Bachelor Computer Science & Economics

Ask it Everything: Drivers of GenAI Overreliance among University Students

Vlad Popa

Supervisors:
Dr. P.W.H. van der Putten, Dr. M.P.H.D. Cleiren

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

01/07/2025

Abstract

Generative AI (GenAI) tools like ChatGPT and Gemini are increasingly used in higher education. However, excessive reliance can undermine critical thinking. This thesis examines how cognitive, educational, and technical factors drive GenAI overreliance among university students. Two conceptual frameworks are proposed: a factor-centered framework that hierarchically organizes environmental, situational and cognitive influence, and a process-centered framework that describes how these factors interact over time to shift students between constructive and compensatory GenAI use and lead to overreliance.

To validate these frameworks, a survey of 99 Dutch university students measured (1) demographic background, (2) frequency of GenAI use in various academic situations, (3) constructive versus compensatory behaviors, (4) cognitive, educational, and technical drivers, and (5) change in usage over time. Descriptive analysis revealed that students turn to GenAI primarily under deadlines and when overwhelmed by study material, and many admit to overreliance despite generally constructive use. Correlations identified low self-efficacy and cognitive laziness as the strongest cognitive links to overreliance, while educational and technical factors showed weaker associations. A Random Forest regression (with SHAP values) confirmed three key predictors of overreliance: high clarity of institutional guidelines, high cognitive laziness, and low problem solving self-efficacy.

These insights highlight that GenAI overreliance arises from interactions of what, when and how. The frameworks and findings suggest that boosting student confidence and engagement, creating awareness around GenAI guidelines, and promoting reflective use can help maintain GenAI as a learning aid rather than a shortcut to bypass meaningful learning.

Contents

1 Introduction **4**
1.1 Problem statement . . . 4
1.2 Thesis overview . . . 5

2 Background **5**
2.1 AI in education . . . 5
2.1.1 Emergence of AI in education . . . 5
2.1.2 GenAI applications for students . . . 6
2.2 GenAI adoption among students . . . 6
2.2.1 Direct motivations behind using GenAI tools . . . 7
2.2.2 GenAI's impact on academic performance . . . 7
2.2.3 Risks of prioritizing efficiency over learning . . . 8
2.3 AI overreliance . . . 9
2.3.1 Cognitive offloading vs. AI overreliance . . . 9
2.3.2 AI overreliance and its manifestation . . . 10

3 Methodology **11**
3.1 Framework Development . . . 12
3.2 Data Collection . . . 12

4 Conceptual Foundations of GenAI Overreliance **12**
4.1 Contributing Factors . . . 13
4.1.1 Cognitive Factors . . . 13
4.1.2 Educational Factors . . . 15
4.1.3 Technical Factors . . . 16
4.2 Theoretical Foundations . . . 17
4.2.1 Metacognitive Model of Cognitive Offloading . . . 17
4.2.2 Social Cognitive Theory . . . 18
4.2.3 Dual Process Theory . . . 19
4.3 Factor-Centered Framework . . . 20
4.4 Process-Centered Framework . . . 21

5 Survey Setup **23**
5.1 Survey Design . . . 23
5.2 Participants and Distribution . . . 25
5.3 Pilot Study . . . 29

6 Survey Results **29**
6.1 Descriptive Analysis . . . 30
6.1.1 Demographic Overview . . . 30
6.1.2 Usage Frequency across Contexts . . . 30
6.1.3 Usage Over Time . . . 32
6.1.4 Nature of Overreliance . . . 33
6.2 Correlation Analysis . . . 36

6.2.1 GenAI Usage Frequency and Overreliance 36
6.2.2 Cognitive Factors and Overreliance 38
6.2.3 Educational Factors and Overreliance 40
6.2.4 Technical Factors and Overreliance 42
6.3 Regression Analysis 43

7 Discussion 48
7.1 Reflection on the Conceptual Frameworks 48
7.2 Reflection on the Survey and Results 49
7.3 Limitations and Future Work 50

8 Conclusion 52

References 56

A Appendix: Visualized Survey Answer Distribution 57

B Appendix: Visualization of Overreliance Score per Factor 60

C Appendix: Visualized Decision Tree 61

1 Introduction

As of recently, Artificial Intelligence (AI) developments have shown its potential to (partially) handle creative tasks that are typically performed by experts and professionals, such as programming and arts (Y. Li et al., 2022). Particularly, the emergence of Generative AI (GenAI), has proven to be highly adaptable. Since the release of ChatGPT, a GenAI tool, in November 2022, the tool had welcomed over 100 million monthly and 13 million daily users within two months (Hu, 2023). By the end of 2023, the North American AI market reached a valuation of $30.9 billion (Gu, 2024), reflecting the tool's ability to expand across various industries (Patil et al., 2024a).

The academic sector is no exception. According to a survey by Klarin et al. (2024), more than half (52.5%) of students report using GenAI for academic purposes, with ChatGPT emerging as the most commonly used tool. Its capabilities include answering exam-style questions, support homework assignments and generate academic literature based on human prompts (Zhai, 2022).

While some research highlights the positive effects of GenAI on reaching learning goals, such as grade improvements (Deng et al., 2024), concerns are raised about the long-term effects of routine usage. For instance, a study by Bastani et al. (2024) states that students who relied on GenAI tools underperformed the peers who had never used it, once access to GenAI tools is subsequently removed. Similarly Gerlich (2025) observed that critical thinking is reduced with frequent use of AI tools.

The act of shifting cognitive tasks to external tools, such as calculators, is known as cognitive offloading, and conserves cognitive effort. However, this behavior can evolve into overreliance. In literature, it is defined as the occurrence in which users trust the AI without determining the extent of it (Passi and Vorvoreanu, 2022). Studies show increased reliance on AI output is caused by the increased trust users place in the AI tool (Klingbeil et al., 2024). This can lead users to, despite conflicting with correct information and own interests, follow AI advice (Klingbeil et al., 2024).

But even accurate output can encourage dependence, reflecting a preference for efficiency over learning (Jo and Bang, 2023). Therefore, we define GenAI overreliance as the state in which a student consistently defaults to using GenAI tools without critically evaluating the need, context, or reliability of the tool. We employ this definition for the remainder of this thesis.

Although the well known documented risks associated with GenAI overreliance, such as the decrease in a person's ability to think critically or solve problems independently without the assistance of AI (Eckhardt et al., 2024), research on its drivers is lagging behind.

1.1 Problem statement

With the increased embedding of GenAI in educational practices, understanding the drivers behind student overreliance on these tools becomes even more important. This thesis addresses this research gap by investigating the cognitive, educational and technical factors that contribute to GenAI overreliance with regard to university students. To this end, the research question states: *How do cognitive, educational and technical factors influence GenAI overreliance among university students?*

This thesis employs a combination of methods in its approach. It begins with a review of literature. Based on these insights, two novel original conceptual frameworks are introduced: a factor-centered framework synthesizing key influences contributing to GenAI overreliance, and a process-centered framework detailing the development of GenAI overreliance. These frameworks are tested through an empirical survey targeting Dutch university students with experience in GenAI tools. Insights from the survey, obtained through descriptive, correlational and regression analyses, have provided valuable evidence on the primary factors that shape students' reliance on GenAI.

