

Students' Experiences with Generative AI for Programming Tasks and Independent Problem-Solving

A Qualitative Study of Computer Science Students

Diman Uzunov[1]

Supervisors: Dr. Martin Skrodzki, Dr. Xiaoqi Feng, Dr. Rafael Bidarra

[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
21 June 2026

Name of the student: Diman Uzunov
Final project course: CSE3000 Research Project
Thesis committee: Dr. Martin Skrodzki, Dr. Xiaoqi Feng, Dr. Rafael Bidarra

Abstract

Generative AI (GenAI) tools such as ChatGPT have become widely used by computer science students for code generation, debugging, and concept clarification. While these tools can improve efficiency in programming tasks, concerns remain about their potential impact on the independent problem-solving skills that computer science education seeks to cultivate. This study investigates how the use of GenAI during programming tasks shapes the way Computer Science students engage with independent problem-solving. Drawing on semi-structured interviews with seven Computer Science students at TU Delft, Through reflexive thematic analysis, the study identifies two themes. First, students framed their reliance on GenAI as a deliberate, self-governed choice rather than a loss of ability: they kept the interpretive work of understanding and planning, and the judgment of whether a solution was correct, for themselves, while delegating execution and The filling of knowledge gaps, and maintained that they could still work unaided if they chose to. Second, students reported a weakening willingness to persist through difficulty without help, both an immediate lowered threshold to begin effortful tasks and a gradual decline of motivation over time. Mapped onto Pólya's four stages of problem-solving, engagement is retained at the early and evaluative stages but surrendered in the middle, Persistence governs whether students return to a problem after a failed attempt. These findings suggest that GenAI reshapes independent problem-solving less by removing ability than by lowering the effort required to avoid it, and contribute to ongoing discussions about the responsible integration of AI tools in computer science education.

Contents

1 Introduction **3**

2 Literature Review **3**
2.1 Independent Problem-Solving in Programming 3
2.2 GenAI as a Programming Aid 4
2.3 Between Scaffold and Crutch 4
2.4 Research Gap 4

3 Methodology **5**
3.1 Participants and Sampling 5
3.2 Data Collection 5
3.3 Data Analysis 5
3.4 Researcher Positionality 5
3.5 Limitations 6

4 Results **6**
4.1 GenAI Reliance as a Deliberate Choice 6
4.2 Reduced Willingness to Persist Through Difficult Tasks 8

5 Discussion **9**
5.1 Future Work 10

6 Conclusions **10**

7 Responsible Research **10**
7.1 Ethical Treatment of Participants 10
7.2 Researcher Reflexivity and Bias 10
7.3 Transparency and Reproducibility 11
7.4 Data Management 11
7.5 Responsible Use of GenAI in the Research 11

A Interview Guide **12**

1 Introduction

The emergence of Generative Artificial Intelligence (GenAI) tools, particularly large language models such as ChatGPT, is rapidly reshaping computer science education across higher education institutions. GenAI helps students complete tasks faster, but programming education also aims to develop reasoning skills that may be affected when such tasks are completed with reduced cognitive effort. These systems provide programming support through code generation, debugging assistance, and natural-language explanations, allowing students to complete programming tasks more efficiently. As these tools become more accessible, they are increasingly used not only as productivity aids but also as learning companions during programming activities (Prather et al., 2023). Research on GenAI in higher education has highlighted both its benefits and its risks (Jafari & Keykha, 2024). While GenAI tools may improve short-term task performance, excessive reliance on them could reduce students' engagement in analytical reasoning, debugging, and independent problem-solving, which are fundamental to programming education. Aruleba et al. (2025) shows that while GenAI tools can significantly improve task performance, they may negatively impact conceptual understanding when overused. Although existing studies have explored student engagement and prompting strategies surrounding GenAI use, there remains a limited understanding of how GenAI reshapes students' engagement at each specific stage of independent problem-solving, and whether growing reliance reflects lost ability or a shift in motivation. Much of the current literature focuses on either performance outcomes or general patterns of tool use (Jafari & Keykha, 2024; Prather et al., 2023), leaving a gap in understanding how perceived dependency on GenAI relates to students' problem-solving engagement and what conditions lead them to turn to GenAI rather than reason independently. This study seeks to address this gap by investigating students' experiences, strategies, and perceptions of reliance on GenAI for programming tasks. In doing so, it contributes to ongoing discussions on the responsible and pedagogically effective integration of these systems in computer science education (Nnamdi et al., 2026; Prather et al., 2023). In the context of this study, independent problem-solving ability refers to a student's capacity to analyse and understand programming assignments, develop a strategy to solve them, and implement solutions without direct AI-generated assistance. Therefore, this study investigates the following research question:

