

The Hands-Up Problem and How to Deal With It: Secondary School Teachers' Experiences of Debugging in the Classroom

Laurie GALE^[0009-0004-4299-6704], Sue SENTANCE^[0000-0002-0259-7408]

Raspberry Pi Computing Education Research Centre, University of Cambridge, United Kingdom

e-mail: lpg28@cst.cam.ac.uk, ss2600@cst.cam.ac.uk

Abstract. Debugging is a vital but challenging skill for beginner programmers to learn. It is also a difficult skill to teach. For secondary school teachers, who may lack time or programming experience, honing students' understanding of debugging can be a daunting task. Despite this, little research has explored their perspectives of debugging. To this end, we investigated secondary teachers' experiences of debugging in the classroom, with a focus on text-based programming. Through thematic analysis of nine semi-structured interviews, we identified a common reliance on the teacher for debugging support, embodied by many raised hands. We call this phenomenon the 'hands-up problem'. While more experienced and confident teachers discussed strategies they use to counteract this, less confident teachers discussed the negative consequences of this problem. We recommend further research into debugging-specific pedagogical content knowledge and professional development to help less confident teachers develop approaches for supporting their students with debugging.

Key words: Debugging education, Secondary teachers, Computing education.

1. Introduction

The process of finding and fixing errors in a computer program is one of the most fundamental but challenging skills to learn for the beginner programmer. Learners must have sufficient understanding of several components of the programming language they are using to effectively and reliably resolve errors, which is rarely present in practice. Instead, beginners are commonly observed utilising ineffective debugging strategies (e.g., (Perkins et al., 1986; Fitzgerald et al., 2008; Liu et al., 2017)) which can easily stall progress and evoke emotional distress (Kinnunen and Simon, 2010; Gorson et al., 2022).

Secondary school teachers play a key role in developing their students' debugging ability. In an ideal classroom environment, they are on hand to teach students effective debugging strategies, support those struggling to debug, and foster positive attitudes towards errors. In practice, this is not always the case, with significant challenges unique to the school context. Teachers may not be confident in their programming ability (Yadav et al., 2016b;

Royal Society, 2017; Sentance and Csizmadia, 2017), lack the lesson time to explicitly teach debugging (Sentance and Csizmadia, 2017; Michaeli and Romeike, 2019a), or have to support a large number of students with fixing errors (Michaeli and Romeike, 2019a). These challenges are arguably exacerbated when teaching text-based programming languages (TBPLs), which is stipulated in many secondary-level national curricula (e.g., (Department for Education, 2013b; British Columbia Ministry of Education, 2016; Sysło, 2020; Ripani and Vazquez-Brust, 2023; Skolverket, 2024)). Although students perceive these languages as more realistic and expressive than block-based languages (Weintrop and Wilensky, 2015), they have a far larger scope for errors and subsequent struggle with resolving them (Jadud, 2006; Denny et al., 2011). Teaching with TBPLs also poses an additional layer of difficulty for time-poor or confidence-lacking teachers. These factors can easily make debugging a significant source of negative experiences for students and teachers in secondary school programming lessons.

While most debugging education research has focused on the student perspective, teachers' classroom experiences are far less studied. This work is always important to conduct; understanding teachers' experiences is essential for ensuring that research remains close to the challenges teachers are facing in the classroom. Focusing on debugging is particularly important as (text-based) programming becomes more widely taught in schools around the world.

This paper reports on secondary computing teachers' experiences with the teaching and learning of debugging in a text-based programming language in their classrooms. Through a thematic analysis of semi-structured interviews with nine teachers, we report on teachers' current approaches for teaching various aspects of debugging, their observations of students engaging in the debugging process, and their challenges with teaching it. The research questions for this study are:

RQ1: What are secondary school computing teachers' experiences and perspectives relating to debugging in the classroom?

RQ2: What are secondary school computing teachers' observations relating to their students' debugging experiences?

Together with related work, we use our findings to conceptualise a significant in-class challenge for computing teachers that derives from student struggle with debugging, titled *the hands-up problem*. We introduce this as a representation of a classroom environment un conducive to positive experiences with learning to program in Section 5. We end with recommendations on how the hands-up problem could be alleviated and further developed as a concept. Our findings can be used by school teachers and researchers seeking to understand other teachers' classroom experiences with debugging and improve the classroom environment of programming lessons.

2. Related Work

An increasing amount of computing education research is being directed towards the

teaching and learning of debugging. This is for good reason: debugging is a core aspect of programming, difficult to master, and has well-documented challenges. Even as generative artificial intelligence becomes more prevalent in programming education, debugging is still, and may be increasingly, a vital aspect of programming (Denny et al., 2024).

2.1 Student Perspectives of Debugging

Most debugging education research has focused on the student perspective. Students learning to program naturally have fragile knowledge (Perkins and Martin, 1986) and misconceptions which often manifest as errors in their programs (Pea, 1986; Ettles et al., 2018; Sorva, 2021). For students learning a TBPL, the scope for errors is naturally far larger than in BBPLs. Even minor syntactical errors can pose a significant barrier to a working program (Jadud, 2006; Denny et al., 2011), not helped by unclear programming error messages (Becker et al., 2019). These challenges arguably make debugging in TBPLs more difficult than in BBPLs.

When trying to fix these errors, strategies that do not aid successful debugging have been observed among students using block-based (Liu et al., 2017; Yan et al., 2025) and text-based (Fitzgerald et al., 2008; Whalley et al., 2023) environments. These often embody a cycle of unreflective tinkering (Perkins et al., 1986), which makes struggle with debugging likely. This struggle can evoke strong negative emotions (Kinnunen and Simon, 2010; Gorson et al., 2022) and ultimately dampen students' experiences of learning to program.

The challenges beginner programmers face when debugging have motivated an increased focus on debugging-specific teaching aids over the last few decades. These include systematic debugging processes (e.g., (Carver and Risinger, 1987; Michaeli and Romeike, 2019b)), debugging-specific tooling (e.g., (Guo, 2013; Luxton-Reilly et al., 2018; Ziemann and Reu 2024)) and beginner-friendly programming error messages (e.g., (Becker, 2016; Denny et al., 2020; Phung et al., 2023)). Many of these have been specifically designed for TBPLs. Summaries of pedagogical debugging 'interventions' have found initial positive effects on measures of debugging ability (Sun et al., 2024; Yang et al., 2024), although evidence of wider adoption is lacking.

2.2 Teacher Perspectives of Debugging in the Classroom

Compared to student-centric research, little work has focused on teachers' perspectives on teaching debugging. Naturally, students' experiences with debugging in the classroom are mediated by the teacher, who may employ a range of strategies to develop students' confidence in resolving errors. However, school teachers face many challenges in providing high-quality debugging instruction. At a subject level, teachers have reported a lack of time and confidence to teach computing (Yadav et al., 2016b; Royal Society, 2017; Sentance and Csizmadia, 2017). As a consequence, teachers may lack time to explicitly teach debugging or be inundated with students in need of support (Michaeli and Romeike, 2019a).

To our knowledge, very few studies have specifically investigated teachers' experiences of debugging in the classroom, despite the aforementioned challenges. Rather, related work has generally investigated teacher-student interactions or concepts related to pedagogical content knowledge.

2.2.1 *Teacher-Student Interactions*

In a classroom setting, teachers frequently assist students with debugging (Michaeli and Romeike, 2019a). These interactions vary widely, but can generally be viewed through the lens of *cognitive apprenticeship* (Collins et al., 1991), an instructional paradigm for teaching problem-solving processes using the principles of traditional apprenticeship.

Teachers' use of *scaffolding* is a major theme in many studies of teacher-student interactions. A popular and general scaffolding technique is to ask students questions, with the content of the questions determining the level of scaffolding provided. Some teachers ask open-ended and generic questions (Nixon et al., 2023; Ryan and DeLiema, 2023; Elliott et al., 2024), providing relatively little scaffolding, while other questions may be far more specific (Yadav et al., 2016a; Flood et al., 2018; Nixon et al., 2023; Elliott et al., 2024; Wachter and Michaeli, 2025). Teachers' initial questions to students may help the teacher to understand the error(s) in students' programs (Elliott et al., 2024; Wachter and Michaeli, 2025). More open-ended questions also encourage students to *articulate* their thoughts, another aspect of cognitive apprenticeship important for debugging.

More debugging-specific techniques include *modelling* the debugging process (Ryan and DeLiema, 2023; Wachter and Michaeli, 2025). Unsurprisingly, students' debugging strategies sometimes differ from their teachers', at which point teachers may decide to tailor their approach to the students' strategies (Ryan and DeLiema, 2023).