1.2 Thesis overview

The following sections outline the thesis' structure that together provide the complete analysis of the research conducted. Section 2 reviews existing literature on GenAI in education, adoption among students and AI overreliance. In Section 3, the thesis discusses the methodology of this study. Next, the thesis addresses the conceptual foundation of GenAI overreliance by constructing two conceptual frameworks (Section 4). Then, Section 5 describes the distributed survey that empirically assesses the influence of several identified factors and the components of these frameworks. Finally, the responses from the survey are analyzed and reported in Section 6, followed by a discussion in Section 7, after which a conclusion is drawn in Section 8.

2 Background

As AI continues to evolve, its role in education is becoming increasingly significant. Emerging technologies such as GenAI are reshaping traditional schooling methods by offering new ways for students to engage with educational content. However, these innovations also introduce new challenges, such as the risk of AI overreliance.

To understand this phenomenon, this section explores the emergence and growing application of GenAI in learning environments, the reasons behind student usage as well as its benefits and challenges, and the concept of AI overreliance.

2.1 AI in education

This subsection examines the emergence of AI in education, highlighting its impact on learning environments. Additionally, it reviews the broad GenAI applications for students, focusing on how conversational assistants facilitate personalized learning and extend support across various fields of study.

2.1.1 Emergence of AI in education

Data-driven AI has become a focus of commercial interest since the beginning of its development in 1956 (Holmes and Tuomi, 2022; McCarthy et al., 2006). This is due to the rapid advances it has made from the increased accessibility to novel processor chip architectures that optimize computations, the increase in data on the internet and the massive human effort in labeling this data (Tuomi, 2019). These advances allowed AI applications to have a massive impact in a variety of sectors by optimizing industry processes (Patil et al., 2024b).

One of these sectors is education. This is not surprising, as there are significant opportunities for the application of AI in education that lessens the burdens of both teachers and students.

As AI continues to expand its role in education, the increasing adoption of AI-powered tools is likely to reshape traditional teaching and learning dynamics. While these technologies offer immense benefits, such as improving accessibility, streamlining administrative processes, and providing personalized learning experiences (Pavlik, 2023; T. Wang et al., 2023; Zhai et al., 2021), they also raise questions about the evolving role of educators. As more AI tools are developed across diverse subjects, the nature of schooling itself may need to be reconceptualized to integrate AI-assisted learning for students effectively while preserving essential aspects of critical thinking, creativity, and human interaction in education (Zhai et al., 2021).

2.1.2 GenAI applications for students

The growing adoption of GenAI tools usage has significantly influenced higher education by offering students novel learning approaches. Although these tools have not been specifically developed for educational use, they can create an interactive and engaging learning environment through their ability to adapt to instructions from human-like conversations (Dai et al., 2023) and beyond, such as image generation, dataset analysis etc. Through real time assistance, GenAI can generate tailored explanations and scaffold learning in more coherent and sensitive ways based on the identification of an individual learners' strong and weak points.

Fong et al. (2014) emphasize GenAI's potential, such as the models developed by OpenAI. They state that GPT models for instance could generate student profiles through assessments that uncover students' thoughts and progress based on prompts. Based on these profiles, learning experiences could be aligned with the students' preferences, engagement, performance and pace of learning.

Therefore, it might be said that these technologies have been 'repurposed' for learning instead (Holmes and Tuomi, 2022). The ability of GenAI tools to answer questions and offer guidance (Kooli, 2023; T. Wang et al., 2023) by processing various types of content such as essays and computer programs (Dai et al., 2023), allows it to enhance their teaching abilities to meet the needs of individual learners.

Dai et al. (2023) also mention that student learning engagement is enhanced through the instant responses and guidance GenAI can provide. The obtained feedback allows for applying new understanding to subsequent tasks in real time, as students are able to learn at their own pace. This can be accredited to the tool's rich database, which allows it to generate ideas, provide suggestions and spark creative thinking among students.

However, GenAI performance varies between different subject areas. Rahman and Watanobe (2023) found that ChatGPT performed well in daily programming exercise and is capable of supporting programmers by offering personalized programming aid based on the code generated. On the contrast, a different study discovered that ChatGPT fails to solve simple mathematical questions and advanced process analysis questions in the absence of human hints (Terwiesch, 2023). This may therefore result in GenAI's usefulness being dependent on the types of questions prompted, subject area, educational level and provided context to facilitate its chain of thought.

It can be said that therefore, ChatGPT and other GenAI tools are no longer passive technologies, but rather active co-creators of the educational experience (Dai et al., 2023).

2.2 GenAI adoption among students

Building on the foundational role of AI in education, this subsection explores the motivations behind GenAI adoption, including its accessibility, efficiency and interface. Furthermore, it examines the impact on academic performance, as well as the risks of prioritizing efficiency over learning.

2.2.1 Direct motivations behind using GenAI tools

One of the reasons behind the use of GenAI tools, such as ChatGPT, Gemini and Claude, is their high online accessibility and free plans (Huang et al., 2022), making them readily available for students.

Another motivation behind student use is the desire for efficiency (Skjuve et al., 2024). Efficiency comes forth from GenAI's ability to offer clear answers to complex questions, which eliminates the need for intricate search queries or sifting through multiple sources. This helps reduce information overload, summarize information, and eliminate choices (Skjuve et al., 2024; Zhai et al., 2021). Skjuve et al. (2024) also found that young people in particular resort to GenAI to support knowledge work by enabling efficient yet elaborate information retrieval for various tasks. The importance of finding productivity in GenAI tools is especially important for students to cope with academic pressures.

The use of GenAI tools is further stimulated among students through the interactive user interface centered around inputs. GenAI interfaces often echo the logic of learning centered around student preferences in higher education, placing students in control of planning, organizing, and personalizing their learning process (Dai et al., 2023; E. Lee and Hannafin, 2016). As students find themselves in a virtual real-world conversational environment, their motivation and engagement are increased (Y. F. Wang et al., 2017). Students are generally more willing to learn in a student-centered learning environment due to perceived choice and value (Chen et al., 2020; Yin et al., 2021). Without the intervention of any other human beings in this environment, embarrassment, nervousness, anxiety and uneasiness decrease (Zhang et al., 2023).

However, the effectiveness of these tools depends on students' ability to interact with them. To maximize usage success, students are required to write structured prompts and employ advanced questioning strategies when using GenAI tools. Furthermore, students should also be aware of the tool's capacity and limitations (Dai et al., 2023). This process follows a "trial and error" approach, as students need to test different practices for interacting with GenAI tools and adjust their expectations.

2.2.2 GenAI's impact on academic performance

When integrated into learning environments, AI-powered tools provide substantial support, particularly for those who initially struggle, helping them achieve performance levels closer to their higher-performing peers (Mollick, 2023). Based on findings from multiple studies, it can be demonstrated that AI can significantly enhance academic performance.

One of these studies conducted research on AI-assisted business consulting tasks, and found that individuals who initially scored in the lower performance bracket improved their output quality by 43% when using AI. On the other hand, top performers saw a comparatively smaller increase of 17%. As a result of implementing AI assistance, the performance gap between the two groups, which initially stood at 22%, was reduced to just 4% (Dell'Acqua et al., 2023).

Similar patterns have been observed in writing-based tasks. One study found that when less proficient writers utilized AI-generated content suggestions, their writing quality improved significantly, closing nearly half of the skill gap between them and stronger writers (Noy and Zhang, 2023). A parallel study on creative writing showed that access to AI-generated story ideas helped individuals with lower creativity scores reach levels comparable to those of naturally more creative writers (Doshi and Hauser, 2023).