> ***RQ1:*** *How does the use of GenAI during programming tasks shape the How do Computer Science students engage with independent problem-solving?*

2 Literature Review

This section explains the concepts and prior research that frame the study. It first describes what is meant by independent problem-solving in programming, then reviews how GenAI tools are reshaping that process, and finally identifies the gap that motivates the research question.

2.1 Independent Problem-Solving in Programming

A useful way to describe independent problem-solving is Pólya's classic account of how problems are solved (Pólya, 1945). Pólya breaks the process into four stages: understanding the problem, devising a plan, carrying out the plan, and Looking back to check the result. Although Pólya's framework originates in mathematics education and dates from 1945, it has been widely applied in programming education research (Prather et al., 2023) and remains a useful lens for mapping the cognitive work involved in solving programming problems. Learning to code is widely understood as learning to carry out this kind of reasoning, not simply to produce working code. Engaging with each stage independently, rather than skipping or delegating steps, is what makes this reasoning educationally significant (Loksa et al., 2016): the value lies not only in reaching correct answers, but in the metacognitive work it requires, since Students who reason through a problem without external help build an awareness of their own understanding, which is hard to develop when a tool removes the need to reason (Zimmerman, 2002). More recent work confirms the framework remains a relevant diagnostic lens: Rahe and Maalej (2025) observed that Students using GenAI tend to skip early reasoning stages, such as problem understanding, handing these directly to the tool, which suggests which stages students engage with independently is a meaningful question,

even as new tools enter the process. While programming-specific models add further stages of self-regulation (Loksa et al., 2016), Pólya's four-stage framework is adopted here because it maps more directly onto the broad phases participants described during interviews.

2.2 GenAI as a Programming Aid

GenAI tools support programming through code generation, debugging assistance, and natural-language explanation, and students have adopted them quickly. In a large multi-country survey, Prather et al. (2023) found that students increasingly turn to GenAI for help, in some cases more than to peers or online forums. Reviews of AI in higher education point to clear benefits, including personalised learning, individual feedback, and reduced workload (Jafari & Keykha, 2024). What makes GenAI different from earlier aids is its reach across the whole problem-solving cycle: a single tool can now interpret a prompt, propose a plan, generate an implementation, and explain an error message. In Pólya's terms, it can take part in every stage rather than supporting just one. Studies also report growing adoption of GenAI among computing and engineering students, alongside concerns about overreliance and the increasing acceptance of AI-generated solutions (Alpizar-Chacon et al., 2025; Ovi et al., 2025). While students generally value GenAI for improving efficiency and understanding, researchers continue to question how sustained reliance may affect learning and independent reasoning. (Ovi et al., 2025). This capacity to take over every stage raises a question about why students delegate work they could do themselves. Research on cognitive offloading describes how people hand cognitive tasks to external aids to reduce mental effort, and crucially shows that the decision to offload is driven by judgements about one's own ability that are not always accurate (Risko & Gilbert, 2016). Applied to programming, this suggests that reliance on GenAI may reflect a deliberate choice to conserve effort rather than an inability to reason, a distinction this study returns to in its analysis.

2.3 Between Scaffold and Crutch

The terms scaffold and crutch have specific pedagogical meanings. Scaffolding was originally defined as support that lets a learner accomplish what they could not manage alone, with the essential feature that it is gradually withdrawn as competence grows (Wood et al., 1976). Support that never fades, by contrast, can leave the learner unable to perform without it. GenAI sharpens this distinction: a tool available at every moment may scaffold reasoning or quietly replace it. This reach is also the source of concern. A tool that can perform every stage of problem-solving may either support a student's reasoning or replace it. Aruleba et al. (2025) captures this directly, describing ChatGPT's "dual role" as both a helpful scaffold and a potential crutch: in their study, some students grew more critical and self-directed over time, while a minority remained passively dependent on generated answers. Reviews voice the same worry at a broader level, warning that AI can lead students to "replace thinking" when its output is over-trusted (Jafari & Keykha, 2024). Work on task design responds to this by trying to keep reasoning with the student: Nnamdi et al. (2026) propose a framework in which GenAI assembles existing knowledge and supports coding, but does not provide finished answers, and in which periods of independent problem-solving without AI are built into the course.

