decision process, at least on the side of the candidates. Of course, deciding to do a PhD and successfully going through the PhD candidates' selection process are two different things. Our focus was on the former. Also, we need to consider that many respondents who reported that they had decided to do a PhD have already completed it or are currently enrolled in a PhD program.

6.2.4. Measurement validity

Most of the questions in our survey were measured with a balanced 5-point Likert scale, presented in the survey as a set of equally-distanced options. While there is some debate on the soundness of computing, for instance, the mean in the distribution of these variables as a summary variable, the balanced, equally distanced scale points,

along with many responses, mitigate that risk. That said, we complement this information with the median, and our hypotheses tests are always done using non-parametric procedures, such as the Mann–Whitney U test, which are adequate for dealing with ordinal data. Also, we always gave respondents the possibility of not answering a particular question to mitigate the risk of randomly selected answers.

6.2.5. Conclusion validity

When testing our hypotheses, we always highlight the extent to which differences are statistically significant, either in isolation when using the Holm–Bonferroni correction or the stricter Bonferroni correction. We also present the computed effect sizes for our hypotheses tests so that we mitigate the risk of overstating the effect of the identified relations.

6.2.6. Internal validity

The internal validity concerns the causal relation between a treatment and its expected results. To do so, we would have to demonstrate correlation, precedence, and the absence of other variables that might be causing identified correlations. Our statistical tests show correlations but not precedence. That said, we specifically asked participants to answer the questionnaire from their perceptions when they were deciding whether to start (or not) a PhD, which preceded the decision itself. We cannot completely discard the possibility of having external variables explaining the found relationships (e.g. we have no direct information on cultural background, including religion or ethnicity, which may affect this kind of decision. We decided not to use nationalities as proxies for this kind of analysis, as we would have to rely on stereotyped visions of the cultural background of people from those countries, possibly introducing another validity threat to our data analysis.

6.3. Implications for researchers and leaders and future work

This work provides new insights into the factors that may be related (either as encouraging or as detrimental) to the decision to start a PhD in Computer Science. We further contrast those factors by segregating answers by gender to get insights into how these factors may differ depending on gender. The study results have allowed us to develop a career lunch program. The program mitigates some of the challenges that were particularly more prevalent in female respondents. Our results also illustrate how this is a multi-factor decision, with several different factors, most with a small to moderate effect size. While our focus in this survey was on the enrolment in a PhD, the Career Lunch program introduced in Section 8 is expected to have a double effect: not only will it help potential candidates to make a more informed decision concerning starting a PhD, but also, because the decision is better informed, should contribute to help to improve the retention rates in PhD programs, vs. assuming people who do decide to start a PhD know better what to expect.

Creating and adapting the Career Lunch program to several languages, with the necessary tunings reflecting the diverse cultures associated with those languages, creates ample opportunities for leaders to adopt this program in their Universities. The next crucial step is to run this program in different instances and continuously improve it to leverage the lessons learned with each new instance of the program. Along with the lessons learned, we will also need to monitor its adoption and its impact on recruiting new Ph.D. students. This monitoring requires the adoption of key process indicators, so the impact of adopting such a program can be assessed via a longitudinal study.

6.4. Lessons learned

The survey design, implementation, and analysis were a joint effort from a large group of researchers from different countries and cultural backgrounds. This has undoubtedly helped with the location-specific advertisement of the survey (e.g. invitation emails were translated into local languages and adapted to target recipients). Also, we leveraged the members of the EUGAIN COST Action and the access to Informatics Europe members to disseminate the questionnaire, leading to a large, diverse sample of participants. Using a research-focused survey tool (in our case, LimeSurvey) helped us understand our participants. For instance, the tool records partial answers as well as complete ones, so we were able to accurately compute the response rates for each question and filter out questionnaire participants whose responses were incomplete.

7. Question catalogue

Following the survey analysis, which focused on general and gender differences between those who progress to PhD studies and those who do not, we constructed a catalogue of questions that will prime a discussion with female students to facilitate conveying the relevant information to them about PhD studies. The questions addressed topics identified as significant from our analysis, which were categorised under the following three categories: (i) basic information and practical issues, (ii) career opportunities, and (iii) emotional support.

7.1. Basic information and practical issues

The category *Basic information and practical issues* includes questions that can start a discussion to provide basic information to the students about PhD studies, the pragmatic issues concerning applying for a PhD position, and the everyday life of a Ph.D. student. It includes the following questions:

Basic information What are the objectives of a PhD? How long does a PhD take? What are the elements of PhD research? What does the supervision of my work look like? How to choose a topic for a PhD? How is the quality of PhD work assessed? What distinguishes a PhD abroad from a PhD in my country?

Requirements What are the requirements for PhD admission? Do the criteria abroad differ from those in my country? What are the requirements for finishing a PhD? What do you do during studies to qualify for a PhD position?

Finding and choosing position, topic, group, supervisor Where do I find PhD positions? How do I find PhD positions abroad? Which institute/department suits me? What topic suits me? Can I choose the topic of my PhD? Can I get to know the group before starting my PhD? Can I ask doctoral candidates abroad for reports on their experiences? How do I like the atmosphere? How big is the group? How do I find good supervision for my PhD? How do I choose a supervisor?

Application and relevant information How do I apply for PhD positions? What do I have to consider when applying for a doctorate abroad? What does the application process look like? What is essential in the written application? What does the PhD application interview look like? Where can I get advice if I want to go abroad? What do I have to think about administratively?

Day-to-day life and expectations What funding possibilities are available? How do I finance my stay abroad? What is the typical day-to-day of a Ph.D. student? What would my future tasks and possibilities be? Do I have to speak the local language if I go abroad?

7.2. Career opportunities

The category *Career opportunities* includes questions that can evoke discussion on career opportunities in general and specifically in academia and industry. This includes the following questions:

Career in general How does the PhD topic determine your career? How important is the topic of a PhD for further career opportunities? Is the salary higher with a PhD than with a master's degree? What skills will be developed during PhD study concerning research, teaching, and leadership? Is the doctorate obtained abroad recognised in my country (and other countries)?

Academic Career What does an academic career look like? How does the PhD prepare you for an academic career? What is essential to progress to the next step in an academic career? Who are your partners throughout an academic career path? With whom will you work together?