Comparable findings have emerged in legal education, where research indicates that law students at the lower end of the class distribution were able to match the performance of top students when given access to AI tools (Choi and Schwarcz, 2023). Likewise, in customer service roles, AI assistance had the most pronounced effect on employees who initially performed at a lower level, helping them close the gap with higher achievers (Brynjolfsson et al., 2025).

Beyond individual case studies, a broader meta-analysis suggests that GenAI models can enhance academic performance across various subjects. As Deng et al. (2024) put it, "short-term experiments involving the usage of ChatGPT result in different gains in academic performance across subject areas, with 44 out of 51 effect sizes being positive".

The majority of studies analyzing AI interventions in educational settings report positive effects, with a significant proportion of experiments demonstrating notable improvements in student achievement. However, these studies tend to share key limitations: they often focus on short-term interventions, involve relatively small sample sizes, and assess performance through straightforward tasks. This raises questions about the long-term implications of AI usage in education, particularly when students face more complex, domain-specific challenges that may exceed the capabilities of current AI models. While AI can act as a valuable tool for efficient, fast academic improvement, its ability to sustain these benefits over time remains uncertain.

2.2.3 Risks of prioritizing efficiency over learning

A key concern with the use of GenAI in education is the tendency for students to prioritize efficiency over deeper learning. This begs the question as to whether GenAI brings students closer to their learning goals, or whether it moves them away.

Research suggests that this motivation for efficiency is often so strong that users overlook explicit warnings from AI service providers regarding the potential inaccuracy of generated responses (Dai et al., 2023).

Another risk of an efficiency focused approach is the encouragement of a "procrastination friendly" environment through GenAI usage (Abbas et al., 2024; Mukhtar et al., 2025). Instead of enhancing work processes, GenAI induces unexpected delays when excessively used. This environment promotes the aversion of engaging in studying, which does not only detriment critical thinking abilities and memory (Abbas et al., 2024; Gerlich, 2025), but also traits such as creativity (Webster and Kruglanski, 1994) and self-efficacy (Mukhtar et al., 2025). Without critical thinking skills, students risk accepting such outputs at face value, which can lead to the propagation of misinformation in their academic work (Spector and Ma, 2019). Developing a discerning mindset is therefore essential, as it enables students to assess the validity of AI-generated content, recognize biases, and make informed decisions about the information they incorporate into their learning. This can be especially dangerous in novel or complex tasks where deep analysis is required (Hiel and Mervielde, 2002).

Empirical studies further highlight the academic performance risks associated with relying on AI tools. Bastani et al. (2024) conducted an experiment in which students were divided into groups and provided with traditional study materials alongside different AI-assisted learning environments: Students were randomly divided in two groups, where they were granted access to their course material, as well as a tutoring program based on OpenAI's GPT-4: "GPT Base" and "GPT Tutor". The first variant was similar to ChatGPT and prompted the practice problem that was being solved, indicating that the model should assist the student as a teacher throughout the prompt. The second program worked the same, but had safeguards implemented in its prompts to mitigate bias in its generated output, and provided hints to the student instead of directly giving them the answer. During their assisted practice sessions, the students would prepare for a math exam using their given resources.

While both GPT Base and GPT Tutor would increase performance (48% and 127%, respectively) on the assisted practice sessions, the subsequent unassisted exam proved otherwise. Based on the assessment, the average control student's performance, who used GPT Base, dropped by 17%, while the performance of students using GPT Tutor had minimal improvements. Academic stressors, including the induced procrastination from uncontrolled and excessive GenAI usage, explain this decrease in academic performance (Abbas et al., 2024). This suggests that while AI can enhance short-term performance when available, excessive dependence on direct answers may impede long-term learning and retention, especially when students are not engaged in critical thinking through safeguards (Bastani et al., 2024).

Students themselves recognize these challenges. Chan and Hu (2023) found that students express that the use of GenAI may weaken critical thinking and creativity. This could be explained by a similar study, where students acknowledge that their efficiency-focused AI approach hinders them from doing the knowledge processing work that teachers expect them to do. Students therefore become unwilling to engage in deep learning (Zhai et al., 2021), as critical thinking could be regarded as "redundant". The issue is further elaborated as students express concerns about the increasing difficulty in assessing validity or identifying falsehoods in GenAI output (Chan and Hu, 2023).

Despite their awareness of GenAI's limitations and consequences, AI-generated content should always be questioned and analyzed by students, rather than consumed passively. Instead of focusing solely on efficiency, students should prioritize effectiveness by using GenAI as a supplement to critical thinking rather than a replacement.

2.3 AI overreliance

To contextualize the risks mentioned in Subsection 2.2, this subsection defines AI overreliance from literature and its relation to cognitive offloading, as well as its manifestation in academic settings.

2.3.1 Cognitive offloading vs. AI overreliance

Cognitive offloading refers to the use of external tools to overcome mental memory capacity limitations, reduce mental effort and enhance problem-solving capacity (Risko and Gilbert, 2016). External tools range from physical gestures and written notes to digital technologies like calculators or navigation systems. This strategy can be particularly beneficial when memory or attention is constrained, as it allows individuals to store or process information in the environment rather than internally. However, individuals do not always benefit from offloading. In many cases, they rely on it based on flawed metacognitive judgments, mistakenly believing it will improve their outcomes (Risko and Gilbert, 2016).

The introduction of computers and the internet has vastly expanded offloading potential, enabling instant access to information that once required internal storage (Bastani et al., 2024; Risko and Gilbert, 2016). Yet, this shift introduces a trade-off: as technologies automate more cognitive work, users are less likely to develop or maintain the associated skills themselves. Tasks delegated to machines may yield short-term efficiency but simultaneously reduce opportunities for skill acquisition and mastery (Bastani et al., 2024).

GenAI tools such as ChatGPT amplify this phenomenon. These systems now handle not only information storage and retrieval but also aspects of decision making, reasoning, and creative generation. While offloading such tasks can free cognitive resources for higher-level thinking, it may also reduce cognitive resilience over time if users habitually defer to AI for mental labor (Gerlich, 2025). The increase in reliance on AI will result in the underdevelopment or even erosion of trivial skills such as memory consolidation, complex reasoning, and independent problem solving. This can be explained by the reduced willingness of engaging in critical thinking (Abbas et al., 2024; Gerlich, 2025). Instead of using critical thinking skills for information gathering, problem solving and task execution, it is instead shifted to nformation verification, AI response integration and task stewardship as a consequence (H.-P. Lee et al., 2025). As routine cognitive engagement diminishes as a consequence, so too might long-term cognitive flexibility and depth.

Recent research suggests this pattern is already emerging in academic settings. Chan and Hu (2023) found that offloading to AI tools plays a significant role in the relationship between critical thinking ability and AI use. Thus, the more students rely on AI to ease mental load, the less they tend to engage in reflective analysis, suggesting that excessive reliance may hinder cognitive development.

While many technologies historically created similar trade-offs, GenAI tools like ChatGPT are fundamentally different in two key ways according to Bastani et al. (2024). First, their scope is significantly broader, often capable of assisting in complex academic domains such as mathematics and writing. Second, their outputs are frequently inaccurate despite being delivered with convincing confidence. When students overtrust these tools, particularly in the presence of mentioned metacognitive beliefs, they are more likely to accept incorrect answers. This blind acceptance marks the transition from cognitive offloading to AI overreliance. The tool is not simply assisting thinking, but replacing it entirely.