2.4 Research Gap

Taken together, these studies show what GenAI tools can do, how students use them, and that overreliance is a common concern. What remains less examined is how, from the student's own perspective, GenAI use reshapes their engagement with independent problem-solving. Two questions in particular are underexplored. First, when students lean on GenAI, does this reflect a genuine loss of problem-solving ability, or a deliberate choice to delegate work they could still do themselves? Second, what happens to their willingness to persist through difficulty when an easier path is always available? This study addresses these questions through the interview-based design described in the next section.

3 Methodology

This study uses a qualitative research design based on individual semi-structured interviews, in order to understand how undergraduate Computer Science students perceive their use of GenAI tools during programming tasks and how this shapes their engagement with independent problem-solving. A qualitative approach is well-suited to this aim, as it allows a rich exploration of students' experiences and reasoning (Braun & Clarke, 2006).

3.1 Participants and Sampling

Participants were Computer Science students at TU Delft, predominantly at the Bachelor's level. A purposive sampling strategy was used to recruit students who already had experience using GenAI tools for their studies, so they could speak directly to the experiences under investigation. Seven students (n = 7) took part, predominantly male, most in their second or third year of Bachelor's studies, with one master's student.

3.2 Data Collection

Data were collected through individual semi-structured interviews, conducted either in person or online and lasting approximately 45 minutes each. With participants' Consent, interviews were audio-recorded and transcribed verbatim. The interview followed a semi-structured guide organised into four thematic areas: **(1) General patterns of GenAI use**; **(2) Interaction strategies and how these had evolved with experience; (3) Perceived cognitive impact of GenAI on independent problem-solving;** and **(4) Institutional and ethical considerations.** Each area contained a small number of open, non-leading questions, supported by follow-up probes (for example, "what do you mean by that?") and deliberate pauses to encourage participants to elaborate in their own words. The full guide is provided in Appendix A. While several questions invited participants to reflect on different points of their problem-solving process, how they approach a task, plan a strategy, implement it, and evaluate the result when working with GenAI, the guide itself was not structured around Pólya's stages. The framework was introduced only later, during analysis, as an organising lens for interpreting the themes (see Section 3.4); The guide was organised thematically.

3.3 Data Analysis

The data analysis followed Braun and Clarke's reflexive thematic analysis (RTA) as described in (Braun & Clarke, 2022). Coding was primarily inductive, with codes created from participants' accounts, while Pólya's four-stage problem-solving framework provided a deductive orienting structure for organising and interpreting the resulting themes. Reflexive thematic analysis was selected because it supports an interpretive exploration of participants' experiences and allows themes to be developed through researcher engagement with the data (Braun & Clarke, 2022). The analysis moved through the six phases of RTA Braun and Clarke, 2022. Transcripts were read repeatedly for familiarisation, then coded openly in ATLAS.ti to generate 47 codes; related codes were combined into candidate themes, which were reviewed against both the coded extracts and the full dataset, then refined and named, before being written up in relation to the research questions. This resulted in two primary themes.

3.4 Researcher Positionality

I conducted the study as an undergraduate Computer Science student at the same institution as the participants and as a user of GenAI tools. This insider position aided understanding of the educational context but also risked shaping interpretation. To address this, coding remained close to participants' own language during the inductive phase, and Pólya's framework was introduced only after initial coding. Findings are therefore presented as participants' perceptions rather than objective measures of behaviour.