Industry Career What can a career in the industry look like? Is there the possibility of doing your PhD in industry? Which skills are acquired during a PhD that are relevant to the industry? What benefits does having a PhD give towards a career in the industry? Can PhD research lead to a Spin-out company? Do domestic and foreign companies consider the merits of having a PhD differently?

7.3. Emotional support

The category *Emotional support* includes prompting questions that address personal doubts, imposter syndrome, lack of self-confidence, gender balance issues, burnout, and sexual harassment. This includes the following questions:

Imposter syndrome and doubts Am I capable of pursuing a PhD? Am I smart enough? What do I do if I am unsure or have problems during my doctorate studies? Will I have enough innovative ideas? What if I am not able to implement my research ideas? Can I change or adjust my research focus during the PhD? What are typical questions and concerns students may have about their doctorate studies?

Burnout What do I do if I “run out of breath” during the doctorate? Is it just me?

Family planning Is doing a PhD compatible with family planning?

Gender issues Am I okay with being one of the few women in the group? What can I do if I receive sexist remarks? Where can I find help for sexual harassment?

Finding support Who can I contact if I have problems? Where can I find advice and support? How are the dignity and respect rights of the Ph.D. student applied in your University? What do I do if I am dissatisfied with my doctoral supervision?

7.4. Survey insights and question catalogue

Table 13 illustrates the connection of each category of questions with the insights and themes identified in the survey. The first row of this table associates each insight identified in Section 5 with the appropriate question catalogue category, illustrating how the questions focus on the areas of insights discovered by the statistical analysis of the survey data. The second and third rows of the table associate the themes from the quantitative analysis identified in Tables 8 and 10 under the relevant question catalogue category, showing how these themes are covered in the questions.

Table 13

Question catalogue and its connection to the insights and themes from the survey.

	Basic information and practical issues	Career opportunities	Emotional support
Key insights	(PhD life information) (encouragement and requirements) (perception of the PhD process) (uncertainty)	(uncertainty) (industry offers)	& (enthusiasm and self-confidence)
Encouraging factors (Table 8)	id 2 (personal goals)	id 3 (interest in academic career and teaching) id 4 (career benefits and opportunities) id 6 (lack of interest in industry jobs)	
Discouraging factors (Table 10)	id 5 (lack of information and guidance) id 8 (requirements to get a good PhD position)	id 1 (career and financial concerns) id 2 (negative perception of academia and research) id 3 (preference for practical experience) id 5 (lack of information and guidance)	id 4 (burnout and stress) id 6 (concerns on family and social life) id 7 (imposter syndrome and lack of confidence) id 9 (perceived sexism and hostile environment)

Note that the insight (*personal strengths*) is not addressed within the catalogue as it refers to the internal characteristics of people who choose to do a PhD. Such characteristics are formed throughout the complete education process and cannot be influenced quickly. This holds for the encouraging factor 1 (*interest and confidence in research*), which is also not included in the table. Encouraging factor themes 5 (*encouragement from mentors or colleagues*) and 7 (*seeing role models*) also are not covered directly by specific questions in the catalogue but are likely to be addressed indirectly through the process of discussion that is prompted by other questions. We address these themes explicitly in the program we propose in the next section. We include all other insights and themes in the proposed question catalogue to prompt discussion in these significant areas.

8. Women career lunch

The *Women Career Lunch (WoCa Lunch)* program is a supporting initiative for female students that aims to acquaint female undergraduate and master's students with the possibilities for advancing their careers if they pursue PhD studies. The program aims to deliver the information identified as significant by the survey to potential female Ph.D. students in a structured, engaging, and collaborative way.

The program was developed by dividing the identified question catalogue (Section 7) into structured modules. Each module represents a cohesive unit suitable for a one-hour discussion on a topic triggered by the questions in the module. Also, this program includes two vital encouraging factor themes that were not explicitly handled in the question catalogue: encouragement from mentors or colleagues and seeing role models (key insight, and encouraging factor themes 5 and 7 from Table 8). These aspects can be provided by the the program facilitators and guests from academia or industry who share their personal experiences. Participation should encourage students and enable them to make well-informed decisions about whether they should enrol in a PhD program.

We prepared a multi-language version of the WoCa program to foster dissemination and assist in implementation. It includes guidance on the structure and format, the practical procedures of the initiative and a comprehensive catalogue of questions, that were refined from the original list, that should be discussed within each module. This catalogue is publicly available (Abraham et al., 2023). It is available in the following languages: English, French, German, Portuguese, Serbian, Spanish, Turkish, and Ukrainian.

WoCa Lunch was successfully implemented at the Department of Computer Science and the Department of Mathematics at RWTH Aachen University, Germany. The following descriptions and recommendations are based on the obtained experiences, and lessons learned including participants' feedback.

Overview. The program has eight modules. The initial and the last modules should be in person. The rest can be virtual. Each module takes place during lunchtime and lasts approximately one hour. One guest is invited and interviewed for each module (except for the first and last modules). The interview guests might be members of the executing institution or external guests. The goal is to address the items from the question catalogue in a familiar environment and further give insights into the career development of female role models.

Organising group. The team organising the program has ideally two moderators and admin support. The moderators should have a PhD and be academic staff members. At least one moderator should be a female, and at least one moderator should be a senior researcher.

Invited guests. Each module has a guest appropriate for the module and at least one moderator. If possible, each guest should be female, and at least one should be an early-stage researcher, while at most one guest should be local to the department/institute running the event. Guests should be suitable for role models (enthusiastic, motivating, communicative, open). Guests should speak a local language, as the modules should be run in the local language. At least two weeks before their module, a guest should receive a catalogue of questions they will be asked. Knowing questions in advance enables guests to prepare well for their module.

Structure. The first module presents the basic ideas of the program and what the students can expect from it. The final module wraps up the program as a whole and gathers feedback. The remaining modules address the most important topics related to PhD studies. Modules aim to initiate discussions and reflections, dispel uncertainties, and clarify the career path for those pursuing this path. Depending on the group, modules can be merged, or new modules can be introduced.

The modules discuss questions identified in Section 7. The questions are grouped around the following topics:

(Module 0) Kick off This initial session introduces the program and what the students can expect.

(Module 1) What is a PhD? This module clarifies the objectives of a PhD, the requirements for PhD admission, and the skills that will be developed during PhD studies. Students get basic ideas about the main elements of PhD research and the requirements for finishing a PhD dissertation. Students also learn about the typical day-to-day life of a Ph.D. student.