2.3.2 AI overreliance and its manifestation

According to Passi and Vorvoreanu (2022), overreliance on AI is defined as users accepting incorrect AI output because they are unable to determine whether or how much they should trust the AI. People tend to overrely on AI advice in scenarios where the advice contradicts available information and their own assessment (Klingbeil et al., 2024).

One possible explanation for the extent of this trust, is the assumption that machines excel at repetitive and mechanical tasks such as objective calculations compared to humans (Klingbeil et al., 2024). Users are led to doubt their domain knowledge rather than validate it, and put their trust in the AI.

Another explanation could be the credibility of the source, where users could perceive an AI's output as credible if they believe it has access to resources (Schreuter et al., 2021). Yet, AI may generate false sources or simply lie about its reasoning to sound credible.

Fostered by the perceived trustworthiness and accessibility of AI tools, the trust encourages a dependence that deems the need for active cognitive effort and engagement unnecessary (Bastani et al., 2024; Gerlich, 2025). However, the problem extends beyond accepting biased output, as students could over rely on AI-generated output while it's correct.

Several studies suggest that the more trust is placed in AI's utility, the more likely users are to entrust cognitive tasks to these tools. This excessive reliance leads to a decline in the users' critical thinking skills, as they become dependent on AI tools for instant answers or insights and thus less proficient at engaging in independent thought (Bastani et al., 2024; Chan and Hu, 2023; Gerlich, 2025; H.-P. Lee et al., 2025). For example, lower scores on structured assessments were observed among students who relied more heavily on AI technologies (Gerlich, 2025).

Another important effect of AI overreliance lies in complacency. Students relying on external support rather than their self-belief and efforts become complacent, diminishing the role of self-efficacy and reduced personal competence as these necessities are reduced (T. Wang et al., 2023).

Gerlich (2025) also highlights how AI overreliance reduces decision making capacities and critical analysis skills due to the convenience of delegating tasks to AI tools. This benefit turns into a consequence: the dependence makes individuals become used to using AI as an escape for decision making and problem-solving. Once access is taken away, cognitive resilience is reduced as users are unable to operate without these tools. This vulnerability was particularly evident in a study by Bastani et al. (2024), where students' performance dropped when access to GenAI study assistants was removed, demonstrating the issues that arise in high pressure situations without technological assistance.

Overreliance will therefore be reflected in diminished quality, misinformation, biases, procrastination, and misuse of AI, resulting in societal complications. While AI tools can improve basic knowledge acquisition, they may not foster the crucial, cognitive skills required for applying insights in unfamiliar or complex situations. It is urgent that users understand the risks involved when utilizing AI tools to present information or suggestions, especially because they can be erroneous or biased for domain-specific knowledge (Skjuve et al., 2024).

3 Methodology

Since (Gen)AI overreliance extends beyond merely trusting (Gen)AI outputs, this thesis introduces a refined definition: *"The state in which a user consistently defaults to using (Gen)AI tools without critically evaluating the need and reliability of the tool within its deployed context."* This definition emphasizes the habitual aspect of reliance, where users bypass critical assessment and reflexively depend on AI tools to fulfill cognitive tasks.

This is different from normal "reliance", where the users are conscientious about the guidance and limitations of the tools they are using. For instance, the reliance on a GPS system is necessary when traveling to new places. In the academic context of student learning, it describes the extent to which learning goals are achieved, or moved away from. For example, using GenAI tools to summarize study material so you don't have to read it, indicates a high degree of overreliance according to the definition.

Following this definition, the methodology consists of two components. The first component centers on the development of two frameworks that synthesize identified drivers of GenAI overreliance and insights from original contributions, and established cognitive models and theory. The second component focuses on the design and deployment of a quantitative survey, aimed at empirically testing key components of these frameworks.

3.1 Framework Development

To identify and structure the key drivers of GenAI overreliance in educational settings, this research draws upon a case study approach. Through original theoretical reasoning and a thorough review of recent and relevant literature, the cognitive, educational, and technical factors are identified that may contribute to students' overreliance on GenAI tools. Based on these factors, a visual factor-centered framework is developed.

Another complementary process-centered framework is constructed to capture how GenAI overreliance develops based on these identified factors. In this framework, the factors are synthesized and mapped against theoretical underpinnings from several behavioral models and theories. Drawing on these theoretical foundations, the framework outlines the mechanisms through which students perceive, interpret and act upon academic tasks using GenAI before becoming overreliant.

The factor-centered framework complements the process-centered framework by providing a hierarchical view of the contributing inputs from which overreliance emerges. Together, they serve as the theoretical backbone for the empirical phase of the study and guide the design of the survey.

3.2 Data Collection

To empirically assess the proposed frameworks, a survey was developed and distributed among students in Dutch higher education. The survey aimed to test the influence of quantitatively measurable factors on GenAI overreliance from the factor-centered framework, aided by the process-centered framework to explore their extent in practice. Due to feasibility constraints, the survey only managed to focus on factors that are both theoretically relevant and measurable. More information regarding the setup of the survey can be found in Section 5.

Participants were recruited from Dutch higher education institutions, including research universities and universities of applied sciences. Distribution occurred via institutional communication channels and widely used social media platforms to ensure accessibility and diversity. Respondents had to use GenAI tools regularly in an academic context to qualify for participation.

The survey mostly contained Likert scaled questions to capture the combined influence of factors in students' attitudes, behaviors and background characteristics in relation to their degree of overreliance. This was done through behavioral items that assessed reliance, with the inclusion of realistic scenarios that show GenAI use, and a minimal sense of change over time. Demographic variables such as program level and field of study were collected to enable stratified analysis. The gathered data supported descriptive, correlational and regression analyses, offering grounding for the conceptual frameworks and facilitating targeted discussion on the generalizability and limitations of the findings.

4 Conceptual Foundations of GenAI Overreliance

This section presents the key contributing factors behind why students become overreliant on (Gen)AI tools. These factors have been categorized into three separate domains, and reflect both insights drawn from existing literature and original theoretical reasoning. These categories help clarify the nature of each influence and lay the foundation for a factor-centered framework and process-centered framework. Combined, they give clarity to why, when and how students over rely on (Gen)AI tools. These insights guided the survey development.

In the following sections, each category is discussed in detail, after which the frameworks are introduced.

4.1 Contributing Factors

This subsection identifies cognitive, educational and technical factors from existing and relevant literature as well as through original theoretical reasoning.

4.1.1 Cognitive Factors

Cognitive factors refer to an individual's psychological traits and tendencies. Some of these involve the tendency to offload tasks based on objective difficulty, cognitive load (such as memory and stress) increases and metacognitive beliefs (Passi and Vorvoreanu, 2022; Risko and Gilbert, 2016). Once the user becomes dependent and is unable to offload, these factors weaken performance as the user struggles without their tools. In this section, we investigate these factors.

Passi and Vorvoreanu (2022) report that the knowledge users possess about their task's domain plays a significant role on overreliance. Domain knowledge therefore ties in with the objective difficulty of a task, where low-expertise users often accept AI recommendations at a high rate. Interestingly, both low and high domain familiarity contribute to overreliance in different ways. Low-expertise users are more susceptible due to insufficient confidence in their own judgments, while high-expertise users may become overconfident, leading to poor integration of AI suggestions. This can also be explained by GenAI's tendency to increase confirmation bias among low-expertise users (Rosbach et al., 2024).