3.5 Limitations

Several limitations qualify these findings. First, the study relies on self-reported perceptions rather than direct observation, so the results reflect how students *perceive* their engagement with independent problem-solving, which may diverge from their actual practice; the recurring claim that they "could still do it" was not, and could not be, tested. Second, the sample is small and homogeneous: seven students from a single institution, predominantly male, and concentrated in the second and third years of one Bachelor's programme, which limits how far the patterns can be generalised to other cohorts, disciplines, or levels of study. Third, as an undergraduate at the same institution who uses these tools myself, I shaped both the interviews and their interpretation; in particular, the central distinction between using and depending on GenAI was at times introduced in my own follow-up questions and then affirmed by participants rather than volunteered unprompted. Where a framing originated with the interviewer, it is treated as weaker evidence, and the analysis foregrounds volunteered accounts accordingly. This interviewer influence is discussed further in Section 3.4.

4 Results

The thematic analysis produced two main themes describing how GenAI shapes students' engagement with independent problem-solving. The first concerns how students frame their reliance: as a deliberate choice rather than a loss of ability (Section 4.1). The second concerns what students report changing: a gradual decline in their willingness to persist through difficult problems without help (Section 4.2). Together, these themes describe two axes of the same self-perception: ability is reported as intact while willingness erodes. In short, students describe being able to reason independently but increasingly choosing not to. To preserve anonymity, participants are referred to as P1–P7, and quotes are lightly edited for readability. Read against Pólya's four stages of problem-solving (Pólya, 1945) — understanding the problem, devising a plan, carrying out the plan, and looking back — the two themes also locate *where* in the cycle engagement is kept or surrendered: students reported retaining the early interpretive stages (understanding and planning) while delegating the later executional ones (carrying out and, in debugging, looking back), with persistence governing the iterative return from a failed result to a new plan. Table 1 sets out this mapping across both themes, and the two subsections that follow develop it.

4.1 GenAI Reliance as a Deliberate Choice

The most prominent theme was that students described their reliance on GenAI as a deliberate choice rather than a loss of ability. Its two categories appear in the upper half of Table 1.

Reliance is by choice and self-governed. Students did not portray GenAI as a default they automatically reach for, but as a tool they decide when and how much to use, staying aware of the difference between using it and depending on it. P3 made this clear by separating the speed the tool gives him from actually getting better at solving problems, and warning that over-relying on it means he ends up not learning:

> **P3:** "That's not really my performance improving [...] it's just my efficiency to solve a problem. Usually, you solve problems to get better yourself, and if you have this thing you can rely on, you can get over-reliant on it and then not learn anything."

P5 described the same control, placing himself clearly on the "use it" rather than "rely on it" side:

> **P5:** "I do feel independent. Compared to before, I'm less independent, but I definitely still feel like it's a tool I use rather than a tool I rely on."

P1 noted that having already done a task with AI made him more likely to repeat that route than to redo the work himself:

> **P1:** "It might be a task that I already know, and I've done before with the help of AI, so I'm gonna repeat that instead of trying by myself."

Theme	Category	Pólya stage(s)	Description
GenAI Reliance as a Deliberate Choice	Reliance is by choice and self-governed	Owns Stages 1–2; tool as a fallback at Stage 3	Students framed reliance as a deliberate choice rather than lost ability, distinguishing "choosing not to" from "not being able to." They tried problems on their own first and used the tool only as a limited, self-directed fallback.
	Confidence in skill is retained, practice is delegated	Owns Stages 1–2; delegates Stages 3–4 Students held that their problem-solving ability remained intact even as they handed over coding or information lookup, keeping the core reasoning and final judgment for themselves.	
Reduced Willingness to Persist Through Difficult Tasks	Lowered threshold to start	Hands off before Stages 1–2 begin	The in-the-moment instinct to hand a task to the tool before engaging with it, triggered by tedium, tiredness, or low interest rather than genuine difficulty.
	Gradual decline in motivation over time	Weakens the Stage 4 → Stage 2 return	A gradual, cumulative drop in the drive to take on effortful learning and push past the hard part of the learning curve, described as a settled habit rather than a one-time decision.

Table 1: Both themes and their categories mapped onto Pólya's four problem-solving stages.

Across these accounts, students describe deciding for themselves when and how much to use GenAI, depending on the situation, rather than relying on it automatically; P3's worry about "not learning anything" is the reason they keep their use limited.