(Module 2) Why a PhD for an Academic Career? This module clarifies the challenges of an academic career and how PhD studies prepare students for that career. Students learn about the process of selecting a research topic, research group, and supervisor for their PhD studies, and the impact of this choice on their future careers. Students should understand the evaluation processes of the quality of their work and the duration of their doctoral studies.

(Module 3) Why a PhD for a Career in Industry? This module clarifies the benefits of obtaining a PhD for a career in industry and the skills acquired during PhD studies relevant to the career path in the industry. It presents different possibilities for doing a PhD in collaboration with industry partners, possible research topics, and requirements. It also addresses the possibility of creating a spin-out company.

(Module 4) How to find a PhD position? This module clarifies the formal and informal aspects of the application process, including finding a suitable PhD position, a group, and a supervisor. Students also learn about funding possibilities and how this choice influences their PhD studies.

(Module 5) How to handle doubts or problems before and during a PhD? This module clarifies typical concerns students have during their studies. This module emphasises the emotional issues that might arise and how to deal with them. Students learn where to find advice and support in case of self-doubts, burnout, and sexual harassment.

(Module 6) Considering a PhD abroad? This module clarifies different aspects of applying for PhD studies abroad. It addresses the application process, administrative issues, financial issues, and differences compared to a PhD in a native country.

(Module 7) Wrap-up This last session is used to wrap up the program and gather feedback from the participants about the program, to support its continuous monitoring and evolution.

In addition to the described topics, each module contains personal experience sharing. Therefore, general information is also observed from a personal point of view. Questions for personal experience sharing are identified based on each topic and its main general questions and should foster further discussion on these topics based on real situations. Personal questions enable getting to know the guest as a person and a role model. As an example, all the topics that should be addressed within Module 5, including the personal experience sharing, are given in [Appendix](#).

Lessons learned. The feedback from the various implementations has been overly positive. A number of lessons have been learned from the different implementations of the WoCa Lunch that have taken place. Organisers of previous lunches found that the questions were beneficial in reducing the preparation effort and should not necessarily be considered prescriptive.

It is crucial to establish with the students and the invited speakers that the communication approach should be informal and interactive. Invited speakers delivering slides do not have the same effect and impact as informal discussions. In certain situations, the WoCa lunch was not part of a departmental strategy and had little or no financial support. Therefore, obtaining funds for the lunch itself required additional effort.

Various experiences have shown that, typically, it is the women in the departments that organise and run these events. This reflects the experience shown with other gender inclusion initiatives where the women are those that are mostly involved ([Armstrong and Sullivan, 2024](#)). Having male allies working side by side with women in these initiatives can have a positive effect, as it helps to frame these initiatives as beneficial for the whole community rather than for women alone ([Farrell et al., 2021](#)).

9. Conclusions and further work

Relatively few women pursue PhD studies in computer science. This is problematic not only for equality reasons but also for the untapped potential that more women with PhDs in computer science could bring to industry and academia. The underlying goal of the research reported in this article is to increase the percentage of females in computer science, focusing on their transition to a PhD.

We conducted an extensive survey to characterise the encouraging and discouraging factors leading to the decision to start a PhD and to determine the extent to which these factors vary by gender. As expected, our results suggest this is a multi-factor decision, with several encouraging and discouraging factors, whose combination ultimately leads to a conclusion.

Encouraging factors towards PhD study include interest and confidence in research, arising from research involvement during earlier studies; enthusiasm for computer science and self-confidence in addition to an interest in an academic career; encouragement from external sources, including academic staff, peers, family and the industry, although the latter is relatively uncommon; and a positive perception towards PhD studies which can involve achieving personal goals. The effect sizes of these factors are mostly medium and small.

Discouraging factors include uncertainty and lack of knowledge of the PhD process, a perception of lower job flexibility and the requirement for long-term commitment, with medium to small effect sizes.

Gender differences highlighted that female students who pursue a PhD have less confidence in their technical skills than males but a higher preference for interdisciplinary areas. Females are less inclined than males to perceive the industry as offering better job opportunities and more flexible career paths than academia, although with a small effect size.

These insights helped develop a question catalogue to initiate discussions with potential female Ph.D. students. The catalogue covers basic information and practical issues, career opportunities, and emotional support for Ph.D. students.

We used the catalogue in the definition of a WoCa Lunch program as a supporting initiative to acquaint female undergraduate and master's students with the possibilities for advancing their careers if they pursue PhD studies. We successfully conducted two pilot iterations of this program at the University of Aachen, with positive feedback from participants. We created localised versions of the WoCa Lunch program guidelines in 8 different languages: English, French, German, Portuguese, Serbian, Spanish, Turkish, and Ukrainian.

We plan to conduct the WoCa lunch program in different countries and monitor its short and long-term effects, using appropriate key process indicators, such as the evolution of the number of enrolled and graduated Ph.D. students following their participation in the program. This will require a longitudinal study, spanning several years. We acknowledge the challenges of having the program in different languages, countries, and cultures. These may require local adaptations. Monitoring the impact of running this program in different countries will be crucial for its continuous improvement.

CRedit authorship contribution statement

Erika Ábrahám: Funding acquisition, Project administration, Validation, Formal analysis, Methodology, Writing – original draft, Conceptualization, Resources, Writing – review & editing, Investigation, Visualization. **Miguel Goulão:** Data curation, Methodology, Visualization, Formal analysis, Writing – original draft, Project administration, Investigation, Validation, Writing – review & editing, Conceptualization. **Milena Vujošević Janičić:** Data curation, Investigation, Validation, Writing – original draft, Conceptualization, Formal analysis, Methodology, Visualization, Writing – review & editing. **Sarah Jane Delany:** Formal analysis, Writing – review & editing, Data curation,