A study by Risko and Gilbert (2016) share similar findings in their experiment. In their experiment, individuals with low confidence in their memory abilities were more probable to offload cognitive tasks since they believed that offloading would improve their performance, despite controlling for any objective difficulty. In a similar study, participants that had subsequent access to the internet were less eager to directly answer a trivia question because they could look up the answer online. The same individuals also reported a lower feeling-of-knowing compared to participants without access. Thus, the willingness to answer questions from knowledge was reduced when users had access to the internet (Ferguson et al., 2015).

These studies therefore suggest that offloading tasks is related to both differences in an individual's objective abilities, as to self-efficacy. Self-efficacy is the belief in one's capabilities to organize and execute actions to achieve specific goals (Bandura and Wessels, 1997). It influences an individual's call for action, their invested effort and their persistency (Bandura and Wessels, 1997). As we saw in the experiment of Ferguson et al. (2015), individuals with low self-efficacy are less inclined to invest effort into cognitive tasks due to the tools at hand (Internet) which can do it for them. GenAI users are exposed to the same risks, where their self-efficacy is lowered through (progressive) GenAI usage, making them believe offloading would improve their performance.

However, these studies treat self-efficacy and task difficulty as isolated predictors, without identifying underlying motivational constructs which may act as a core driver behind these behaviors.

The preference for less effort through offloading is caused by the human's design, which is to conserve as much energy as possible, both physically and mentally (cognitive effort) (Garbarino and Edell, 1997). This makes cognitive closure, which is the human desire to arrive at fast, unambiguous conclusions (Kruglanski et al., 2010), very favorable to a high degree.

Garbarino and Edell (1997) found that in situations where benefits are compared to costs, people experience more negative affect when they exert more effort in cognitively processing an alternative equivalent. Therefore the option with less costs or increased benefits is often chosen, as the benefits outweigh the costs.

Cognitive closure then explains when and why people choose to engage in or avoid cognitive engagement altogether (Kruglanski et al., 2010). One could argue that the need for cognitive closure is proportionate in magnitude to the perceived benefits of closure and the costs of lacking closure, which can vary situationally, and are represented by individual differences (Kruglanski et al., 2010).

Students with a high need for closure may be more inclined to accept AI outputs uncritically because they can reduce ambiguity and offer fast, (seemingly) authoritative answers when instructed, minimizing cognitive effort. For instance, an art student, which is typically seen as a low cognitive closure individual, could require unambiguous conclusions when faced with a deadline. Additionally, situational factors can be taken advantage of due to the accessibility of GenAI tools and its high cognitive closure potential. This can be especially dangerous in novel or complex tasks where deep analysis is required (Hiel and Mervielde, 2002).

A study by Stadler et al. (2024) is in line with this. The authors found that GenAI tools significantly lower the time and effort individuals invest in an activity, compared to other modern day offloading tools such as search engines. Consequently, a risk for students is created where users substitute this time and effort gained for useless, non-educational benefits instead of engaging in deeper critical thinking. This encourages maximizing the wrong rewards while minimizing the essential costs of study.

It has also become clear that the use of GenAI tools causes students to present justifications of lower quality (Stadler et al., 2024). Acording to Stadler et al. (2024), the interaction students undergo with challenging information is known to encourage deep engagement with the content, enhancing their learning experience. However, the study's outcome suggests that using GenAI is not cognitively challenging enough, and that students regulate their learning behavior and engage the content methodically in order to minimize cognitive effort while still passing the course.

This indicates the potential development of mental laziness due to the decreased cognitive effort. Mental laziness, according to Kriger (2024), is defined as "[...] the cognitive tendency to avoid deep or effortful thinking, opting instead for simple and superficial solutions." The described tendency makes individuals default to less resource-intensive modes of thinking and becoming prone to higher cognitive closure, while minimizing cognitive effort, which we have observed occuring in prior discussed research.

Over time, what begins as a strategic decision to offload can become an unaware habitual pattern of thinking. Utilizing the Dual Process Theory (DPT), we can hypothesize that repeated reliance on AI tools can shift cognitive engagement from deliberate to automatic processes. This is explained by the two types of thinking; system 1, which operates intuitively and relies on subconscious pattern recognition based on past experiences, and system 2 that is characterized by deliberative, analytical, and effortful reasoning (Kahneman, 2011; Tay et al., 2016). System 1 plays a vital role in everyday decision making, and is heavily influenced by situations that require quick judgments such as time constraints, stress, or familiar conditions, which students can find themselves often in. On the other hand, system 2 is invoked when situations are complex, unfamiliar, or demand careful evaluation of alternatives. This mode of thinking requires working memory and conscious cognitive resources, making it significantly slower but often more reliable (Kahneman, 2011). However, as proficiency increases, even complex tasks initially managed by system 2 may become automated instead of deliberate, and shift into the domain of system 1 through learned pattern recognition.

This renders overreliance largely unconscious, further entrenching mental laziness and shallow engagement, as each factor contributing to (Gen)AI overreliance influences and amplifies another in a continuous feedback loop manner. The loop is in line with the behavior captured in the Metacognitive Model of Cognitive Offloading, which suggests that cognitive offloading decisions are driven by (subconscious) metacognitive evaluations (Risko and Gilbert, 2016). In our loop, the mentioned factors end up determining the self-efficacy of an individual and their perceived of difficulty a task.

4.1.2 Educational Factors

The degree of cognitive offloading and engagement is furthermore shaped by educational factors. Educational factors describe the institutional, instructional, and pressure related characteristics of the student's academic environment that shape learning behaviors.

A study conducted by Dai et al. (2023) found that educational level, age, and occupation significantly contribute to the size of cognitive offloading. In their study, higher levels of deep thinking were recorded among the participants that had advanced education levels, were in managerial roles or above the age of 25. These patterns suggest that exposure to cognitively demanding environments over time fosters metacognitive discipline and confidence in personal problem-solving skills, reducing the need to rely on external tools (Dai et al., 2023) as self-efficacy is increased. In contrast, younger participants appear to lack the internal motivation or confidence to critically evaluate AI-generated responses (Dai et al., 2023), increasing their tendency to offload.

In an academic context, these findings can be extended by hypothesizing that students in analytical or text-heavy domains (e.g., law, humanities) may use GenAI tools differently than students in STEM or arts faculties. This is based on both the cognitive demands of their field and the tool's capacity of providing relevant and rich domain-specific content. The degree to which language models like ChatGPT could assist students then also differs, as these models would provide higher quality responses to students from certain study backgrounds compared to others due to its training data richness for a domain (T. Wang et al., 2023). This results in a difference in perceived satisfaction (Jo and Bang, 2023) and is therefore hypothesized to induce overreliance.

Other faculty bound factors involve pressure related characteristics. In a study by Swaroop et al. (2024), time pressure and AI overreliance are found to be positively correlated. In an experiment, individuals often accepted a decrease in accuracy at the cost of saving time. Similar results have been found by Rosbach et al. (2024), where time pressure in an AI-assisted environment increased confirmation bias, which lead to increased dependence on GenAI (Abbas et al., 2024). This is particularly relevant for students, as they are are often faced with time pressure through deadlines and high study load, which is also assumed to be different for each year of study. For instance, Dutch universities have a Binding Study Advice (BSA) for first year students, requiring students to pass a handful of courses in order to move to the next academic year and avoid expulsion.