Confidence in skill is retained, practice is delegated. Alongside this control, students consistently said their problem-solving ability remained, even as they handed over coding or information lookup. P4 gave the fullest account, separating the factual knowledge a programming language requires — which he draws from the tool — from the problem-solving ability itself, which he held was unchanged:

> **P4:** "I don't think [it changed] my problem-solving. The knowledge of how a programming language works, yes [that's where AI helps]. But if I already had that knowledge, I think I could solve the problem with my problem-solving skills. So it's more for filling in the knowledge gaps. [...] There's not really a way to test whether I can solve problems just as well as before, but I would think so, if I have all the information I need."

He extended the same stance to how he treats the tool's output, keeping the final judgment for himself and trusting his own assessment over the AI's, discarding suggestions he judged to be wrong:

> **P4:** "I can be independent of the AI, but the trust in my work would be lower [...] I don't trust AI too much, I trust myself better. So if I think the AI's ideas are not correct, then I won't use it."

P3 showed the same kept ability in how he starts a problem: he keeps the understanding and planning stages (Stages 1–2) for himself and turns to the tool only after trying on his own, once he is stuck while writing the solution (Stage 3):

> **P3:** "It doesn't change how I approach the problem itself, because I first approach it on my own still [...] in competitive programming, I think on my own, and then I ask AI if I get stuck. Before, if I got stuck, I would ask someone else or look up the solution."

P2 separated copying a fix — which happens at the carrying-out and looking-back stages (Stages 3–4), where the code is already written and failing — from the skill itself, which he described as still there and still improving:

> **P2:** "If something doesn't work with the library, I'm going to copy and paste and ask, 'Why is this wrong?' [...] I think my problem-solving skills are still there [...] I'm improving gradually throughout my program. [...] If I give myself the time, I think I can still do it."

Because this study captures perceptions rather than observed behaviour, this confidence is Itself a finding: students' belief in their ability stays strong even as they practise it less often. The recurring phrase "I could still do it if I wanted to" works both as a judgment of their own ability and as reassurance. Taken together, the stages students keep for themselves come earlier in Pólya's cycle than the work they hand over: understanding and planning are kept (Stages 1–2), while carrying out and the checking that goes with debugging (Stages 3–4) are the parts most often handed off.

4.2 Reduced Willingness to Persist Through Difficult Tasks

The second theme concerned what students reported changing: not their ability, but their willingness to push through the hard part of a problem without help. Its two categories appear in the lower half of Table 1.

Lowered threshold to start. This category captures the in-the-moment instinct to hand a task to the tool before engaging with it, triggered by tedium, tiredness, or low interest rather than by genuine difficulty. P2 gave the fullest account, watching himself bypass thinking as the bar to begin dropped:

> **P2:** "The barrier to do something now is much lower [...] you barely need to think in some cases, and that's very bad. I see myself doing that sometimes."

P5 described the same instinct as an almost automatic response to tedium, reaching for the tool before engaging with the work:

> **P5:** "Sometimes if I'm given more and more tedious tasks, I just instinctively think 'I'm just going to give this to the AI' [...] instead of getting started and putting in the work."

For P2, the same lowered threshold surfaced around fatigue and low interest, conditions under which the intended norm of working unaided gave way:

> **P2:** "Uni work, you should be able to do mostly yourself with AI assistance [...] But sometimes when I'm tired, or I don't really feel like learning stuff that I'm not very interested in [I lean on it more]."

In each case, the tool is reached for in place of starting the work, with the decision driven less by the difficulty of the task than by an unwillingness to engage with it in the moment.

Gradual decline in motivation over time. Beyond the immediate instinct, students described a gradual, cumulative decline in the drive to take on effortful learning. P1 named it as the single largest downside and framed it as a worsening trend rather than a one-off shift:

> **P1:** "that is the biggest negative impact for me [...] the motivation to get past the hard part of the learning curve [...] it happens less and less."

P6 reported a settled new default for demanding work, while remaining careful about how much to attribute to GenAI:

> **P6:** "I don't know if the willpower has been affected by [AI] or other external factors as well, that's one of the harder things to say. [. . .] I will say it probably has, at least a tiny bit, for the harder or longer tasks, and now I just usually go and ask for that initial skeleton directly. But there could also be other factors, so I can't really say how much."