While general techniques used by teachers in their interactions with students are useful to know, the support provided is incredibly contextual. Nixon et al. (2023) frame teacher-student interactions as improvisational, meaning teachers' general approaches must be adapted to a specific interaction. Other in-class factors, such as the time remaining in the lesson, the number of students in the classroom, and the type of error in question, also influence the support the teacher will provide Elliott et al. (2024). In the same vein, these studies each have their own context, some of which are not representative of an actual classroom environment. Some were conducted in summer schools (Ryan and DeLiema, 2023) or simulations of classrooms (Nixon et al., 2023; Wachter and Michaeli, 2025), so not all findings necessarily reflect teachers' actual in-class interactions with students.

2.2.2 *Pedagogical Content Knowledge and Professional Development*

The approaches that teachers adopt when providing debugging support, and the knowledge required to employ them, form part of their 'debugging pedagogical content knowledge' (PCK) (Shulman, 1986; Carlson et al., 2019). Shulman (1986) defines this as "the ways of representing and formulating the subject that make it comprehensible to others" (p. 9).

PCK can be developed in several ways: teaching experience, research-informed teaching resources, or professional development (PD) (Shulman, 1986). Some debugging-specific PD activities have been developed in previous work, involving frequent teacher reflection (Michaeli and Romeike, 2022) and simulating the student perspective (Elliott et al., 2022; Nixon et al., 2023). Although analyses of these activities only included small cohorts of teachers, they reported benefitting from a low-stakes environment to improve their knowledge of debugging strategies, common errors, and teaching approaches (Elliott et al., 2022; Michaeli and Romeike, 2022; Nixon et al., 2023).

Despite the volume of pedagogical and content knowledge required to teach debugging effectively, we are not aware of a model of PCK specifically for debugging. One study analysed school teachers' level of debugging PCK (Tsan et al., 2022), but did not propose a formal model for it and focused on Scratch. Related work has focused on developing computing or programming PCK more generally (Saeli et al., 2011; Hubwieser et al., 2013). A related concept that could inform the development of a debugging PCK is the idea of debugging learning trajectories (Rich et al., 2019; Du et al., 2021). Although designed for curriculum development, they could be used to intuitively visualise content knowledge to teachers. This would be particularly helpful for non-specialist teachers wanting to understand fundamental debugging concepts.

Overall, relatively little research has been conducted into teachers' experiences of debugging in the classroom. Studies on teacher-student debugging interactions indicate that some teachers are aware of a variety of teaching approaches that can be applied to debugging. However, several of these studies do not take place in the classroom, where teachers' time for interactions with students may be severely limited (Michaeli and Romeike, 2019a). For less experienced or confident teachers, debugging-specific PD is a helpful way of developing teachers' debugging PCK, which itself is a concept that could be further developed.

3. Method

To explore both teachers' experiences and perspectives with teaching and supporting students with debugging in the classroom, we conducted semi-structured interviews with nine teachers.

3.1 Participants

Teachers were purposively sampled for the study, with the criteria that they must be teaching a text-based programming language at a lower secondary level¹. Eligible teachers were invited to partake in the study through a teacher research network newsletter, hence incorporating an element of convenience sampling.

Nine teachers took part in the study. Seven of these were from England, where the average class size for state-funded schools is 26 students (Department for Education, 2025), and

¹By lower secondary, we refer to the first three years of secondary education, known as key stage three in England, equivalent to grades 6-8 and ages 11-14.

two were from other countries. Each of these teacher² taught computing as a compulsory subject to their lower secondary students. Teachers from state-funded and nonEnglish schools followed the English computing curriculum, which mandates the “use

two or more programming languages, at least one of which is textual” (Department for Education, 2013b, p. 2). Teachers from private or independent schools were able to implement their own programme of study. Python was taught by all teachers, with some also teaching other TBPLs such as Swift. Prior to learning these TBPLs, their students would have been taught a BBPL in primary school (Department for Education, 2013a). Each teacher received a bookshop voucher worth £10 as a token of appreciation for participating. Table 1 reports the demographic information of the teachers interviewed.

Table 1
Demographic Information For Participating Teachers

Characteristic	Demographic Breakdown
Gender	Male: 4 Female: 5
Age	30-39: 2 40-49: 3 50-59: 3
Years of computing teaching experience	0-5: 1 6-10: 2 11-15: 3 16-20: 2 Over 20 years: 1
School information	Mixed-gender: 8 State-funded: 5 Private or independent: 2 Grammar: 2

3.2 Procedure and Data Collection

We conducted semi-structured interviews to allow teachers ample time and flexibility to discuss matters relevant to their experiences of debugging in the classroom. Moreover, as per our research questions, we were particularly interested in teachers’ *reported* feelings and experiences, which suited interviews well.

Interviews took place virtually from October to December 2023, using Google Meet or Microsoft Teams. Each teacher was interviewed individually, apart from Chloe and Deborah, who were interviewed together. Before being interviewed, each teacher completed a consent form and demographic questionnaire, in accordance with the study procedure approved by our department’s ethics committee. Interviews lasted between 35-65 minutes.

The interviews facilitated two main points of discussion. Firstly, teachers’ *own experiences and perspectives* with debugging in the classroom. That is, their personal thoughts and recollections related to debugging in programming lessons. This could include teaching approaches they used, general feelings they have experienced about debugging, or beliefs around the role of debugging when teaching programming. Second, we asked teachers

²Apart from one teacher, who was between teaching jobs at the time of interviewing.

about their *observations of students' experiences* with debugging, including any common errors, behaviours, or reactions that they commonly see among students. The interview schedule, along with the codebook definitions and agreement procedure, is available to view in the study repository (Gale and Sentance, 2025). We did not have consent from the teachers to publish the transcripts.

Each interview was recorded and automatically transcribed. The transcripts were then manually read through in conjunction with the recording to correct any mistakes and remove any mentions of personal or school-specific information, while also being used for initial memoing. These transcripts were sent to teachers within two weeks of the interview, who were given two weeks to mention anything they wished to be omitted or corrected. For the teachers who took this opportunity, a final version of the transcript was sent to them for their records. After this period, transcripts were disassociated with the teachers' contact details.

3.3 Analysis

We analysed teachers' interviews using 'codebook' thematic analysis (TA) (Braun and Clarke, 2006, 2021). This is a 'school' of TA that maintains the core values of reflexive TA, in which the subjectivity and experiences of the researcher are key aspects of the analysis, while incorporating some pragmatic approaches that seek reproducible coding (Braun and Clarke, 2021). In practice, we conceptualised latent themes that identified meaning beyond teachers' responses, as discussed in Braun and Clarke (2019, 2023), while incorporating approaches to ensure the codebook was consistently applied across the interview data (see Section 3.4).

This began with multiple passes of the interview transcripts. The first of these, as previously mentioned, was used both for correcting the automatically generated transcripts and initial memoing. This was performed by the first author. The second author then read through the corrected transcripts and wrote initial memos.

After this, the interview transcripts were migrated to NVivo 12, which was used for the rest of the analysis. Each transcript was then read through again by the first author to inductively generate open descriptive codes and note further memos. Through further inspection of the interview data, open codes were inductively abstracted into higher-level latent themes and sub-themes that aimed to provide an interpretive, and hence subjective, account of the data. This process of theme generation involved repeatedly restructuring the codebook, redefining the themes and sub-themes, and rewriting the report of the codebook, encompassing steps 4-6 in Braun and Clarke (2006). This process was conducted by the first author with feedback from the second author after some of the iterations. Our approach was therefore more iterative than linear, which is considered a normal part of TA (Braun and Clarke, 2006).

The analysis culminated in a codebook containing six themes, each containing subthemes, some of which contained their own sub-themes. Only the lowest-level sub-themes were

used to code the data. The final commentary of the results aims to provide a rich description of all portions of the interviews that are relevant to our research questions.

Regarding positionality, the first author is a doctoral student with limited previous experience in conducting interviews and qualitative analysis. The regular feedback of the second author was hence important, who is an experienced computing education researcher with expertise in qualitative analysis.

3.4 Reliability and Validity

The content validity of the interview questions was strengthened through two procedures. First, the questions were iteratively refined by both authors to ensure they appropriately discussed teachers' experiences, perspectives, and observations of debugging. Second, a pilot interview was conducted with a teacher to assess the wording of the questions and the timings of the interview. This teacher was personally contacted by the first author and received the same bookshop voucher as the other participating teachers. Their transcript was not included in the analysis.

The reliability with which the codebook was applied was improved by seeking agreement between the authors' coding patterns without calculating an interrater reliability (IRR) coefficient. Campbell et al. (2013) call this *negotiated agreement*. We wanted to verify the consistency with which the first author coded the transcripts according to the codebook definitions, also known as *intracoder reliability* (O'Connor and Joffe, 2020). Seeking agreement by identifying disagreements in our coding was appropriate for this case (Campbell et al., 2013; McDonald et al., 2019). This was particularly important as the first author was inexperienced in qualitative analysis. Negotiated agreement hence improved the clarity of definitions and the consistency with which the first author coded.