Investigation, Validation, Writing – original draft, Conceptualization, Methodology, Visualization. **Amal Mersni**: Formal analysis, Visualization, Writing – review & editing, Methodology, Conceptualization, Writing – original draft, Data curation, Investigation, Validation. **Oleksandra Yeremenko**: Investigation, Validation, Writing – original draft, Data curation, Conceptualization, Formal analysis, Methodology, Writing – review & editing, Visualization. **Özge Büyükdaglı**: Data curation, Methodology, Visualization, Conceptualization, Formal analysis, Writing – original draft, Validation, Writing – review & editing, Investigation. **Karima Boudaoud**: Conceptualization, Methodology, Writing – original draft, Formal analysis, Validation, Investigation, Visualization, Writing – review & editing. **Caroline Oehlhorn**: Writing – original draft, Conceptualization, Validation, Resources, Writing – review & editing, Data curation. **Ute Schmid**: Investigation, Resources, Visualization, Conceptualization, Data curation, Validation, Methodology, Writing – review & editing, Writing – original draft. **Christina Büsing**: Conceptualization, Resources, Investigation, Writing – review & editing. **Helen Bolke-Hermanns**: Conceptualization, Methodology, Resources, Investigation, Writing – review & editing. **Kaja Köhnle**: Resources, Writing – review & editing, Investigation, Conceptualization, Methodology. **Matilde Pato**: Methodology, Investigation, Writing – review & editing, Writing – original draft. **Deniz Sunar Cerci**: Conceptualization, Investigation, Methodology. **Larissa Schmid**: Data curation, Writing – original draft, Methodology, Conceptualization, Visualization, Formal analysis, Investigation, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by COST Action CA19122 - EUGAIN (European Network for Gender Balance in Informatics). The authors also acknowledge the financial support by the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia, grant number 451-03-47/2023-01/ 200104, the support from UID/04516/NOVA Laboratory for Computer Science and Informatics (NOVA LINCIS) with the financial support of FCT/IP, and by the pilot program Core Informatics at KIT (KiKIT) of the Helmholtz Association (HGF).

Appendix. Module 5: How to handle doubts or problems before and during a PhD?

General questions

- Am I capable of pursuing a PhD? Am I smart enough?
- Is doing a PhD compatible with family planning?
- What do I do if I am unsure or have problems during my doctorate studies?
- What are typical questions and concerns students may have about their doctorate studies?
- Will I have enough innovative ideas? What if I am not able to implement my research ideas? Can I change, or adjust, my research focus during the PhD?
- What do I do if I “run out of breath” during the doctorate? Is it just me? (Everyone knows this feeling)
- What do I do if I am dissatisfied with my doctoral supervision? (Time, support, appreciation)
- Am I okay with being one of the few women in the group? What can I do if I receive sexist remarks? Where can I find help for sexual harassment?

- Who can I contact if I have problems? Where can I find advice and support?
- How are the dignity and respect rights of the Ph.D. student applied in your University?

Personal experience sharing

- How did you deal with any doubts you had? How did you deal with the feeling that you might not be smart enough for a PhD?
- What obstacles/uncertainties did you face during your doctorate?
- Who did you get support from?

Data availability

The data is available in a companion site referenced in the text. <https://doi.org/10.5281/zenodo.16571934>.

References

- Abraham, E., Goulão, M., Vujošević Janićić, M., Delany, S.J., Mersni, A., Yeremenko, O., Schmid, L., Nebot, A., Buyukdagli, O., Boudaoud, K., Sunar Cerci, D., 2023. Women's Career Lunch — Catalog of questions for the speakers. <http://dx.doi.org/10.5281/zenodo.10079419>.
- Adamecz-Völgyi, A., Jerrim, J., Pingault, J.-B., Shure, D., 2023. Overconfident boys: The gender gap in mathematics self-assessment. IZA Discuss. Pap. <http://dx.doi.org/10.2139/ssrn.4464593>.
- Alfermann, D., Holl, C., Reimann, S., 2021. "Should I stay or should I go?" indicators of dropout thoughts of doctoral students in computer science. Int. J. High. Educ. 10 (3), 246–258.
- Alvarado, C., Dodds, Z., Libeskind-Hadas, R., 2012. Increasing women's participation in computing at harvey mudd college. Acm Inroads 3 (4), 55–64. <http://dx.doi.org/10.1145/2381083.2381100>.
- Amir, B., Ralph, P., 2018. There is no random sampling in software engineering research. In: Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings. pp. 344–345. <http://dx.doi.org/10.1145/3183440.3195001>.
- André, B.F., 2022. Gender Balance in Computer Science: How do Women View the Transition into a PhD?. (Master's thesis). Universidade Nova de Lisboa, Lisbon, Portugal, Available at https://run.unl.pt/bitstream/10362/155510/1/Andre_2022.pdf.
- Anon, 2023. Grad cohort workshop for women. 10.10.2023. <https://cra.org/cra-wp/grad-cohort-for-women>.
- Armstrong, J., Sullivan, A., 2024. A critical analysis of athena swan as a policy-scoring scheme. Br. Educ. Res. J.
- Baltes, S., Ralph, P., 2022. Sampling in software engineering research: A critical review and guidelines. Empir. Softw. Eng. 27 (4), 94. <http://dx.doi.org/10.1007/s10664-021-10072-8>.
- Beaubouef, T., Mason, J., 2005. Why the high attrition rate for computer science students: some thoughts and observations. ACM SIGCSE Bull. 37 (2), 103–106.
- Beyer, S., 2014. Why are women underrepresented in computer science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. Comput. Sci. Educ. 24 (2–3), 153–192. <http://dx.doi.org/10.1080/08993408.2014.963363>.
- Beyer, S., Rynes, K., Perrault, J., Hay, K., Haller, S., 2003. Gender differences in computer science students. In: Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education. pp. 49–53. <http://dx.doi.org/10.1145/611892.611930>.
- Bizot, B., Zweben, S., 2020. 2020 taulbee survey. URL <https://cra.org/wp-content/uploads/2021/05/2020-CRA-Taulbee-Survey.pdf>.
- Bonferroni, C., 1936. Teoria statistica delle classi e calcolo delle probabilita. Pubbl. Del R Ist. Super. Sci. Econ. E Commerciali di Firenze 8, 3–62.
- Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qual. Res. Psychol. 3 (2), 77–101. <http://dx.doi.org/10.1191/1478088706qp0630a>.
- Bravata, D.M., Watts, S.A., Keefer, A.L., Madhusudhan, D.K., Taylor, K.T., Clark, D.M., Nelson, R.S., Cokley, K.O., Hagg, H.K., 2020. Prevalence, Predictors, and Treatment of Impostor Syndrome: a Systematic Review. J. Gen. Intern. Med. 35 (4), 1252–1275. <http://dx.doi.org/10.1007/s11606-019-05364-1>.
- at Brown University, C.S., 2023. Helping students from underrepresented groups apply to CS phd programs. <https://cs.brown.edu/degrees/doctoral/applications/helpful-resources-applying-computer-science-phd-programs/>.