While high study loads and academic policies like the BSA increase short-term performance pressure, the long-term effects from overreliance is hypothesized to be mainly moderated by students' coping mechanisms, self-regulation abilities, and study habits. Students lacking these skills then become more likely to adopt AI tools as reactive (system 1 thinking), rather than strategic (system 2 thinking) in aiding learning, resulting in habitual reliance even in low pressure contexts.

Another interesting factor is students' approach to learning, which is responsible for their academic performance (Duff et al., 2004; Lavin, 1965). In order to engage in meaningful learning, students must be successful in problem solving by applying novel knowledge, and by developing necessary cognitive processes (Mayer, 2002). When these constructs are avoided, students only add new information to their memories instead of understanding its deeper implications and application in problem solving. This study approach is known as rote-learning (Mayer, 2002), and results in low academic performance (Duff et al., 2004). GenAI tools, which prioritize quick and unambiguous outputs, can easily cater to this surface level processing. For students already inclined toward rote-learning, these tools offer an efficient shortcut to complete tasks without engaging in critical evaluation, reducing cognitive effort. Over time, this reinforces shallow processing habits and makes it harder for students to transition to deeper learning methods.

Lastly, Jo and Bang (2023) explain that organizational culture has the potential to influence students' view on, and intentions toward the adoption of technology. An academic institution that promotes technological openness provides students with both the encouragement and permission to experiment with AI tools. This factor could create an academic environment in which technologies like ChatGPT become embedded. However, without adequate scaffolding and digital literacy training, this can inadvertently normalize superficial usage in support of rote-learning as described prior. A balanced institutional approach should therefore teach both effective tool usage and critical reflection (Jo and Bang, 2023), enabling students to optimize GenAI use without becoming dependent.

4.1.3 Technical Factors

Technical factors capture the properties and design affordances of GenAI tools themselves, such as ease of use, accessibility, or the quality of outputs which define user satisfaction.

According to Jo and Bang (2023), there exists a correlation among university students between user satisfaction and the desire to utilize GenAI tools. The correlation delves deeper into how characteristics of the AI tools, such as interface, satisfy academic aspirations, expectations, and overall user experience. When satisfaction with the tool is high, the usage also increases (Jo and Bang, 2023). Contrarily, any malfunctions during student-GenAI interaction may disrupt their academic flow, reducing their satisfaction and resulting in reduced usage.

As satisfaction is measured differently per user, users may continue relying on these tools not due to output quality, but because the experience itself aligns with their need for efficiency and closure. For example, students may favor fast but weak answers over slower, yet accurate responses in pressured academic contexts (Swaroop et al., 2024). This behavioral trade-off reflects a preference for cognitive closure rather than critical reflection, making speed and fluency more influential than output validity. As a result, user satisfaction may not align with learning gains but rather with the perceived efficiency of the AI tool. The frequent reliance on GenAI to complete tasks with minimal friction out of convenience, may gradually reduce users' cognitive engagement, creating a feedback loop that accelerates overreliance.

As mentioned in Section 4.1.2, the quality of a model's output, which could be measured through accuracy and depth of knowledge, depends on their consumed training data (T. Wang et al., 2023). Hence, a model's performance and effectiveness is inherently tied to its training data, which varies per domain or topic. As such, GenAI tools may struggle in highly specialized or niche domains compared to a dedicated domain-specific AI tool (Dai et al., 2023).

In line with this are the AI models used. Because GenAI models vary in training breadth, they also vary in performance (B. Li et al., 2024). While high-performing, premium models are said to offer greater fluency and accuracy, their accreditation may lead users to overestimate the reliability of its outputs. Students granted access to these premium models may develop stronger trust in AI responses. We therefore hypothesize that increased technical performance correlates with a higher degree of overreliance. Even the top prestigious models can generate biased or plagiarized content, implying that satisfaction or trust alone should not be mistaken for reliability. Over time, this misplaced trust may foster habitual overreliance, especially when students begin to defer judgment to the AI without verifying results.

Yet, students unfamiliar with the described model limits may still trust the tool. This unfamiliarity has to do with the knowledge users possess about AI, and is known as AI literacy. AI literacy affects users' attitudes towards AI and is partly responsible for the development of reliance (Passi and Vorvoreanu, 2022). Based on the findings from the study of Passi and Vorvoreanu (2022), low AI literacy is often associated with an increased proneness to overreliance due to AI recommendations. This suggests that exposure to these technologies and hands-on experience through educational bodies may help in enhancing students' understanding and acceptance of GenAI (Jo and Bang, 2023), preventing or slowing down the development of overreliance.

Dell'Acqua (2022) identifies another risk tied to model quality and performance. According to him, higher quality AI tools result in users to "fall asleep at the wheel". In other words, he suggests that high quality AI tools cause individuals to become more reliant on AI output, and therefore less engaged in their work efforts based on several cognitive factors such as cognitive effort, domain knowledge, and learning approaches in response to their experience with using the AI. He concludes by saying human-AI interactions are maximized when the AI performs worse compared to industry standards, emphasizes how we need to move away from maximizing technical capabilities of new technologies and instead focus on the exerted human behaviors in response to them.

These technical factors, where many are measurable and subject to interface or policy interventions, collectively shape how students engage with GenAI.

4.2 Theoretical Foundations

This subsection outlines key theoretical models that underpin the conceptual frameworks created in this study. These theories provide a foundation for understanding students' interaction with GenAI tools, particularly in relation to the identified factors.

4.2.1 Metacognitive Model of Cognitive Offloading

The Metacognitive Model of Cognitive Offloading (MMCO), developed by Risko and Gilbert (2016), offers a useful view on examining the transfer of reliance from internal to external cognitive processing. The model addresses how individuals make choices between solving tasks mentally versus relying on external aids (cognitive offloading), which is especially relevant in modern environments where digital tools (like GenAI tools) are readily available and become increasingly efficient. In such settings, the model suggests that the more we offload, the more we're likely to continue doing so (Risko and Gilbert, 2016).

The framework's states that beliefs and experiences on how well our mind works versus how well external tools perform, is the deciding factor in choosing to offload cognitive tasks. For example, when deciding how to get to a location, we subconsciously weigh our internal memory against the

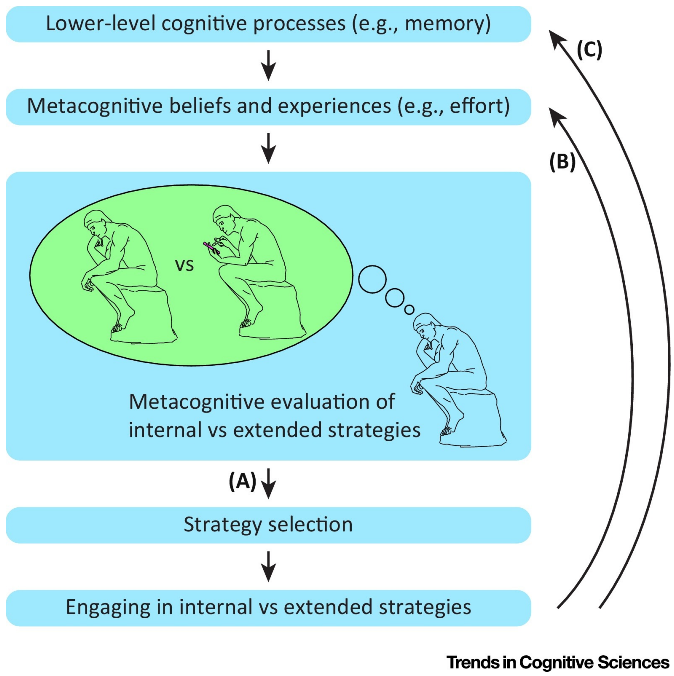

Figure 1: Metacognitive Model of Cognitive Offloading: Investigates the metacognitive aspects behind the processes that trigger cognitive offloading (Risko and Gilbert, 2016).

perceived reliability of a GPS. These judgments are often based on past experiences, and beliefs about the tool's accuracy and fluency. Then, the choice to offload may become a subconscious, automatic response triggered by familiarity or convenience rather than a deliberate decision. This process is visualized in Figure 1.