Calculating an IRR coefficient was not required for the above process and was considered detrimental for our study. Our interview transcripts were long and conversational, including frequent tangents and interruptions. Naturally, this makes it difficult for coders to agree on units of meaning, which is required for calculating agreement with IRR coefficients. Combined with the complexity of the codebook, any IRR coefficient is likely to be positively skewed and unrepresentative of the reliability of the codebook (Campbell et al., 2013). Furthermore, there is little consensus on acceptable IRR coefficient values for semi-structured interview data, which undoubtedly differs from other qualitative data with different research objectives (Krippendorff, 2004; Campbell et al., 2013).

Agreement was hence sought as follows. As described in the previous section, the first author developed the codebook with regular input from the second author. Two interview transcripts (Gideon and Hannah) were then randomly selected to be coded by both authors, in line with the guidelines of O'Connor and Joffe (2020). For this, the second author was provided with the codebook, complete with definitions, and an uncoded version of the transcripts. Once we both coded the transcripts, disagreements in coding were identified using the coding stripes feature in NVivo. Disagreements and ambiguous sub-themes were

then discussed. Where necessary, revisions to the structure and definitions of the codebook were made.

3.5 Limitations

A limitation of our method was that we kept the final themes close to the research questions (i.e., closely related to debugging). This was done to avoid generating an overly large codebook, but in the process likely missed some information that contributes to teachers' experiences of debugging in the classroom.

Additionally, our sample was, within certain constraints, self-selecting. Although we were not seeking a generalisable population for an in-depth qualitative study, the participating teachers were likely engaged with computing education research to some degree. As a result, they may be more reflective or intentional about their teaching practice than most teachers, potentially affecting the themes we identified.

4. Results

Six themes related to teachers' experiences, perspectives, and observations of students were yielded from the interviews, which are:

- Barriers to successful debugging
- The emotional nature of debugging
- Reliance on the teacher
- Varying levels of scaffolding for helping students to debug
- Improving students' debugging ability
- Promoting a positive error culture

The codebook structure and code definitions are available to view in the study repository (Gale and Sentance, 2025). We now report each of the themes, with sub-themes reported in italics, supported by teachers' quotes³ from the interviews.

4.1 Barriers to successful debugging

A particularly widespread theme was the obstacles that impede students from successfully debugging their programs. These challenges related to schools' allotted time for teaching programming, students' attitudes, and the nature of the Python language.

A barrier outside of the teachers' control was the *lack of time for sufficient debugging practice*. Teachers were generally dissatisfied with the lesson time they had to teach programming, sometimes constrained to as little as six hours per year. As well as making any progress beyond basic programming constructs difficult, this has effects for debugging. Teachers simply do not have the lesson time for their students to frequently encounter common errors, nor teach effective debugging strategies. As succinctly put by Adam:

³All teachers have been given pseudonyms and all quotes have had fillers and contractions removed.

“Do I have frustrations teaching students how to debug? Yeah, we don’t have enough time to do it!” - Adam

Barriers related to the difficulties of learning to program (*programming-related*), some of which were specific to learning a TBPL, also make debugging difficult for students. The *difficulty of getting syntax exactly right and conceptual difficulties*, sometimes manifesting as logical errors, were both mentioned as struggles for students. Many teachers described the common errors students encounter, mostly related to minor incorrect formatting, illustrating the ‘syntax barrier’ (Denny et al., 2011) present for many beginner programmers. Esther makes the point that such syntax errors may not be minor ‘slips of the hand’, but a *“a lack of the deeper understanding of what the language is actually trying to do”*.

Some teachers also commented on the hard to interpret programming error messages in the Python environment students were using⁴. While they can be positively utilised (see Section 4.5), teachers expressed the unclear and intimidating nature of them for students, not helped by the red text they usually appear in. Hannah explains that the content of the error message is not pitched at a level that most students understand, in turn increasing the difficulty of resolving syntax or runtime errors.

“The error message that comes back is really quite cluttered ... It could just be, ‘here’s the line number’. And it’s not very in student speak, so it will say like, ‘identification error’.” - Hannah

Not all the barriers to successful debugging identified by teachers were external to the student. Students may also have unhelpful attitudes towards computing or programming which impede their ability or motivation to debug.

“Why don’t they like to debug? It’s because ... they don’t like coding, they just don’t like the lesson, they don’t like the subject ... they’ve never known the success because they’ve given up too early or they’ve been put off it.” - Gideon

Gideon highlights that students’ negative attitudes span from debugging as a skill to computing as a subject. He also alludes to a lack of success and motivation as drivers for these negative attitudes. The former relates to the ‘performance accomplishment’ determinant of self-efficacy (Bandura, 1978), which is derived from a learner’s initial experiences with a subject. For most students in England, initial experiences of programming come at primary school in block-based environments. However, unhelpful attitudes could also arise from students’ experiences of learning more generally, which Isabel believes to be the case.

“Maybe some of them do not expect themselves to finish successfully because ... of their previous experiences in, not just in computing, but across the board of other subjects.” - Isabel

The effect of these attitudes often manifests as a lack of motivation to debug their programs. This has consequences for the teacher, who has the challenge of encouraging their students to persevere with debugging.

⁴At the time of the interviews, the most recent stable version of Python was 3.12.

4.2 *The emotional nature of debugging*

In addition to attitudes, teachers discussed a range of emotional reactions they observe when their students are debugging. Observed reactions were generally physical or verbal, occurring at the point of encountering an error or after attempting to resolve an error. Visible reactions are not universal, however. Some teachers specifically pointed out the *reserved reactions of quieter students*.

Most observed reactions represented *negative emotion during debugging struggle*, conveying feelings such as frustration and disappointment. An interesting observation mentioned by several teachers is the emotional attachment students develop towards their programs and the distress this can cause when errors occur. For students, encountering errors in their own programs can be far more costly than debugging an erroneous program written by the teacher, leading to more dire emotional reactions. Felix describes this as potentially depressing for students.

“When they’re starting to write upwards of ten lines of code, you’re invested in getting it right and then if it fails, that can flip and that can be quite depressing.” - Felix

Repeatedly experiencing such negative emotions when debugging is difficult for students to deal with. Some students may therefore disengage from the debugging process if they become stuck too often.

“Once they’ve hit the brick wall, that’s kind of the green light to go, ‘I’m just gonna do anything’. [They] Sit there in silence, they’re not causing any commotion.”

- Gideon

Despite the lack of visible frustration, disengagement may indicate more severe emotional distress, as students may not be able to cope with any more errors in their programs. Alternatively, students may disengage due to the difficulty of the error. Either way, some students’ motivation to debug is reduced after repeatedly encountering errors. It can be difficult for busy teachers to notice these students who, as Gideon mentions, are *“not causing any commotion”*.

Not all the emotional reactions reported were negative. The elation when solving an error was mentioned by some teachers. These reactions were also acute, indicating feelings of intense joy and pride.

“That dopamine hit when you get it right . . .I’ve seen kids struggling and struggling and struggling and all of a sudden, the light goes on or something happens . . .they even shake with excitement.” - Felix

Felix’s observations allude to the physical enactment of students’ emotions, demonstrating the sheer joy that can be experienced from debugging. However, such positive reactions were rarely mentioned, perhaps due to the perseverance needed to experience them.

We also found some teachers alluded to the *resilience of motivated students*, who do not experience such polarising emotions when debugging. These students remain composed

and explore different solutions, not afraid to add errors to their programs. Hannah explains that most of her students have such resilience.

“They are all interested in finding a solution to a problem. And I think that that’s really important in having the resilience to bother to debug the program.” - Hannah

The resilience to debug, to a large extent, comes from the student. From Hannah’s perspective, this derives from students’ interest in programming. Other factors are undoubtedly at play, such as the debugging-specific teaching strategies teachers use (see Section 4.5), their work in normalising errors in the classroom (see Section 4.6), and students’ access to computers at home. Hannah teaches in a highly selective school, where high levels of interest are perhaps not as surprising. Generally, however, negative emotions and disengagement were the most widely discussed observations we identified in this theme.

4.3 *Reliance on the teacher*

A combination of the *barriers to successful debugging* and the *emotional nature of debugging* contributes to many students relying on the teacher for debugging support. This is typically embodied by students raising their hands when encountering errors. Often, there are enough hands up to make it unrealistic for the teacher to effectively help every student requesting support.

Several teachers discussed *the commonness of teacher reliance*. When asked about the atmosphere of her programming lessons, Esther simply stated “*chaotic . . . lots of hands up and down all the time*” and that her students’ problems are “almost always about errors in code”. Chloe similarly reports “*I think at least half the class will ask all the time*”; far beyond any teacher’s capacity to meaningfully assist every student in classes of 25-30 students (Yadav et al., 2016b; Department for Education, 2025).