- Cohoon, J., 2003. Must there be so few? Including women in CS. In: 25th International Conference on Software Engineering, 2003. Proceedings. IEEE, pp. 668–674.
- Cohoon, J.M., Wu, Z., Chao, J., 2009. Sexism: toxic to women's persistence in CSE doctoral programs. *Acm Sigcse Bull.* 41 (1), 158–162. <http://dx.doi.org/10.1145/1539024.1508924>.
- Department of Computer Science, S.B.U., 2023. Celebrating the success of women PhD students in computer science. <https://www.cs.stonybrook.edu/about-us/News/Celebrating-Success-Women-PhD-Students-Computer-Science>.
- computerscience.org, 2023. Celebrating the success of women PhD students in computer science. <https://www.computerscience.org/degrees/phd/online/information-technology/>.
- Corallo, S., Mazkatli, M., Rapp, M., Hajiabadi, H., Kaplan, A., Kuehn, R., Schmid, L., Singh, S., 2022. Toward awareness creation of common challenges women are facing in academia: A study from a german perspective. In: Scandurra, P., Galster, M., Mirandola, R., Weyns, D. (Eds.), *Software Architecture*. Springer International Publishing, Cham, pp. 47–62. http://dx.doi.org/10.1007/978-3-031-15116-3_3.
- Corbett, C., Hill, C., 2015. Solving the equation - AAUW. URL <https://www.aauw.org/app/uploads/2020/03/Solving-the-Equation-report-nsa.pdf>.
- Cramér, H., 1946. *Mathematical Methods of Statistics*. Princeton University Press.
- Cronin, C., Roger, A., 1999. Theorizing progress: Women in science, engineering, and technology in higher education. *J. Res. Sci. Teach.* 36, 637–661. [http://dx.doi.org/10.1002/\(SICI\)1098-2736\(199908\)36:6<637::AID-TEA4>3.0.CO;2-9](http://dx.doi.org/10.1002/(SICI)1098-2736(199908)36:6<637::AID-TEA4>3.0.CO;2-9).
- Cunningham, M., Wells, M., 2017. Qualitative analysis of 6961 free-text comments from the first National Cancer Patient Experience Survey in Scotland. *BMJ Open* 7 (6), e015726. <http://dx.doi.org/10.1136/bmjopen-2016-015726>.
- de Aguiar Gomes, L.L., Carvalho, J.R.H., Lauschner, T., Nakamura, F.G., de Freitas, R., 2018. Encouraging women to pursue a computer science career in the context of a third world country. In: 2018 IEEE Frontiers in Education Conference. FIE, IEEE, pp. 1–5.
- Denisco, A., 2017. The state of women in computer science: An investigative report. Tech Repub..
- Di Nitto, E., García-Varea, I., Jazayeri, M., Tamburri, D.A., Tikhonenko, S., 2023. Informatics higher education in Europe: A data portal and case study. *Commun. ACM* 66, 61–67. <http://dx.doi.org/10.1145/3591107>.
- Diogo, S., Gonçalves, A., Cardoso, S., Carvalho, T., 2022. Tales of doctoral students: Motivations and expectations on the route to the unknown. *Educ. Sci.* 12 (4), 286.
- Europe, I., 2023. Best practices in supporting women. URL <https://www.informatics-europe.org/society/minerva-informatics-equality-award/best-practices-in-supporting-women.html>.
- Farrell, L., Petzel, Z.W., McCormack, T., Turner, R.N., Rafferty, K., Latu, I.M., 2021. When you put it that way: Framing gender equality initiatives to improve engagement among STEM academics. *BioScience* 71 (3), 292–304.
- Fisher, A.J., Mendoza-Denton, R., Patt, C., Young, L., Eppig, A., Garrell, R.L., Rees, D.C., Nelson, T.W., Richards, M.A., 2019. Structure and belonging: Pathways to success for underrepresented minority and women PhD students in STEM fields. *PLoS One* 14 (1), e0209279.
- Funk, C., Parker, K., 2020. Women and men in STEM often at odds over workplace equity. URL <https://www.pewresearch.org/social-trends/2018/01/09/women-and-men-in-stem-often-at-odds-over-workplace-equity/>.
- Genut, S., Ori, B., Kolikant, Y.B.-D., 2019. Factors influencing women's decision to study computer science: Is it context dependent? *Issues Informing Sci. Inf. Technol.* 16, 127–141. <http://dx.doi.org/10.28945/4296>.
- Happe, L., Buhnova, B., 2021. Frustrations steering women away from software engineering. *IEEE Softw.* 39 (4), 63–69.
- Hill, C., Corbett, C., St. Rose, A., 2010. Why so few? - AAUW. URL <https://www.aauw.org/app/uploads/2020/03/why-so-few-research.pdf>.
- Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 65–70. <https://www.jstor.org/stable/4615733>.
- Horta, H., Cattaneo, M., Meoli, M., 2018. Phd funding as a determinant of Phd and career research performance. *Stud. High. Educ.* 43 (3), 542–570. <http://dx.doi.org/10.1080/03075079.2016.1185406>.
- Hyrnsalmi, S.M., 2019. The underrepresentation of women in the software industry: Thoughts from career-changing women. In: 2019 IEEE/ACM 2nd International Workshop on Gender Equality in Software Engineering (GE). pp. 1–4. <http://dx.doi.org/10.1109/GE.2019.00008>.
- in Informatics, E.N.F.G.B., 2022. 1st international EUGAIN summer training school (EUGAIN-STS 2022). <https://sites.google.com/view/eugain-training-schools/eugain-sts-2022>.
- in Informatics, E.N.F.G.B., 2023. 2nd international EUGAIN summer training school: Research recharge (june 14-15, 2023 – rome, Italy). <https://eugain.training/>.
- Jaccheri, L., Pereira, C., Fast, S., 2020. Gender issues in computer science: lessons learnt and reflections for the future. In: 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing. SYNASC, IEEE, pp. 9–16. <http://dx.doi.org/10.1109/SYNASC51798.2020.00014>.
- Koc University, G.E.O., 2023. Teaching. <https://geo.ku.edu.tr/teaching/>.
- Lagesen, V.A., 2007. The strength of numbers: Strategies to include women into computer science. *Soc. Stud. Sci.* 37 (1), 67–92.
- Lagesen, V.A., 2008. A cyberfeminist utopia?: Perceptions of gender and computer science among Malaysian women computer science students and faculty. *Sci. Technol. Hum. Values* 33 (1), 5–27. <http://dx.doi.org/10.1177/0162243907306192>.
- Lindner, R., 2020. Barriers to doctoral education, equality, diversity and inclusion for postgraduate research students at UCL. <https://www.grad.ucl.ac.uk/strategy/barriers-to-doctoral-education.pdf>.
- Lockett, V., 2023. Supporting women in PhD programs: Scholarships and resources to optimize success. <https://www.phds.me/financial-aid/scholarships-and-grants/women/>.
- London, J., Cox, M.F., Ahn, B., Branch, S., Zephirin, T., Torres-Ayala, A., Zhu, J., 2014. Motivations for pursuing an engineering PhD and perceptions of its added value: A US-based study. *Int. J. Dr. Stud.* 9, 205.
- Lovitts, B., 2005. Being a good course-taker is not enough: a theoretical perspective on the transition to independent research. *Stud. High. Educ.* 30, 137–154. <http://dx.doi.org/10.1080/03075070500043093>.
- McCallum, C.M., 2012. Understanding the relationships and experiences that contribute to African Americans' decision to enrol in doctoral education (Ph.D. thesis). University of Michigan.
- Miliszewska, I., Barker, G., Henderson, F., Sztendur, E., 2006. The issue of gender equity in computer science – what students say. *J. Inf. Technol. Educ.: Res.* 5 (1), 107–120, URL <https://www.learntechlib.org/p/111535>.
- Moreno, M.d.C.C., Kollanus, S., 2013. On the motivations to enroll in doctoral studies in computer science—A comparison of PhD program models. In: 2013 12th International Conference on Information Technology Based Higher Education and Training. ITHET, IEEE, pp. 1–8.
- Moskal, B., 2002. Female computer science doctorates: what does the survey of earned doctorates reveal? *ACM SIGCSE Bull.* 34 (2), 105–111. <http://dx.doi.org/10.1145/543812.543841>.
- Motogna, S., Alboia, L., Todericiu, I.A., Zaharia, C., 2022. Retaining women in computer science: the good, the bad and the ugly sides. In: Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering. pp. 35–42.
- Mullen, A.L., Goyette, K.A., Soares, J.A., 2003. Who goes to graduate school? Social and academic correlates of educational continuation after college. *Sociol. Educ.* 143–169. <http://dx.doi.org/10.2307/3090274>.
- Munoz-Hernandez, S., Osman, I., 2016. Using a productive distance PhD program to empower women in academia. In: 2016 IST-Africa Week Conference. IEEE, pp. 1–7. <http://dx.doi.org/10.1109/ISTAFRICA.2016.7530698>.
- Pantic, K., 2020. Retention of Women in Computer Science: Why Women Persist in Their Computer Science Majors (Ph.D. thesis). Utah State University, URL <https://www.proquest.com/dissertations-theses/retention-women-computer-science-why-persist/docview/2454679255/se-2?accountid=10673>.
- Pantic, K., Clarke-Midura, J., 2019. Factors that influence retention of women in the computer science major: A systematic literature review. *J. Women Minor. Sci. Eng.* 25 (2), <http://dx.doi.org/10.1615/JWomenMinorSciEng.2019024384>.
- Pappas, I.O., Giannakos, M.N., Jaccheri, L., 2016. Investigating factors influencing students' intention to dropout computer science studies. In: Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education. pp. 198–203.
- Perneger, T.V., 1998. What's wrong with Bonferroni adjustments. *BMJ* 316 (7139), 1236–1238. <http://dx.doi.org/10.1136/bmj.316.7139.1236>.
- QBIT, 2023. How Many Programmers are there in the World and in the US? [2023]. Tech. Rep., QBIT Labs, URL <https://qubit-labs.com/how-many-programmers-in-the-world/>.
- Roberts, E.S., Kassianidou, M., Irani, L., 2002. Encouraging women in computer science. *ACM SIGCSE Bull.* 34 (2), 84–88.
- Rowe, N., 2021. The realities of completing a phd: How to plan for success. In: Routledge Research in Education, Taylor & Francis, URL <https://books.google.pt/books?id=Id0NEAAQBAJ>.
- Rubegni, E., Penzenstadler, B., Landoni, M., Jaccheri, L., Dodig-Crmkovic, G., 2023. Owning your career paths: Storytelling to engage women in computer science. In: Gender in AI and Robotics: The Gender Challenges from an Interdisciplinary Perspective. Springer, pp. 1–25. http://dx.doi.org/10.1007/978-3-031-21606-0_1.
- Ryan, R.M., Deci, E.L., 2000a. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemp. Educ. Psychol.* 25 (1), 54–67.
- Ryan, R.M., Deci, E.L., 2000b. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55 (1), 68.
- of Science, N.U., Technology, 2023. Gender equality and diversity at NTNU. <https://www.ntnu.edu/genderequality>.
- Soe, L., Yakura, E.K., 2008. What's wrong with the pipeline? Assumptions about gender and culture in IT work. *Women's Stud.* 37 (3), 176–201. <http://dx.doi.org/10.1080/00497870801917028>.