What makes this model particularly relevant to this thesis is its indirect explanatory power in relation to the underlying cognitive factors identified in Section 4.1.1. The model can be applied to provide insight into how students might grow increasingly dependent on GenAI tools, as the process of offloading impacts how they evaluate their own cognitive abilities. After repeatedly using a tool like ChatGPT for academic tasks, a student might not only rely on it more but also lose confidence in their own cognitive abilities, perpetuating a reliance loop. It therefore also reasons to reshape cognitive processes altogether.

Still, the model primarily focuses on internal evaluations and does not consider external or situational influences like time pressure, tool accessibility, or institutional culture. These influences are closely tied to the factor-centered framework in Section 4.3, which consists of the educational and technical factors identified in Sections 4.1.2 and 4.1.3. These factors are essential to understanding how and when students choose to offload cognitive effort. As such, this model served as a foundational piece in constructing the thesis' broader process-centered framework.

4.2.2 Social Cognitive Theory

The Social Cognitive Theory (SCT) offers a lens through which to explore how students interact with GenAI tools. The SCT provides a framework for understanding how humans acquire and maintain behavior, as the product of continuous interaction between cognitive, behavioral and environmental factors (Bandura, 2001; Luszczynska and Schwarzer, 2015).

At the core of the SCT is the concept of forethought, which is the human capacity to anticipate outcomes and act accordingly (Bandura, 2001). Forethought guides behavior by shaping one's expectations about the outcomes of certain actions, and one's belief in their capabilities of successfully performing those actions (self-efficacy) (Bandura and Wessels, 1997). These evaluations influence the interpretation of challenges and opportunities and determine whether to engage in a behavior. It therefore builds onto the strategy selection phase within the MMCO.

Setting goals is then responsible for behavior regulation. According to the theory, people are more likely to pursue goals if they believe the required efforts will yield favorable outcomes, and if they believe they have the capabilities to succeed. External impediments or enablers in the person's environment shape which goals are set and how reachable they appear. People with higher self-efficacy are more likely to notice opportunities in their environment compared to those with lower confidence (Bandura and Wessels, 1997).

In the context of GenAI use, the SCT helps explain how students' cognitive characteristics, expectations about its utility and the surrounding technical and academic environment combine to influence usage behaviors. For instance, high grades as a result of GenAI usage can reinforce the students' belief in its utility and their own self-efficacy, encouraging future use. On the contrary, negative outcomes may weaken these beliefs and generate aversion towards these tools as a result. By accounting for these corresponding influences, the SCT provides a grounded explanation for the influence of cognitive, educational and technical factors on GenAI overreliance.

4.2.3 Dual Process Theory

The Dual Process Theory (DPT) proposes that human reasoning and decision making are governed by two distinct but interacting systems with each representing a fundamentally different mode of cognitive processing. As described in Section 4.1.1, system 1 relies on subconscious pattern recognition to act fast and reflexive, while system 2 is characterized by effortful reasoning and engagement in novel situations. The transitions between system 1 and 2 thinking change and interact according to the task at hand and external pressure, but also as a result of an increase or decrease in proficiency (Kahneman, 2011). For instance, a student uses GenAI carefully (system 2) for their schoolwork, checking outputs critically. Yet, after seeing good grades, with little to no checking, they start trusting GenAI more at each iteration. Eventually, they begin pasting full answers without much thought, regardless of the situation (system 1).

In theory, the distinction between these two systems are deeply connected to individual cognitive differences. Factors, such as working memory capacity, influence the degree to which individuals can engage in system 2 processing (Kahneman, 2011), as they shape the frequency and effectiveness of strategy selection across both systems. Individual differences could also account for a preference in cognitive effort made, despite remarkable cognitive abilities.

Building on the DPT, we can interpret common GenAI usage among university students in academic contexts. Students who view AI as a means to maximize efficiency and reduce effort may default to system 1 thinking. These students choose to use GenAI tools out of habit or convenience, independent of the situation they are in without consideration of the educational consequences.

Contrarily, students critically evaluate the potential benefits and limitations of GenAI when engaging more reflectively with their learning materials. It could then be said that these students are likely operating in a system 2 mode. System 2 students either consciously avoid offloading to GenAI tools to preserve learning, or use it strategically. For example, instead of letting GenAI

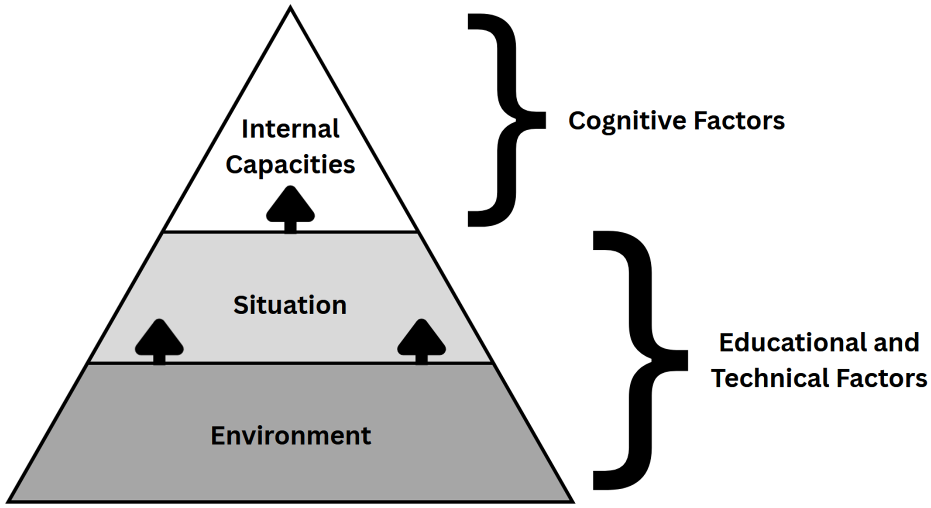

Figure 2: Factor-Centered Framework: Educational and technical factors form the base, shaping situational perceptions, which are filtered by cognitive factors to drive GenAI usage behavior.

write out an entire assignment for them, they deploy GenAI to support their writing by checking for grammar mistakes.

Therefore, the cognitive processing behind the intention of GenAI usage and the individual's learning goals define how problematic reliance on GenAI is.

However, the binary nature of the two systems risks oversimplifying the nuanced strategies students employ in real educational settings. There are more factors involved aside from character traits and high pressure situations, which are not easily reducible to internal cognitive modes. Additionally, these systems are dynamically intertwined, rather than isolated. Its complexity therefore makes it difficult to determine the precise moment when a student exchanges information between the systems or when used both in tandem.

Regardless of these limitations, the DPT extends the theory behind the strategy selection stage of the MMCO. Originally, individuals select whether to offload based on internal beliefs. By integrating the DPT and the foundations of the MMCO, we obtain a better understanding of the engaged cognitive system and whether offloading happens automatically or deliberately as we have seen in the framework.