Several teachers confidently posited that the cause of such reliance is a lack of student resilience or motivation, rather than the difficulty of the error they are debugging. Some students want their errors to be instantly resolved, which they cannot always do themselves. This creates a preference for waiting for teacher support over sustained independent effort.

“They want you to give them the answer (laughs). They don’t want to think too much. Only very few of them would want to think and figure out the errors.” - Deborah

Once a student raises their hand, teachers generally stated that students do not continue debugging until their teacher approaches them. This is naturally detrimental for the student, with emotional reactions likely to ensue. However, this reliance also has a toll on the teacher. Some teachers reported that “*programming lessons are really, really busy lessons*” (Chloe), which can “*erode their [teachers’] professional confidence*” (Adam). Esther talks about the obligation they feel to help every student, while Gideon discusses the challenge of not being able to do so.

“You’re always thinking about, ‘while their hands are up, they’re not doing anything or making progress’. So you sort of feel this pressure to go and help everybody.” - Esther

“This is the problem with debugging I suppose in general ... I’ll have my program and then I got 20 other programs, so I’m potentially debugging 25 programs in a lesson ... This is beyond the capability of the human.” - Gideon

The impact of this pressure on the teacher is multifaceted. As well as the “*pressure to go and help everybody*”, many teachers also feel a pressure to prove that their students have learnt something every lesson. This is very difficult when so many students’ progress depends on the teacher.

Some teachers accept that it is not possible to personally help every student in need of debugging support. Instead, teacher reliance must be carefully managed. As Adam states, “*there has to be a more scalable model than going around and solving individual people’s code problems*”. Consequently, some teachers utilise strategies that motivate teacher-independent debugging. These aim to provide students with the confidence to debug without the teacher by their side. For example, some teachers tell their students to consult other sources of information before approaching them. These may be their peers (*encouraging helpful collaboration*) or resources that teachers have curated (*signposting to resources for assistance*). The former does not just reduce the in-class burden on the teacher; explanations from peers “*just seem to go in better*” (Hannah). Alternatively, some teachers will simply refuse to help students until they have taken some effort to debug themselves, such as Hannah.

“If they ask me any questions and they haven’t tried to solve it themselves, I will just tell them, ‘you need to try it yourself first’. So there’s not like an easy solution of, ‘I’m lazy. I’m just gonna ask for the answer’.” - Hannah

As a result, the teachers’ role regarding debugging support becomes less diagnostic. In Isabel’s words, “*they do ask [for debugging support], we just redirect them*”. Students exposed to these strategies become accustomed to *teacher-independent* debugging instead of teacher-dependent debugging. Teachers who employ such approaches explained how this helps to develop resilience among students that is often lacking at the beginning of their programming journey.

“They get better at these things, but there’s just as many older students who make the same mistakes and they’re like ‘oh, that was a silly error’ rather than ‘that was a killer’.” - Benjamin

4.4 Varying levels of scaffolding for helping students to debug

There are of course times when teachers do need to provide individualised debugging support to students. This initiates a teacher-student interaction, where teachers mentioned using a range of scaffolding techniques. Some of these are generic teaching tactics applied to the context of debugging, while others are programming- or debugging-specific. The level of scaffolding varies across a continuum, from very little, placing the onus on the student, to near-complete scaffolding of the problem, placing the onus on the teacher. The amount a teacher applies to a particular interaction partially depends on the teacher’s

knowledge of their student(s), as well as practical factors such as remaining lesson time (Elliott et al., 2024).

When putting the onus on the student, some teachers simply ask students to share their thoughts about the error in question (*encouraging externalisation of ideas or asking students questions*). Both of these approaches encourage students to outwardly reflect on their understanding of the problem, which might be what is stopping them from successfully resolving errors.

“It’s that rubber ducking process, sometimes students don’t explain it to themselves, and when they go back and explain their problem they spot it themselves.” - Benjamin

The questions teachers mentioned asking students were naturally of varying specificity, which again shifts the level of scaffolding. Adam deliberately uses generic questions at the beginning of interactions with his students.

“You’re gonna have to tell me what do you want it to do, what does it do at the moment and what have you tried already?” - Adam

These questions are useful for specialist and non-specialist teachers alike, as they do not require expert knowledge or time spent comprehending students’ code. Furthermore, once students understand that they will be asked generic questions, they may be more inclined to debug independently of the teacher.

Heavier forms of scaffolding may be used if a student has been debugging for a sustained period of time, or if the error is difficult to localise. Some teachers may give students the solution or give students hints specific to the error(s) the student is struggling with.

“If they’re struggling to find it, I can see they’re getting frustrated, first thing I would say is, ‘have a look at that particular line’. This is assuming I know where the error is.” - Benjamin

4.5 Improving students’ debugging ability

As well as general scaffolding, most teachers reported using approaches designed to help students become more successful debuggers. A more general sub-theme of this theme is teaching *programming strategies to make debugging easier*. These are patterns of programming that prevent errors from accumulating or help to keep a student aware of the purpose of their program. Adam and Hannah, for example, encouraged an iterative process of small changes and testing.

“Every five minutes, they need a reminder, ‘you need to save and you need to debug’. And so that process of make one change, test it, make another change, test that, is really, really important.” - Hannah

Additionally, some teachers *guide students through debugging strategies*. This typically involves explicitly modelling the debugging process with an example or student program.

Live coding was a popular medium for this. Some teachers, such as Gideon, model to groups of students who are struggling with a particular error, in turn helping several students at once.

“Sometimes if I’m getting two or three students with the same error, then I’ll be like, ‘right come round the board. I need to demonstrate, I need to model’.” - Gideon

A few teachers were strong advocates of *utilising programming error messages* (PEMs). Although they are not always intuitively phrased, Adam and Felix viewed PEMs as valuable sources of immediate feedback.

“In Python, although it might look intimidating, that red line of text, actually that’s really helpful because it’s instant and it tells you where to start looking.” - Adam

This positive outlook on PEMs drives these teachers to frequently discuss them. Some teachers also mentioned teaching students how to ‘decode’ PEMs.

Through repeated exposure and practice, guided by the strategies they teach, teachers reported a positive change in their students’ debugging behaviours. Benjamin’s students, who are *“pre-cued to look for common errors”*, become drilled in doing this themselves.

“They get better at looking for the obvious errors ... the top five syntax errors, they will look for those straight away as they have the practice.” - Benjamin

This positive change appears to be more frequently reported by the more confident or experienced teachers we interviewed. Regardless, such improvement is to be celebrated. These teachers show how their teaching strategies help students to go about debugging with independence and confidence. With this comes more of the aforementioned *elation of solving an error*, as well as positive initial experiences for students learning to program.

4.6 Promoting a positive error culture

Students’ improvement in debugging is not solely due to instructional strategies. Cultivating a positive error culture, an environment where students feel comfortable making errors in their programs, also helps students become more resilient when debugging. One useful method for promoting a positive error culture is *getting used to erroneous code*. This may involve sabotaging activities (Benjamin and Hannah), where students actively add errors to their peers’ programs, or giving students erroneous code (Chloe and Gideon). Hannah points out that sabotaging is useful because it exposes students to errors in a depersonalised format, reducing the potential emotional angst around debugging.

“It doesn’t matter if it doesn’t work. It’s not yours. And so that was a really nice way to kind of break down the anxiety about that ‘there will be errors’.” - Hannah

Another method of normalising errors in the classroom is by highlighting that errors are to be expected rather than dreaded. Felix explains how ‘failures’ should not just be reported, but celebrated to remove the stigma around them.

“Spectacular failures. I want those reported and celebrated as well. If something

should have gone right and [went] badly wrong but somebody found something interesting on the way ... celebrate it. Take the fear out of it." - Felix

Felix's response also highlights the practice of celebrating the effort as well as the outcome. If students are actively rewarded for their sustained debugging efforts, as in Adam's case, this can act as a further motivator for teacher-independent debugging.

"If we have a culture in a classroom where students know, 'ok, sir's not just going to give me a reward if I get things right ... he's gonna give me a reward if I try and get it right' and that's the big breakthrough for debugging." - Adam

These approaches, combined with verbal encouragement throughout the debugging process, help to create an environment where both encountering and struggling with errors are natural experiences. Such measures help make debugging a healthy, normal activity that can be a source of joy rather than despair.

5. Discussion

Teachers' experiences of debugging in the classroom spanned a range of insights and issues. A major motif we found among these experiences was the common reliance students have on their teacher for debugging support. This tended to be the case for less motivated students, who would raise their hands at the point of or soon after encountering an error.

Our findings, along with previous reports of teachers' experiences with debugging in the classroom, led us to conceptualise the hands-up problem, the in-class difficulty deriving from the common student tendency to raise their hand when struggling to complete a task. Although our definition frames *the hands-up problem* as a potential challenge for teachers of any subject, we focus on programming classrooms due to the lack of similar studies with other subject teachers. Therefore, in this paper, 'struggling to complete a task' refers to struggling to debug when learning to program.