- Stanford Engineering, C.S., 2023. Incoming CS PhD students. <https://www.cs.stanford.edu/student-services-overview/new-student-orientation/incoming-cs-phd-students>.
- Steele, H., 2023. Salary outlook with PhD in computer science. URL <https://www.businessstudent.com/careers/salary-outlook-phd-computer-science/>.
- Stross, R., 2008. What has driven women out of computer science. N. Y. Times 15, URL <https://www.nytimes.com/2008/11/16/business/16digi.html>.
- Tarvid, A., 2014. Motivation to study for PhD degree: Case of Latvia. *Procedia Econ. Financ.* 14, 585–594.
- Varma, R., 2010. Computing Self-Efficacy Among women in India. *J. Women Minor. Sci. Eng.* 16 (3), 257–274. <http://dx.doi.org/10.1615/JWomenMinorScienEng.v16.i3.40>.
- Vitores, A., Gil-Juárez, A., 2016. The trouble with ‘women in computing’: a critical examination of the deployment of research on the gender gap in computer science. *J. Gend. Stud.* 25 (6), 666–680. <http://dx.doi.org/10.1080/09589236.2015.1087309>.
- Waaaijer, C.J., 2017. Perceived career prospects and their influence on the sector of employment of recent PhD graduates. *Sci. Public Policy* 44 (1), 1–12. <http://dx.doi.org/10.1093/scipol/scw007>.
- Wagner, S., Mendez, D., Felderer, M., Graziotin, D., Kalinowski, M., 2020. Challenges in survey research. *Contemp. Empir. Methods Softw. Eng.* 93–125. http://dx.doi.org/10.1007/978-3-030-32489-6_4.
- Wang, Y., Redmiles, D., 2019. 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. <http://dx.doi.org/10.1109/ICSE-SEIS.2019.00009>.
- Webber, L., Dismore, H., 2021. Mothers and higher education: Balancing time, study and space. *J. Furth. High. Educ.* 45 (6), 803–817. <http://dx.doi.org/10.1080/0309877X.2020.1820458>.
- Wollast, R., Aelenei, C., Chevalère, J., Van der Linden, N., Galand, B., Azzi, A., Frenay, M., Klein, O., 2023. Facing the dropout crisis among PhD candidates: the role of supervisor support in emotional well-being and intended doctoral persistence among men and women. *Stud. High. Educ.* 48 (6), 813–828.
- Yamane, T., 1973. *Statistics: An Introductory Analysis*. Harper & Row New York.
- Yates, J., Plagnol, A.C., 2022. Female computer science students: A qualitative exploration of women’s experiences studying computer science at university in the UK. *Educ. Inf. Technol.* 27 (3), 3079–3105. <http://dx.doi.org/10.1007/s10639-021-10743-5>.