The phrasing in both accounts, "less and less," "now I just usually", frames the change as a habit forming over time rather than a single decision. This matters for Pólya's framework: persistence is what brings a student back from a failed attempt to devise a new plan, so as it weakens, the middle stages of the cycle are handed to the tool exactly when They are hardest, and therefore most instructive. P6's account is, in this respect, a partial disconfirming case: by allowing that the change might owe as much to fatigue and competing demands as to GenAI, he resists the tidy causal story, the theme otherwise tells, and that hesitation is reported here rather than resolved away.

Theme	**Category**	**Associated codes**
GenAI Reliance as a Deliberate Choice	Reliance is by choice and self-governed	conditional dependency; dependent in preference, not skill; feels between dependent and independent; tries independently before turning to AI
	Confidence in skill is retained, practice is delegated	confidence in own skills; retains problem-solving ability, delegates knowledge gaps; can work unaided but chooses to delegate
Reduced Willingness to Persist Through Difficult Tasks	Lowered threshold to start	lowered barrier to begin a task; offloading of tedious tasks; delegating under fatigue
	Gradual decline in motivation over time	effect on motivation/willpower to engage; declining motivation to climb the learning curve; gradual decline recognised in hindsight

Table 2: Codes grouped under each theme and category.

5 Discussion

T

The findings suggest that GenAI does not remove independent problem-solving so much as change when and why students engage in it. Within the first theme, students framed their reliance as a deliberate, self-governed choice rather than a loss of ability: they decided when to let the tool in, kept the core reasoning and final judgment for themselves, and delegated execution or knowledge gaps while maintaining that their underlying ability was intact. This refines the "dual role" described by Aruleba et al. (2025): scaffold And crutches are not two types of students but two modes available to the same student, selected on a case-by-case basis. The concern that AI may "replace thinking" (Jafari & Keykha, 2024) is thus only partly borne out — students described replacing *effort*, by choice, more than ability — which mirrors the account of cognitive offloading as a deliberate move to conserve mental effort, rather than a sign of incapacity (Risko & Gilbert, 2016), and echoes the selective, goal-dependent help-seeking observed by Prather et al. (2023). If ability stays intact, the more pressing concern lies in the second theme: a reduced willingness to persist through difficulty. This took two forms. The first was immediate. Faced with tedium, fatigue, or low interest, students lowered their threshold for starting and handed the task over almost before engaging with it. The second was slower. Over time, some students experienced a gradual fading in motivation, and settled into a habit of reaching for the tool on the hardest or longest work, a drift they recognised only in hindsight. Durable learning requires active engagement with difficulty (Marton & Säljö, 1976); by lowering the cost of avoiding that work, GenAI may shift students from a learning orientation towards a performance one: not because they cannot reason independently, But because the motivation to do so weakens, both in the moment and over time, when a faster The route is always available. These two themes describe the same self-perception from two sides: the self-governance of The first is exactly what erodes in the second. Students who stayed engaged

were, in effect, regulating their GenAI use around personal learning goals (Zimmerman, 2002), But this regulation broke down under precisely the conditions that most test self-regulatory capacity — deadline pressure, tedium, fatigue, and the difficulty of starting. For teaching practice, this points less towards restricting GenAI than towards protecting the motivation and space to reason independently, for example, through tasks that ask students to Critique and improve AI output, or designated stages of work, without it (Nnamdi et al., 2026). These interpretations rest on a small, single-institution sample and on self-reported perceptions rather than observed behaviour, with some risk of social desirability bias.

5.1 Future Work

Future research could combine self-reported perceptions with observed behaviour, for example, through task-based studies or conditions with and without AI access, to test whether perceived dependency matches actual ability. A larger and more diverse sample, drawn from several institutions and stages of study, would help show how widely these patterns hold. Longitudinal work could then track how reliance and independent problem-solving change over time, clarifying whether the motivational effects students describe deepen, stabilise, or reverse as they gain experience.

6 Conclusions

This study examined how the use of GenAI during programming tasks shapes the way Computer Science students engage with independent problem-solving. The interviews suggest that GenAI reshapes engagement rather than simply removing it. Students framed their reliance as a deliberate, self-governed choice rather than a loss of ability, keeping ownership of their reasoning and final judgment while delegating execution and knowledge gaps, and maintaining that they could still work unaided if they chose to. At the same time, they reported a weakening willingness to persist through difficulty without help — both an immediate, lowered threshold to begin a task and a gradual decline in motivation for effortful learning over time. In answer to the research question, GenAI appears to shape independent problem-solving chiefly by lowering the effort required to avoid it, which students then navigate according to their goals and the demands of the moment. The central tension is therefore not that GenAI makes students unable to reason independently, but that it makes independent reasoning the costlier option, one they increasingly decline to pursue. Whether an ability that is rarely exercised stays intact is a question this study raises but cannot answer.