We now introduce the hands-up problem as a challenge that computing teachers face in their programming lessons as a result of students struggling with debugging. We also propose ways to alleviate the problem. In addition to prior work, we reference theme names from the results section in italics where appropriate.

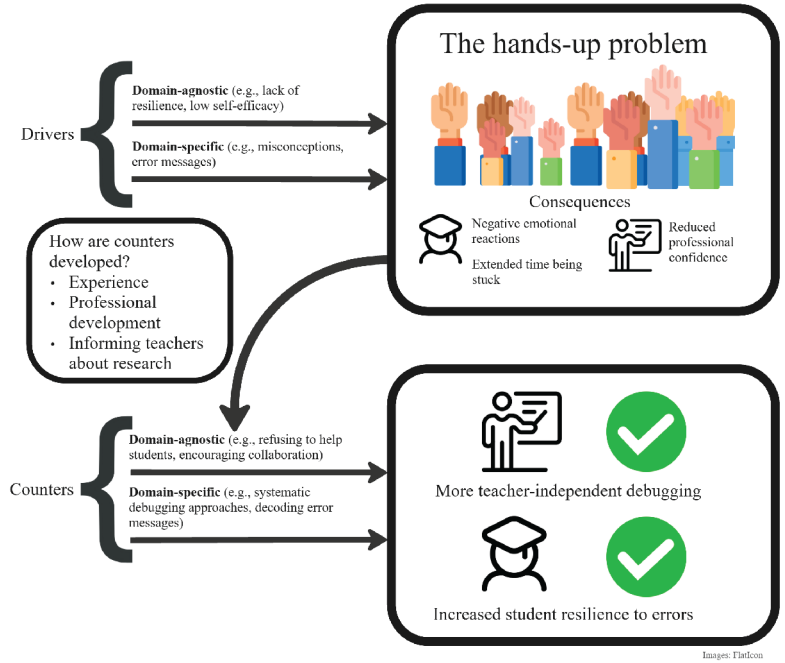


Figure 1. A Visualisation of the Hands-Up Problem

5.1 The Hands-Up Problem

The hands-up problem is the in-class difficulty deriving from the common student tendency to raise their hand when struggling to complete a task. Since the need for teacher support is reportedly common (*see the commonness of teacher reliance* and Bagge (2015); Michaeli and Romeike (2019a); Uppal (2020); Kerslake (2024)), the problem derives from many students raising their hands at once. Therefore, the severity of the problem scales with the size of the classroom, making it difficult for school teachers whose classes contain 25-30 students (Yadav et al., 2016b; Department for Education, 2025).

According to the teachers we interviewed, students generally raised their hands at the point of encountering an error. Hand-raising may also occur after students have put some initial effort into debugging (Michaeli and Romeike, 2019a; Wachter and Michaeli, 2025). Either way, the consequences for teachers and students are generally negative. We now describe some drivers, consequences, and counters to the hands-up problem, visualised in Figure 1.

5.1.1 The Drivers

There is a range of drivers that prompt a student to raise their hands during debugging struggle. These may be *domain-agnostic* or *domain-specific*. The list below is not meant to be exhaustive, but gives a general categorisation of some factors at play.

Domain-agnostic drivers for hand-raising are factors that can contribute to the handsup problem in any subject. We argue that these are generally related to students' affect. Teachers in our study often discussed a general lack of resilience to work through a problem, leading to reliance on the teacher. This problem is also highlighted by teachers in related work (Sentance and Csizmadia, 2017; Wachter and Michaeli, 2025) and likely relates to students' motivation levels. However, the very act of hand-raising must indicate some motivation to progress, otherwise students would disengage (see *negative emotion during debugging struggle*). Therefore, the drivers at play here may be a lack of motivation to problem-solve independently of the teacher, or attitudes of low self-efficacy.

Several *domain-specific* drivers for debugging may also contribute to the hands-up problem. For example, students learning a block- or text-based programming language may have misconceptions (see *conceptual difficulties* and (Sorva, 2021)) that make errors difficult to spot without the help of a more knowledgeable other. Teachers in our study also discussed drivers specific to learning a TBPL, such as the *difficulty of getting syntax exactly right and difficult to interpret PEMs*. Again, previous work has repeatedly identified these as difficulties with debugging in a TBPL (Denny et al., 2011, 2012; Becker et al., 2019). We suggest how these could be alleviated in Section 5.1.3.

Generally, domain-specific drivers can be viewed through cognitive load theory (Sweller, 1988; Sweller et al., 2019). If a student's working memory is near-saturated when programming, errors and difficulty with problem-solving follow (Sweller, 1988). Hand-raising is thus a logical behaviour, as students lack the cognitive capacity to debug themselves. However, if many students are experiencing a high load on working memory, the hands-up problem quickly follows.

5.1.2 The Consequences

Most teachers in our study alluded to many students raising their hands for debugging support, particularly when first learning to program. This is why we call it the *hands-up problem*; in typical classroom settings, teachers simply cannot go around and effectively support every student (Yadav et al., 2016b; Michaeli and Romeike, 2019a). This has both *student-side* and *teacher-side* consequences.

The *student-side consequences* are multi-faceted. Once students raise their hands, teachers in our study explained how unlikely it was for students to progress until they came to help. Given the drivers discussed, this is not surprising, as students may not have the working memory to continue their debugging efforts. However, this means some students will linger in a state of being stuck or emotional angst for an extended period. Even when a teacher does come to provide support, the teacher-student interaction will likely be rushed if there is a time pressure or other students with their hands up (Michaeli and Romeike, 2019a; Elliott et al., 2024). Such a classroom atmosphere makes it very difficult for teachers to meaningfully guide students through the errors they are struggling with.

The emotional consequences of debugging are also important to consider. Feelings of frustration, anxiety, and disappointment are documented in this study (*the emotional nature of*

learning to debug) and previous work (Kinnunen and Simon, 2010; Gorson et al., 2022; Ryan and DeLiema, 2023). If these feelings occur repeatedly, students may simply disengage in future encounters with errors (see *negative emotion during debuggingstruggle*). If experiences of emotional arousal and lack of success are students' first with learning to program, they can negatively form students' programming self-efficacy (Bandura, 1978; Gorson and O'Rourke, 2020). Such experiences are therefore important to limit, but hard to do so if students are heavily reliant on teacher support.

There are also *teacher-side* consequences of the hands-up problem. Some teachers described their programming lessons as 'chaotic' and 'busy', in part due to the number of hands raised. Some also reported feeling a duty to help every student requesting assistance. Failing to do so evoked feelings of inadequacy or disappointment (see *the toll on the teacher of this reliance*). Furthermore, if teachers are regularly assessed or inspected, a sea of raised hands may give the impression of an unproductive lesson. This can again evoke negative feelings in teachers, who may overcompensate by providing too much assistance to students. As Uppal states in their reflections on their in-class debugging practice, "*in the midst of a busy classroom where as a teacher I was multitasking on many levels, it was all too easy to over-scaffold and provide too much direction, to give the appearance of a seemingly successful lesson*" (Uppal, 2020, p. 15).

5.1.3 The Counters

So far, we have characterised the hands-up problem in fairly negative terms. However, it can be counteracted. Several teachers in our study report using teaching approaches that encourage teacher-independent debugging (see *motivating teacher-independent debugging* and related work (Bagge, 2015; Michaeli and Romeike, 2019a; Wachter and Michaeli, 2025)). We refer to these as *counters* to the hands-up problem. Like the drivers, the counters can be divided into *domain-agnostic* and *domain-specific*.

Domain-agnostic counters are strategies that teachers of any subject can apply to alleviate the hands-up problem. For instance, teachers may refuse to help students until they evidence some teacher-independent problem-solving (see *refusing to help students* and Bagge (2015); Michaeli and Romeike (2019a)). Some may also encourage collaboration with peers before providing support (see *encouraging helpful collaboration*).

There are also a variety of *domain-specific* counters that support students with the debugging process. Teachers reported referring students to appropriate resources and tools (see *signposting to resources for assistance* and Nixon et al. (2023); Ryan and DeLiema (2023); Elliott et al. (2024)) and modelling the debugging process to the whole class (see *guiding students through debugging strategies* and Ryan and DeLiema (2023)). Some also taught heuristics for 'decoding' PEMs due to their often-ambiguous nature (*utilising programming error messages*). An obvious benefit of these counters is that the whole class can benefit from them simultaneously. An attitudinal counter mentioned by some teachers is the enforcement of a positive error culture (see *promoting a positive error culture* and Hennessey Elliott et al. (2023)).