Erika Ábrahám graduated from the Christian-Albrechts-University Kiel (Germany), and received her Ph.D. from the University of Leiden (the Netherlands) for her work on the development and application of deductive proof systems for concurrent programs. Then she moved to the Albert-Ludwigs-University Freiburg (Germany), where she started to work on the development and application of SAT and SMT solvers. Since 2008 she has been a professor at RWTH Aachen University (Germany), with her main research focus on SMT solving for real and integer arithmetic, and formal methods for probabilistic and hybrid systems.

Miguel Goulão is an Associate Professor at the NOVA University of Lisbon. He is an integrated member of the Software Systems group of NOVA LINES. He has a Ph.D. (2008) in Informatics from Universidade Nova de Lisboa. The broad aim of his research is to improve the software developers experience, to better deal with software development complexity. Miguel uses Evidence-Based, Empirical Software Engineering, and User Experience evaluation techniques to identify the strengths and shortcomings in languages, tools, and approaches. He is particularly interested in improving the understandability of Requirements Engineering and Domain-Specific Languages. Miguel is also active in research on Equity, Diversity and Inclusion in Software Engineering, and is a working group leader in the EUGAIN COST Action. Miguel published over 70 papers in peer-reviewed journals and conferences and served as guest editor of special issues in international journals, as a PC member, and as PC and Organising Chair in several events.

Milena Vujošević Janičić is an Associate Professor at the Department of Computer Science and Head of the Computer Science Study Program at the Faculty of Mathematics, University of Belgrade. She also works as a principal researcher at Oracle Labs. She received a Ph.D. in Computer Science from the University of Belgrade in 2013. Her main research interests are in compiler optimisations, automated bug finding, and the application of software verification techniques in different fields. Milena is also interested in research on Equity, Diversity and Inclusion in Computer Science and is an active member of the EUGAIN COST Action. Milena has published over 40 papers in peer-reviewed journals and conferences.

Sarah Jane Delany is a Professor of Inclusive Computer Science in the School of Computer Science at TU Dublin. She has a BA (Mod) in Mathematics from Trinity College Dublin and received her Ph.D. from Dublin Institute of Technology in 2006.

Before joining TU Dublin she worked for Accenture and ESBI Computing as an IT consultant. Her research interests are in the area of Machine Learning, focussing on the application of machine learning to real world problems, with specific interest in the areas of supervised machine learning, text analytics, bias in machine learning, concept drift and case-based reasoning. She is the TU Dublin lead of ML-Labs, the SFI Centre for Research Training in Machine Learning and a co-founder of the TU Dublin Applied Intelligence Research Center (AIRC). She was the TU Dublin lead for CeADAR, the National Centre for Applied AI for a number of years and has worked with many companies on the application of data science and machine learning to their business. Recent research projects include identifying and mitigating bias in training data for machine learning models and exploring methods in supervised machine learning and natural language processing to automate identifying gender issues and bias in text content. Sarah Jane has also received funding from the HEA Gender Equality Enhancement Fund for the TechMate project which will develop an online resource to help HEIs work towards gender equality in technology-focused courses.

Amal Mersni is an Assistant Professor in the Department of Engineering at the International University of Sarajevo. She earned her Ph.D. in Telecommunication Systems and Networks from Kharkiv National University of Radio Electronics in 2017. Her main research focuses on a variety of topics, including routing, traffic engineering, quality of service, cyber-resilience, network security, fault tolerance, and software-defined networking. She has contributed to over 40 publications in peer-reviewed journals and conferences and holds several international and industry certifications in networking and cybersecurity. Amal is a certified Cisco Instructor and an active member of IEEE. She is dedicated to promoting the involvement and success of women in computing through various activities under the auspices of ACM-W. Amal is committed to research on equity, diversity, and inclusion in computer science and actively participates in the EUGAIN COST Action.

Oleksandra Yeremenko is a Professor of V.V. Popovskyy Department of Infocommunication Engineering at the Kharkiv National University of Radio Electronics in Kharkiv, Ukraine. She has a Ph.D. (2008) and D.Eng.Sc. (2018) in Telecommunication Systems and Networks from Kharkiv National University of Radio Electronics. The broad aim of her research is to improve the cyber resilience of softwarized network architectures. Additional areas of focus include network resilience, fault-tolerant routing, and Quality of Service. Oleksandra is actively involved in several Erasmus+ Jean Monnet modules focused on integrating EU cybersecurity frameworks, policies, and resilience strategies in Ukraine. Oleksandra is a participant in the COST projects CA15127 Resilient Communication Services Protecting End-User Applications from Disaster-Based Failures (RECODIS) and CA19122 European Network for Gender Balance in Informatics (EUGAIN). She is an IEEE Senior Member. Oleksandra published over 100 papers in peer-reviewed journals and conferences, served as an editorial board member in international journals, and was an organising and program committee member at several conferences.