7 Responsible Research

This section reflects on the ethical and methodological responsibilities of the study, following the principles of research integrity and reproducibility. It builds on the procedures described in the Methodology by addressing participant protection, researcher reflexivity, transparency, data management, and the responsible use of GenAI in conducting the research itself.

7.1 Ethical Treatment of Participants

The study received approval from TU Delft's Human Research Ethics Committee. (HREC). All participants received clear information about the purpose of the study and how their data would be used, and gave informed consent before taking part. Participation was voluntary, and participants could withdraw at any point without consequence. To protect privacy, all data were anonymised: participants are referred to by pseudonyms (P1–P7), and identifying details were removed during transcription. Recordings were stored securely and used only for analysis.

7.2 Researcher Reflexivity and Bias

Qualitative interviews are shaped by the researcher who conducts and interprets them, and reflexive thematic analysis explicitly recognises this role in constructing themes (Braun & Clarke, 2006).

Two sources of bias deserve particular attention. First, because the study relies on self-reported perceptions, students may understate their reliance on GenAI or overstate their independence, whether through social desirability or limited self-awareness; the findings should therefore be read as how students *perceive* their engagement, not as a direct measure of behaviour. Second, the interviewer's own framing could steer responses, and this risk is concrete rather than abstract: some of the vocabulary the analysis turns on, the distinction between "using" and "relying on" the tool, and the "scaffold" versus "crutch" framing, originated with the interviewer and was at times agreed to by participants rather than volunteered. To guard against reading agreement as evidence, formulations that participants produced unprompted (for example, P1's "it happens less and less" and P2's "you barely need to think") were weighted more heavily in developing the themes than phrasings that merely echoed the interviewer's wording. Participants were also assured that their answers were confidential and had no bearing on their grades, and the interviewer aimed to keep a neutral stance and to ask open rather than leading questions. In practice, this was not always achieved: in several interviews, the distinction between *using* and *depending on* GenAI was first articulated by the interviewer and then affirmed by the participant rather than raised independently. Such moments are treated as weaker evidence in the analysis, with greater weight given to formulations participants offered in their own words.

7.3 Transparency and Reproducibility

Qualitative research cannot be reproduced in the same way as a controlled experiment, since themes are interpreted rather than measured. Responsibility here lies in transparency about how conclusions were reached. To support this, the interview guide, the coding approach (inductive coding of emerging themes, later interpreted through Pólya's problem-solving framework), and anonymised interview excerpts will be made available so that another researcher can follow and scrutinise the reasoning behind each theme.

7.4 Data Management

Data were handled in line with TU Delft's data management policy. Audio recordings and transcripts were stored in secure, access-restricted locations, and no personally identifying information appears in the reported results. Only anonymised material is presented, and quotes are attributed to pseudonyms.

7.5 Responsible Use of GenAI in the Research

GenAI tools (specifically Claude) were used during the writing of this thesis to suggest rephrasings and improve the clarity and concision of prose I had drafted, to help reorganise the structure of the Literature Review and Results, to check the manuscript for internal inconsistencies, and as a sounding board for framing arguments. The tools were not used to conduct the thematic analysis, generate codes or themes, or interpret the interview data: Coding was carried out manually in ATLAS.ti, and all analytic decisions, interpretations, and conclusions are my own. Every passage produced with AI assistance was reviewed, edited, and verified by me. This use is disclosed in line with TU Delft policy.

References

Alpizar-Chacon, I., Keuning, H., de Jong, I., Lykourentzou, I., & Rings, S. (2025). Excited, skeptical, or worried? A multi-institutional study of student views on generative AI in computing education. *Proceedings of the 25th Koli Calling International Conference on Computing Education Research*. https://doi.org/10.1145/3769994.3770008

Aruleba, K., Sanusi, I. T., Obaido, G., Ogbuokiri, B., & Mienye, I. D. (2025). Beyond the prompt: Student strategies, ethical reflections, and learning with chatgpt in computer science. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-025-09899-7

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE Publications.