As well as the domain-specific counters that teachers mentioned, debugging-specific teaching approaches developed in prior work could be useful counters. Systematic debugging processes (e.g., (Carver and Risinger, 1987; Michaeli and Romeike, 2019b)) or debugging-specific tooling (e.g., (Guo, 2013; Luxton-Reilly et al., 2018; Ziemann and Reu 2024)) may provide students with effective debugging strategies to employ when encountering errors. Additionally, features of the programming environment students use could alleviate some of the drivers specific to TBPLs. For example, frame-based environments allow students to write text-based programs while reducing the aforementioned syntax barrier (Kölling et al., 2015; Weill-Tessier et al., 2024), and more informative PEMs (e.g., (Becker, 2016; Denny et al., 2020; Phung et al., 2023)) would help reduce the dependency on the teacher for feedback. While the curriculum taught by teachers in our study mandates the teaching of a TBPL, secondary school teachers could instead teach a BBPL as a counter in itself, though this may come at the cost of perceived authenticity (Weintrop and Wilensky, 2015).

These strategies may take time for students to become accustomed to (Bagge, 2015). However, the teachers who report consistently using these counters also report their students becoming more empowered to attempt teacher-independent debugging (Michaeli and Romeike, 2019a). Ultimately, this gives the teacher more time to provide more meaningful support to a smaller number of struggling students.

Among the teachers we interviewed, those who reported employing counters tended to be more confident and had decades of experience. This was less true for the less confident teachers, leaving them more susceptible to the consequences of the hands-up problem. More generally, the prevalence of these counters in computing classrooms is unclear. Results from Sentance and Csizmadia (2017) suggest only a fraction of teachers have strategies for teaching debugging and resilience to errors, with a lack of time to teach programming probably a limiting factor.

5.2 *The Implications*

Although we only interviewed nine teachers, we expect our findings align with other teachers' experiences for two reasons. First, previous researchers and teachers have reported struggles in line with the hands-up problem (Bagge, 2015; Sentance and Csizmadia, 2017; Michaeli and Romeike, 2019a; Uppal, 2020; Kerlake, 2024). Second, the commonality of non-specialist computing teachers and reported lack of content knowledge by some (Yadav et al., 2016b; Sentance and Csizmadia, 2017) hints at a lack of debugging PCK necessary for effectively supporting students with debugging.

For teachers who report counters to the hands-up problem, experience or professional development seemed to be the reason. This begs the question: how else can counters to the hands-up problem be developed? Or, in other words, how can less experienced or confident teachers teach debugging more confidently? We did not identify any themes related to this in our analysis and thus only provide initial recommendations for future research.

One suggestion is through professional development. Some studies have already reported successful debugging-specific PD, which have been participatory in nature (Elliott et al., 2022; Michaeli and Romeike, 2022; Nixon et al., 2023). However, given that some teachers still report findings aligned with the hands-up problem, debugging-specific PD is either not widespread enough or not being effectively translated into practice. If the former is the case, debugging-related PD needs to be developed or expanded. This does not need to solely focus on debugging; more emphasis on debugging could be included in programming PD workshops, given the intertwined nature of these skills.

Similarly, informing teachers about pedagogical approaches for debugging developed by prior work has huge potential. Although long-term evidence for many debuggingspecific teaching approaches is lacking (Yang et al., 2024), the increasing amount of research in this area at least provides teachers with some suggestions. Some teachers would likely adapt these to their classroom practice (Michaeli and Romeike, 2022), which could provide further insight into the efficacy and applications of these approaches. How teachers can be informed about debugging-specific approaches in a way that encourages classroom adoption is an important research-to-practice consideration, but outside the scope of this article.

These suggestions relate to the improvement of teachers' 'debugging PCK'. Despite the challenges debugging presents in programming classrooms, we are not aware of a conceptualisation of debugging PCK in the research literature. A debugging PCK would inevitably have some commonalities with conceptualisations of programming PCK (Saeli et al., 2011), such as the motivation for teaching it and common student difficulties. It would also include debugging-specific pedagogy, techniques for effectively managing teacher-student interactions, and other aforementioned counters to the hands-up problem. Therefore, even if a teacher does not have time to explicitly teach the debugging process, a more developed debugging PCK would equip them to deal with the hands-up problem. Once the concept of a debugging PCK is more well-developed, it can be spread through professional development.

6. Conclusions and Further Work

This paper has reported secondary school teachers' experiences with the teaching and learning of debugging in their programming lessons, with a specific focus on textbased programming. Through thematically analysed semi-structured interviews, we report teachers' strategies for supporting students with debugging, their observations of how students engage with the debugging process, and some challenges they face related to debugging.

A particular trend among our themes was the common student reliance on the teacher when encountering and struggling to resolve errors. This corroborates with previous reports of computing teachers' classroom experiences (Bagge, 2015; Sentance and Csizmadia, 2017; Michaeli and Romeike, 2019a; Uppal, 2020; Kerslake, 2024). We conceptualised our results through the hands-up problem, the in-class difficulty deriving from the common student tendency to raise their hand when struggling to complete a task. There are several

drivers for the frequent hand-raising that contributes to the hands-up problem, which have generally negative *consequences*. Students are often left stuck and frustrated, and teachers flustered. Usually through experience, some teachers developed *counters* to the hands-up problem. These consisted of general classroom management techniques and debugging-specific teaching approaches. Over time, these counters help students to become less reliant on the teacher and more resilient to errors in their programs.

This paper presents the hands-up problem as an initial conceptualisation of a significant challenge teachers face in programming lessons as a result of student struggle with debugging. This conceptualisation should be developed through further studies of teachers' experiences with teaching programming in the classroom. Where previous studies have used interviews (Michaeli and Romeike, 2019a) and simulations of teacher-student interactions (Elliott et al., 2022; Nixon et al., 2023; Wachter and Michaeli, 2025), methods such as classroom observations would help to paint a more well-rounded picture of teachers' lived experiences. This work is increasingly important to conduct given the increasing number of school curricula mandating the teaching of text-based programming (e.g., (Department for Education, 2013b; British Columbia Ministry of Education, 2016; Sysło, 2020; Ripani and Vazquez-Brust, 2023; Skolverket, 2024)).

Developing a model for debugging PCK is another area of further work we recommend. This would be a useful way of synthesising research into beginner debugging education at a school level and beyond. A debugging PCK should include approaches developed by both researchers and teachers, which could be iteratively tested and refined in the classroom. This could then inform further work on debugging-specific professional development and pedagogy.

Finally, further investigation into pedagogical approaches that facilitate positive experiences with learning to debug should be conducted. This is already happening, but future research should investigate these approaches on a larger scale over longer periods of time to better understand their effectiveness (Yang et al., 2024). Supporting teachers' debugging pedagogies will help to cultivate classroom environments where errors are confidently handled rather than feared.

Acknowledgments

We are very grateful to all of the teachers we interviewed for sharing their invaluable experiences. Without their participation, this research would not have been possible.

The first author gives all praise to the Lord Jesus Christ for blessing him with the wisdom and opportunity to conduct this research.

References

- Bagge, P. (2015). Eight steps to promote problem solving and resilience and combat learnt helplessness in computing. <https://philbagge.blogspot.com/2015/02/eight-steps-to-promote-problem-solving.html>
- Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy* 1, 4 (1978), 139–161. doi:10.1016/0146-6402(78)90002-4
- Becker, B.A. (2016). An effective approach to enhancing compiler error messages. In SIGCSE 2016 -Proceedings of the 47th ACM Technical Symposium on Computing Science Education. Association for Computing Machinery, Inc, New York, NY, USA, 126–131. doi:10.1145/2839509.2844584
- Becker, B.A., Denny, P., Pettit, R., Bouchard, D., Bouvier, D.J., Harrington, B., Kamil, A., Karkare, A., McDonald, C., Osera, P.M., Pearce, J.L. & Prather, J. (2019). Compiler error messages considered unhelpful: The landscape of text-based programming error message research. Technical Report. Association for Computing Machinery, New York, NY. 177–210 pages. doi:10.1145/3344429.3372508
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. doi:10.1191/1478088706QP0630A
- Braun, V. & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (8 2019), 589–597. doi:10.1080/2159676X.2019.1628806
- Braun, V. & Clarke, V. (2021). Conceptual and Design Thinking for Thematic Analysis. *Qualitative Psychology* 9, 1 (5 2021), 3–26. doi:10.1037/QUP0000196
- Braun, V. & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International Journal of Transgender Health* 24, 1 (2023), 1–6. doi:10.1080/26895269.2022.2129597
- British Columbia Ministry of Education. (2016). Area of Learning: Applied Design, Skills & Technologies. Technical Report. 22/03/2024. https://curriculum.gov.bc.ca/sites/curriculum.gov.bc.ca/files/curriculum/adst/en_adst_k-9.pdf
- Campbell, J.L., Quincy, C., Osserman, J. & Pedersen, O.K. (2013). Coding In-depth Semistructured Interviews: Problems of Unitization and Intercoder Reliability and Agreement. *Sociological Methods and Research* 42, 3 (8 2013), 294–320. doi:10.1177/0049124113500475
- Carlson, J., Daehler, K.R., Alonzo, A.C., Barendsen, E., Berry, A., Borowski, A., Carpendale, J., Chan, K.K.H., Cooper, R., Friedrichsen, P., Gess-Newsome, J., Henze-Rietveld, I., Hume, A., Kirschner, S., Liepertz, S., Loughran, J., Mavhunga, E., Neumann, K., Nilsson, P., Park, S., Rollnick, M., Sickel, A., Schneider, R.M., Suh, J.K., Driel, J. & Wilson, C.D. (2019). The refined consensus model of pedagogical content knowledge in science education. *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (1 2019), 77–92. doi:10.1007/978-981-13-5898-2
- Carver, S.M & Risinger, S.C. (1987). Improving children's debugging skills. In *Empirical studies of programmers: second workshop*. Ablex Publishing Corp., Norwood, New Jersey, 147–171. <https://dl.acm.org/doi/10.5555/54968.54978>
- Collins, A., Brown, J.S. & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator* 15, 3 (1991), 1–18.
- Denny, P., Luxton-Reilly, A. & Tempero, E. (2012). All syntax errors are not equal. In *Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Sci-*