Özge Büyükdaglı is an Assistant Professor at the International University of Sarajevo. She earned her Ph.D. in Industrial Engineering from Yasar University, Türkiye in 2018. Her research interests include mathematical modelling, optimisation algorithms, and performance evaluation for production systems. She has actively participated in numerous industrial and research projects throughout her career. Currently, she serves as the Program Coordinator for the Software Engineering and Artificial Intelligence and Data Engineering programs and holds a current role as a member of the Sustainable Development Goals committee at her university. She is passionate about gender equality research and contributes to the European Network for Gender Balance in Informatics (EUGAIN).

Karima Boudaoud is Associate Professor at Université Côte d’Azur. Her main research focus on Human-centric cybersecurity and privacy putting the Human in the heart of Cybersecurity and privacy challenges and issues. She has participated in EU, Swiss & French security research projects. Member of TPCs of national & international conferences and workshops, she has published more than 90 scientific publications, conference and workshop papers. Karima is also involved in gender balance activities for more than 10 years and an active member of the Femmes & Sciences Association since 2013. She is a working group leader in the EUGAIN (European Network For Gender Balance in Informatics) COST Action.

Caroline Oehlhorn is a Ph.D. student in information systems at the University of Bamberg, Germany. Her research focuses on women in IT, gender diversity as well as IT human resource management. She received the universitys Bettina Paetzold Award for Good Gender Education in 2022.

Ute Schmid is a full professor of cognitive systems at the University of Bamberg. She holds a diploma in psychology and a diploma in computer science, both from Technical University Berlin, Germany. Ute received her doctoral degree in computer science in 1994 and her habilitation in computer science in 2002. From 1994 to 2001 she was an assistant professor at the Methods of AI/Machine Learning group,

Department of Computer Science. Afterwards, she worked as a lecturer for Intelligent Systems at the Department of Mathematics and Computer Science at the University Osnabrück and was a member of the Cognitive Science Institute. She has more than 20 years of experience in AI research, especially in the domain of interpretable machine learning and explainable AI. Since 2005, Ute Schmid has been the faculty's Women representative and dedicates a significant amount of her time to supporting women in computer science as well as to promoting computer science in elementary, primary, and secondary education. She won the Minerva Gender Equality Award of Informatics Europe 2018 for her university.

Christina Büsing studied Mathematics at the WWU Münster, the Universidad Complutense de Madrid and the Technical University of Berlin. She did her dissertation on the topic of Recoverable Robustness in Combinatorial Optimisation. She then worked as a PostDoc in Aachen, Lancaster and Vienna. From 2016 to 2022 Christina Büsing served as Junior Professor for Robust Planning in Medical Care at the RWTH Aachen and since 2022 she is Full Professor for Combinatorial Optimisation at RWTH Aachen University. She leads the Teaching and Research Group on Combinatorial Optimisation and her research interests include optimisation under uncertainty, robust optimisation, scenario generation, combinatorial optimisation, complexity theory, optimality criteria and both exact and heuristic algorithms.

Helen Bolke-Hermanns holds a degree in Mathematical and Technical Software Engineering (MATSE), RWTH Aachen University. Her practical experience spans system and network administration, as well as the organisation of internships and academic courses. She is also actively involved in coordinating DFG Research Training Groups and contributes to the Computer Science Study Centre at RWTH Aachen University, with a focus on public relations, communication, and reporting. As a woman with a degree in a mathematical-technical field, she has been a passionate advocate for equal opportunities for women in computer science from the very beginning of her career.

Graduated as a mathematical-technical assistant (MATSE). She has many years of experience at the Chair of Computer Science 4 with practical experience in system and network administration, and implementation of internships and courses. She works on the Coordination of DFG RTGs with a focus on IT, PR and communication, and collection and evaluation of statistics in the Department of Computer Science. She has a strong commitment to women's equality in computer science.

Kaja Köhnle received her Master's degree in Gender Studies at the Albert-Ludwigs-University Freiburg (Germany). Since 2020 she has been the project coordinator of "Career Support and Networking for Early-Career Women Researchers" at the Equal Opportunity Office at RWTH Aachen University (Germany). She specialises in supporting female students, doctoral candidates and Post Docs in the development of their academic careers and professional networks.

Matilde Pato (Ph.D. in Biomedical Engineering from IST-UTL in 2010) is an Adjunct Professor at ISEL-IPL and a researcher at IBEB and LASIGE (FCUL) and NOVA LINES (FCT-UNL). She has supervised and co-supervised several M.Sc. theses in Computer Science and is currently co-supervising two Ph.D. students at FCT-NOVA in Informatics. She has co-authored dozens of papers in national and international journals and conference proceedings. Currently, she is part of the Management Committee of the COST ACTION EUGAIN (European Network For Gender Balance in Informatics Management Committee) in WG2 (from Bachelor/Master studies to Ph.D.). She participates in European and national R&D projects with industrial partners and other universities. She participated in organising a Lecture-Style Tutorial on KDD 2021 entitled "Creating Recommender Systems Datasets in Scientific Fields". Furthermore, she has a talk where teachers and researchers can discuss computer science topics. Currently, her research focuses on creating a Telemonitoring system for chronically ill patients and developing recommender datasets for healthcare.

Deniz Sunar Cerci is a Professor at Yildiz Technical University, in Turkey, a member of the European Organization for Nuclear Research in Genève, Switzerland and of the Compact Muon Solenoid (CMS) Collaboration (Switzerland). She has a Ph.D. in Physics from the University of Antwerp, in the Netherlands. Her research interests focus on High Energy and Particle Physics. She has participated in several international research projects and is a member of the CA19122 European Network for Gender Balance in Informatics (EUGAIN). Deniz published over 200 journal and congress papers.

Larissa Schmid obtained her PhD at Karlsruhe Institute of Technology (KIT) on performance modeling and analysis of configurable scientific software and workflows. Larissa is a postdoctoral researcher at the Department of Computer Science, Division of Theoretical Computer Science, of the KTH Royal Institute of Technology, in the group of Martin Monperrus. Her current research interests are within software supply chain security and software plagiarism detection. Larissa is also interested in research on Equity, Diversity, and Inclusion in Computer Science. Her contribution to this work was made while affiliated with KIT.