Jafari, F., & Keykha, A. (2024). Identifying the opportunities and challenges of artificial intelligence in higher education: A qualitative study. *Journal of Applied Research in Higher Education*, *16*(4), 1228–1245. https://doi.org/10.1108/JARHE-09-2023-0426

Loksa, D., Ko, A. J., Jernigan, W., Oleson, A., Mendez, C. J., & Burnett, M. M. (2016). Programming, problem solving, and self-awareness: Effects of explicit guidance. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4420–4431. https://doi.org/10.1145/2858036.2858252

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—outcome and process. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Nnamdi, M. C., Tamo, J. B., Marteau, B., Shi, W., & Wang, M. D. (2026). Advancing problem-based learning in biomedical engineering in the era of generative ai. *IEEE Transactions on Education*, *69*(2), 112–124. https://doi.org/10.1109/TE.2026.3658007

Ovi, J. A., Fierro, G., & Smith, C. E. (2025). Assessing student adoption of generative artificial intelligence across engineering education from 2023 to 2024. *arXiv preprint arXiv:2503.04696*.

Pólya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton University Press.

Prather, J., Denny, P., Leinonen, J., Becker, B. A., Albluwi, I., Craig, M., Keuning, H., Kiesler, N., Kohn, T., Luxton-Reilly, A., MacNeil, S., Petersen, A., Pettit, R., Reeves, B. N., & Savelka, J. (2023). The robots are here: Navigating the generative ai revolution in computing education. *Proceedings of the 2023 Working Group Reports on Innovation and Technology in Computer Science Education (ITiCSE-WGR '23)*, 108–159. https://doi.org/10.1145/3623762.3633499

Rahe, C., & Maalej, W. (2025). How do programming students use generative AI? *Proceedings of the ACM on Software Engineering*. https://doi.org/10.1145/3715762

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

A Interview Guide

Logistics

Before recording, participants were thanked for taking part and informed that the The interview would be audio-recorded and later transcribed, and notes might be taken during the conversation, and that the interview was expected to last around 45 to 60 minutes. Participants were invited to ask any questions, and the consent form was presented and signed before recording began. After starting the recording, the following points were explained to each participant:

- The purpose of this interview is to understand students' experiences using GenAI tools in their studies, in relation to problem-solving and learning.
- I want to hear your own experiences and how you perceived them. There is no right or wrong answer.
- I want to understand your perspective, so I will ask questions like "what do you mean by that?" I will purposefully leave long pauses after my questions.
- Please take your time, and if you need clarification, please let me know.

Category 1: Warm-up, general questions (5 min)

- How often do you use GenAI tools for your studies?
- What types of tasks do you find GenAI most helpful for?

Category 2: Interaction strategies and evolution (12 min)

Purpose: to capture the mechanics of how they use the tools and whether their habits have changed as they have become more experienced.

- What factors influence your decision to use GenAI immediately versus trying it yourself first?
- Walk me through a typical situation where you use AI to find a solution to a programming task. What does it look like?
- How did you use GenAI for your studies in the beginning (when ChatGPT came out), and how does that compare with the way you use it now? What has changed?
  - *Follow-up:* If you were to go back, how would you use ChatGPT? What would you change?

Category 3: Cognitive impact and problem-solving (12 min)

Purpose: to explore the students' perception of their own skill development and whether they feel the AI is a "crutch" or a "scaffold."

- Do you feel that using GenAI has changed your independent problem-solving skills for learning tasks (e.g. programming tasks)? How?
- Do you think AI has changed the way you approach problems before attempting to solve them? How?
- When do you feel AI helps you learn best, and when does it prevent learning?

Category 4: Institutional frameworks and ethics (12 min)

Purpose: to gather actionable advice on how the university and professors should handle AI in the future.

- In your opinion, whose responsibility is it to ensure that students do not become overly dependent on AI, the student's own self-discipline, the professor's assignment design, or the university's regulations?
- Do you think there are situations where using GenAI in coursework becomes unfair or unethical? Why or why not?

Conclusion

- Do you have any questions for me?
- Is there anything about your experience that we didn't cover but you think is important?

Participants were thanked for taking part and sharing their experiences.