- ence Education. Association for Computing Machinery, New York, NY, USA, 75–80. doi:10.1145/2325296.2325318
- Denny, P., Luxton-Reilly, A., Tempero, E. & Hendrickx, J. (2011). Understanding the syntax barrier for novices. In ITiCSE'11 -Proceedings of the 16th Annual Conference on Innovation and Technology in Computer Science. Association for Computing Machinery, New York, NY, USA, 208–212. doi:10.1145/1999747.1999807
- Denny, P., Prather, J. & Becker, B.A. (2020). Error Message Readability and Novice Debugging Performance. In Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE. Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education, New York, NY, USA, 480–486. doi:10.1145/3341525.3387384
- Denny, P., Prather, J., Becker, B.A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B.N., Santos, E.A. & Sarsa, S. (2024). Computing Education in the Era of Generative AI. *Commun. ACM* 67, 2 (1 2024), 56–67. doi:10.1145/3624720
- Department for Education. (2013a). Computing programmes of study: key stages 1 and 2. 3 pages. [https://assets.publishing.service.gov.uk/media/5a7c576be5274a1b00423213/ PRIMARY_national_curriculum_-_Computing.pdf](https://assets.publishing.service.gov.uk/media/5a7c576be5274a1b00423213/PRIMARY_national_curriculum_-_Computing.pdf)
- Department for Education. (2013b). Computing programmes of study: key stages 3 and 4. 3 pages. [https://assets.publishing.service.gov.uk/media/5a7cb981ed915d682236228d/ SECONDARY_national_curriculum_-_Computing.pdf](https://assets.publishing.service.gov.uk/media/5a7cb981ed915d682236228d/SECONDARY_national_curriculum_-_Computing.pdf)
- Department for Education. (2025). Schools, pupils and their characteristics. <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics/2024-25#dataBlock-50ad5b91-16e6-4159-bc9e-4dac3be7bbae-tables>
- Du, H., Xing, W. & Zhang, Y. (2021). A Debugging Learning Trajectory for Text-Based Programming Learners. In Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 2. ACM, New York, NY, USA. doi:10.1145/3456565
- Elliott, C.H., Nixon, J., Bush, J.B., Chakarov, A.G. & Recker, M. (2022). "Do I need to know what I am doing if I am the teacher?" Developing teachers' debugging pedagogies with physical computing. In Proceedings of the 16th International Conference of the Learning Sciences. International Society of the Learning Sciences, 1565–1568. <https://repository.isls.org/handle/1/8959>
- Elliott, C.H., Nixon, J., Chakarov, A.G., Bush, J.B., Michael J Schneider & Recker, M. (2024). Characterizing Teacher Support of Debugging with Physical Computing: Debugging Pedagogies in Practice. *ACM Transactions on Computing Education* 24, 4 (12 2024). doi:10.1145/3677612
- Ettles, A., Luxton-Reilly, A. & Denny, P. (2018). Common Logic Errors Made By Novice Programmers. In ACM International Conference Proceeding Series. Association for Computing Machinery, 83–89. doi:10.1145/3160489.3160493
- Fitzgerald, S., Lewandowski, G., McCauley, R., Murphy, M., Simon, B., Thomas, L. & Zander, C. (2008). Debugging: finding, fixing and flailing, a multi-institutional study of novice debuggers. *Computer Science Education* 18, 2 (6 2008), 93–116. doi:10.1080/08993400802114508
- Flood, V.J., Deliema, D., Harrer, B.W. & Abrahamson, D. (2018). Enskilment in the Digital Age: The Interactional Work of Learning to Debug. In Proceedings of International Conference of the Learning Sciences, ICLS. ICLS, 1405–1406. <https://escholarship.org/uc/item/0ws6d6vp>
- Gale, L. & Sentance, S. (2025). Study Repository for "The Hands-Up Problem and How to Deal With It: Secondary School Teachers' Experiences of Debugging in the Classroom". (2025).

- doi:10.17863/CAM.123194 Accessed: 17-11-2025.
- Gorson, J., Cunningham, K. & Worsley, M. (2022). Using Electrodermal Activity Measurements to Understand Student Emotions While Programming. In Proceedings of the 2022 ACM Conference on International Computing Education Research -Volume 1. Association for Computing Machinery, New York, NY, USA, 105–119. doi:10.1145/3501385.3543981
- Gorson, J. & O'Rourke, E. (2020). Why do CS1 Students Think They're Bad at Programming?: Investigating Self-efficacy and Self-assessments at Three Universities. In Proceedings of the 2020 ACM Conference on International Computing Education Research. Association for Computing Machinery, New York, NY, USA, 170–181. doi:10.1145/3372782.3406273
- Guo, P.J. (2013). Online Python Tutor: Embeddable Web-Based Program Visualization for CS Education. In Proceeding of the 44th ACM Technical Symposium on Computer Science Education. Association for Computing Machinery, New York, NY, USA, 579–584. doi:10.1145/2445196.2445368
- Elliott, C.H., Chakarov, A.G., Bush, J.B., Nixon, J. & Recker, M. (2023). Toward a debugging pedagogy: helping students learn to get unstuck with physical computing systems. *Information and Learning Sciences* 124, 1/2 (6 2023), 1–24. doi:10.1108/ILS-03-2022-0051
- Hubwieser, P., Mühling, A., Magenheimer, J. & Ruf, A. (2013). Towards a conceptualization of pedagogical content knowledge for computer science. In Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/2493394.2493395
- Jadud, M.C. (2006). Methods and tools for exploring novice compilation behaviour. In ICER 2006 -Proceedings of the 2nd International Computing Education Research Workshop, Vol. 2006. Association for Computing Machinery, New York, NY, USA, 73–84. doi:10.1145/1151588.1151600
- Kerslake, C. (2024). Stump-the-Teacher: Using Student-generated Examples during Explicit Debugging Instruction. In Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1. Association for Computing Machinery, New York, NY, USA, 653–658. doi:10.1145/3626252.3630934
- Kinnunen, P. & Simon, B. (2010). Experiencing programming assignments in CS1: the emotional toll. In Proceedings of the Sixth International Workshop on Computing Education Research. Association for Computing Machinery, New York, 77–86. doi:10.1145/1839594.1839609
- Kölling, M., Brown, N.C.C. & Altmir, A. (2015). Frame-Based Editing: Easing the Transition from Blocks to Text-Based Programming. In ACM International Conference Proceeding Series, Vol. 0911-November-2015. Association for Computing Machinery, New York, NY, USA, 29–38. doi:10.1145/2818314.2818331
- Krippendorff, K. (2004). Reliability in Content Analysis: Some Common Misconceptions and Recommendations. *Human Communication Research* 30, 3 (7 2004), 411–433. doi:10.1111/J.1468-2958.2004.TB00738.X
- Liu, Z., Zhi, R., Hicks, A. & Barnes, T. (2017). Understanding problem solving behavior of 6–8 graders in a debugging game. *Computer Science Education* 27, 1 (2017), 1–29. doi:10.1080/08993408.2017.1308651
- Luxton-Reilly, A., McMillan, E., Stevenson, E., Tempero, E. & Denny, P. (2018). Ladebug: An online tool to help novice programmers improve their debugging skills. In Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education. Association for Computing Machinery, New York, NY, USA, 159–164. doi:10.1145/3197091.3197098

- McDonald, N., Schoenebeck, S. & McDonald, N. (2019). Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019), 1–23. doi:10.1145/3359174
- Michaeli, T. & Romeike, R. (2019a). Current Status and Perspectives of Debugging in the K12 Classroom: A Qualitative Study. In *2019 IEEE Global Engineering Education Conference (EDUCON)*. IEEE Computer Society, Washington, DC, 1030–1038. doi:10.1109/EDUCON.2019.8725282
- Michaeli, T. & Romeike, R. (2019b). Improving debugging skills in the classroom -The effects of teaching a systematic debugging process. In *Proceedings of the 14th Workshop in Primary and Secondary Computing Education*. Association for Computing Machinery, New York NY, USA, 1–7. doi:10.1145/3361721.3361724
- Michaeli, T. & Romeike, R. (2022). “I Now Feel that this is Unfair” A Case Study on the Effects of Professional Development for Debugging in the K-12 Classroom. In *Informatics in Schools. A Step Beyond Digital Education. ISSEP 2022.*, Andreas Bollin and Futschek Gerald (Eds.), Vol. 13488 LNCS. Springer International Publishing, Cham, 78–89. doi:10.1007/978-3-031-15851-3_{ }7
- Nixon, J., Elliott, C.H., Schneider, M., Bush, J.B., Bhaduri, S. & Recker, M. (2023). Teachers’ Learning to Support Students During Science Inquiry: Managing Student Uncertainty in a Debugging Context. In *Proceedings of the 17th International Conference of the Learning Sciences -ICLS 2023*. International Society of the Learning Sciences, 601–608. doi:10.22318/icls2023.247673
- O’Connor, C. & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods* 19 (1 2020). doi:10.1177/1609406919899220
- Pea, R.D. (1986). Language-Independent Conceptual “Bugs” in Novice Programming. *The Journal of Educational Computing Research* 2, 1 (1986), 25–36. doi:10.2190/689T-1R2A-X4W4-29J2
- Perkins, D. & Martin, F. (1986). Fragile knowledge and neglected strategies in novice programmers. In *Papers Presented at the First Workshop on Empirical Studies of Programmers on Empirical Studies of Programmers*. Educational Center for Technology, Ablex Publishing Corp., Cambridge, MA, 213–229. <https://dl.acm.org/doi/abs/10.5555/21842.28896>
- Perkins, D.N., Hancock, C., Hobbs, R., Martin, F. & Simmons, R. (1986). Conditions of Learning in Novice Programmers. *Journal of Educational Computing Research* 2, 1 (2 1986), 37–55. doi:10.2190/ gujt-jcbj-q6qu-q9pl
- Phung, T., Cambroner, J., Gulwani, S., Kohn, T., Majumdar, R., Singla, A. & Soares, G. (2023). Generating High-Precision Feedback for Programming Syntax Errors using Large Language Models. In *Proceedings of the 16th International Conference on Educational Data Mining*, Mingyu Feng, Tanjaand KÄöser & Partha Talukdar (Eds.). International Educational Data Mining Society, Bengaluru, India, 370–377. doi:10.5281/zenodo.8115652
- Rich, K.M., Strickland, C., Binkowski, T.A. & Franklin, D. (2019). A K-8 Debugging Learning Trajectory Derived from Research Literature. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. Association for Computing Machinery, New York, NY, USA, 745–751. doi:10.1145/3287324.3287396
- Ripani, M.F. & Vazquez-Brust, D. (2023). Programming and Coding Skills in Latin America. Technical Report. UNESCO, Accessed: 10/04/2025. 1–27 pages. <https://unesdoc.unesco.org/ark:/48223/pf0000386204>

- Royal Society. (2017). After the reboot: computing education in UK schools. Technical Report. Accessed: 07/04/2025. 1–116 pages. <https://royalsociety.org/~media/policy/projects/computing-education/computing-education-report.pdf>
- Ryan, Z.D. & DeLiema, D. (2023). Reflections on sustained debugging support: conjecture mapping as a point of departure for instructor feedback on design. *Instructional Science* 51, 6 (12 2023), 1043–1078. doi:10.1007/s11251-023-09629-5
- Saeli, M., Perrenet, J., Jochems, W.M.G. & Zwaneveld, B. (2011). Teaching Programming in Secondary School: A Pedagogical Content Knowledge Perspective. *Informatics in Education* 10, 1 (4 2011), 73–88. doi:10.15388/infedu.2011.06
- Sentance, S. & Csizmadia, A. (2017). Computing in the curriculum: Challenges and strategies from a teacher’s perspective. *Education and Information Technologies* 22, 2 (3 2017), 469–495. doi:10.1007/s10639-016-9482-0
- Shulman, L.S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher* 15, 2 (2 1986), 4–14. doi:10.3102/0013189X015002004
- Skolverket. (2024). Curriculum for Compulsory School, Preschool Class and School-Age Education. Technical Report. Skolverket, Accessed: 07/04/2025. <https://www.skolverket.se/getFile?file=13128>
- Sorva, J. (2021). Misconceptions and the Beginner Programmer. In *Computer Science Education* (2 ed.). Bloomsbury Academic, Chapter 20, 171–188. doi:10.5040/9781350057142.ch-013
- Sun, C., Yang, S. & Becker, B. (2024). Debugging in Computational Thinking: A Meta-analysis on the Effects of Interventions on Debugging Skills. *Journal of Educational Computing Research* 62, 4 (1 2024). doi:10.1177/07356331241227793
- Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science* 12, 2 (4 1988), 257–285. doi:10.1016/0364-0213(88)90023-7
- Sweller, J., Merriënboer, J.J.G. & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review* 31, 2 (6 2019), 261–292. doi:10.1007/s10648-019-09465-5
- Sysło, M.M. (2020). The new computer science curriculum in Poland-challenges and solutions. Technical Report. Warsaw School of Computer Science, Accessed: 10/04/2025. <https://ceur-ws.org/Vol-2755/paper10.pdf>
- Tsan, J., Weintrop, D. & Franklin, D. (2022). An Analysis of Middle Grade Teachers’ Debugging Pedagogical Content Knowledge. In *Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education Vol. 1*. Association for Computing Machinery, New York, NY, USA, 533–539. doi:10.1145/3502718.3524770
- Uppal, G. (2020). Exploring and Understanding Pupils’ Lack of Perseverance and Autonomy With Debugging in Computing. *Research in Teacher Education* 10, 1 (2020), 12–16. doi:10.15123/ucl.88z8q
- Wachter, H. & Michaeli, T. (2025). Analyzing Teachers’ Diagnostic and Intervention Processes in Debugging Using Video Vignettes. In *Informatics in Schools. Innovative Approaches to Computer Science Teaching and Learning*, Vol. 15228 LNCS. Springer, Cham, 167–179. doi:10.1007/978-3-031-73474-8_{ }13
- Weill-Tessier, P., Brown, N.C.C. & Kölling, M. (2024). Strype: a Versatile Python-like Programming Environment for Novices. In *Proceedings of the 19th WiPSCE Conference on Primary and Sec-*

- ondary Computing Education Research. Association for Computing Machinery, New York, NY, USA, 1–2. doi:10.1145/3677619.3678132
- Weintrop, D. & Wilensky, U. (2015). To Block or not to Block, That is the Question: Students' Perceptions of Blocks-based Programming. In Proceedings of the 14th International Conference on Interaction Design and Children. Association for Computing Machinery, New York, NY, USA, 199–208. doi:10.1145/2771839.2771860
- Whalley, J., Settle, A. & Luxton-Reilly, A. (2023). A Think-Aloud Study of Novice Debugging. *ACM Transactions on Computing Education* 23, 2 (6 2023), 1–38. doi:10.1145/3589004
- Yadav, A., Berges, M., Sands, P. & Good, J. (2016a). Measuring computer science pedagogical content knowledge: An exploratory analysis of teaching vignettes to measure teacher knowledge. In Proceedings of the 11th Workshop in Primary and Secondary Computing Education. Association for Computing Machinery, New York, NY, USA, 92–95. doi:10.1145/2978249.2978264
- Yadav, A., Gretter, S., Hambrusch, S. & Sands, P. (2016b). Expanding computer science education in schools: understanding teacher experiences and challenges. *Computer Science Education* 26, 4 (2 2016), 235–254. doi:10.1080/08993408.2016.1257418
- Yan, W., Luo, F., Israel, M., Liu, R. & Chandler, L.T. (2025). How Do Elementary Students Apply Debugging Strategies in a Block-Based Programming Environment? *Education Sciences* 2025, Vol. 15, Page 292 15, 3 (2 2025), 292. doi:10.3390/EDUCSCI15030292
- Yang, S., Baird, M., O'Rourke, E., Brennan, K. & Schneider, B. (2024). Decoding Debugging Instruction: A Systematic Literature Review of Debugging Interventions. *ACM Transactions on Computing Education* 24, 4 (2024), 45. doi:10.1145/3690652
- Ziemann, F. & Reu, F. (2024). Systematic Debugging of Logical Errors in Source Code. In Proceedings of the 19th WiPSCE Conference on Primary and Secondary Computing Education Research. Association for Computing Machinery, New York, NY, USA, 1–4. doi:10.1145/3677619.3678105