4.3 Factor-Centered Framework

To synthesize the identified factors from the previous sections contributing to students' overreliance on GenAI tools, this section introduces a factor-centered framework. Building mainly on the SCT, this framework integrates cognitive, educational and technical influences into a vertical causal pathway that mirrors the academic contexts students find themselves in that lead to a behavioral response.

The visual taxonomy of the framework in Figure 2 reflects a cascading influence, with foundational features shaping situations, and situations interpreted through personal cognition.

At the base of this framework are the **environmental factors**, which consist of educational and technical factors. These include academic level, faculty, and curriculum, which have been found to influence offloading behaviors through differing cognitive demands (Dai et al., 2023; T. Wang et al., 2023). Additionally, the institutions' acceptance of GenAI tools and their emphasis on certain learning approaches influence the degree of engagement with GenAI (Duff et al., 2004; Jo and Bang, 2023; Lavin, 1965). In parallel, access to high performing (or premium) GenAI tools can inadvertently increase overreliance. Students may in those cases equate output fluency or technical sophistication with reliability, despite underlying limitations (B. Li et al., 2024; Passi and Vorvoreanu, 2022). These environmental factors can create specific situational contexts. For example, a user can send a limited number of prompts to an advanced GPT model within a 3-hour time frame before the model is no longer accessible.

The middle layer consists of **situational factors**, which are the occasional, temporary, context specific conditions that directly influence offloading behavior. These may arise from environmental inputs, such as deadlines, and are positively correlated with overreliance (Abbas et al., 2024; Swaroop et al., 2024). The pressures that arise also increase confirmation bias tendencies due to the positive affirmations of GenAI (Abbas et al., 2024). In these moments, students may continue turning to GenAI tools as a coping mechanism, prioritizing efficiency over understanding and accuracy (Jo and Bang, 2023; Swaroop et al., 2024). Conversely, technical difficulties can reduce user trust and disrupt flow, potentially resulting in aversion or frustration (Jo and Bang, 2023).

At the top of the framework are the **internal capacities** which ultimately drive behavioral responses, and consist of cognitive factors. These refer to the factors that influence how students perceive situations within given contexts, and therefore determine their interaction with GenAI tools. Several of these characteristics, such as domain knowledge, which modulates susceptibility to AI suggestions (Passi and Vorvoreanu, 2022) and self-efficacy, which influences reliance on external aids (Bandura and Wessels, 1997) are strongly correlated to cognitive offloading (Gerlich, 2025). Complex personality traits like cognitive effort, cognitive laziness and the need for cognitive closure, explain the underlying motivational constructs of these behaviors (Garbarino and Edell, 1997; Kriger, 2024; Kruglanski et al., 2010; Stadler et al., 2024), which may act as core drivers, especially when considering their interactions. Additionally, user satisfaction from student-GenAI interactions can reinforce habitual use regardless of output quality (Jo and Bang, 2023).

The factor-centered framework emphasizes how the influence on GenAI usage behavior progressively becomes more internalized and subjective through the vertical causal pathway. By the time environmental and their situational influences reach the top layer, it is filtered through internal capacities, which are inherently subjective and differ per individual. Therefore, vertical layout of the framework reflects how students' responses to GenAI tools are ultimately shaped by increasingly individualized factors, which is on a personal level.

4.4 Process-Centered Framework

The process-centered framework in Figure 3 is visually designed to describe how overreliance develops with each iteration, building on the components of the factor-centered framework (Figure 2). More specifically, this framework visually explains how individual, environmental and situational characteristics influence students' tendency to rely on GenAI tools. These dynamic factors shape the students' interpretation of their learning context and lead to a behavioral response. The framework is inspired by, but distinct from, the MMCO (Risko and Gilbert, 2016) and applies insights from the DPT (Kahneman, 2011) and SCT (Bandura, 2001) to tailor to educational contexts.

The factor-centered framework defined in Section 4.3 shapes the input characteristics of

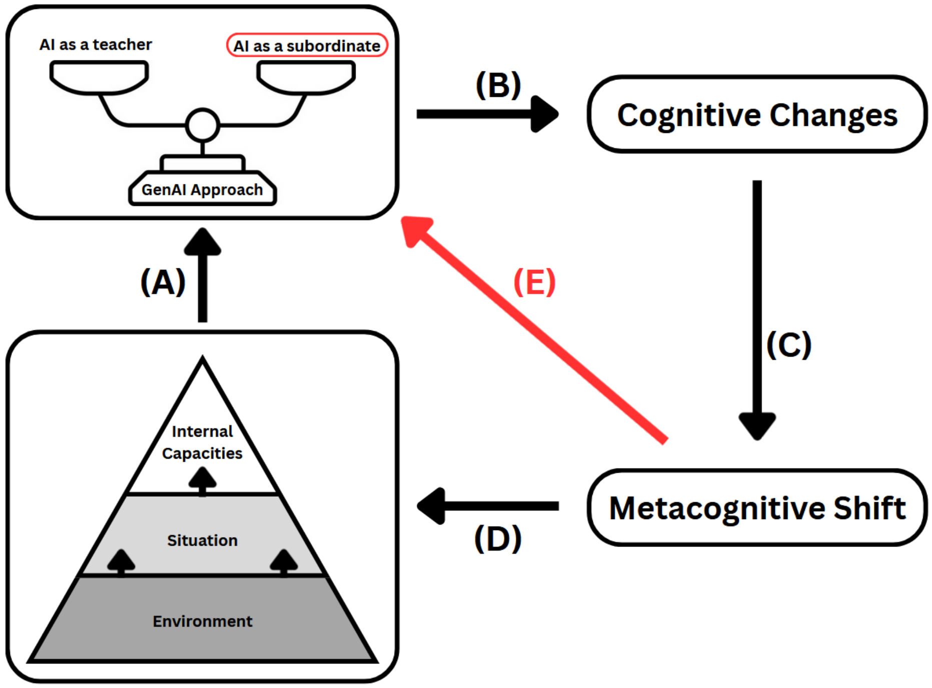

Figure 3: Process-Centered Framework: The different interactions between offloading and the responsible dynamic mechanisms that lead to GenAI overreliance.

the process-centered framework. These characteristics are divided into **internal capacities**, **environmental factors** and **situational factors**, with their combined effect shaping how students perceive academic tasks and respond to them. Their interpretations then ultimately inform their tendency to offload (certain) cognitive tasks to GenAI tools. In the framework's logic, environmental factors generate specific situations, while cognitive abilities determine how these situations are interpreted (arrow A).

Influenced by the student's state of mind, tool accessibility and academic urgency, two approaches to GenAI usage are proposed by Cleiren et al. (2025), which are inspired by the DPT:

1. **AI as a subordinate**: prioritize speed and resource conservation to minimize cognitive effort (system 1 thinking).

2. **AI as a teacher**: prioritize mastery and learning to ensure internalization of study material (system 2 thinking).

The approach determines whether students are inclined to learn or simply get the task done. Acting as a dynamic mediator, the model visualizes how students can (consciously) reflect and switch between mindsets (systems 1 and 2). However, this does not imply that students are always conscious of their decisions. The flexibility to change in approach depends on the student's awareness and reflective ability when faced with a choice of approach. Subsequent evaluations from prior successes and beliefs about performance and reliability give weight to usages preferences. Depending on the interpretation, the extent of cognitive offloading is determined.

Cognitive offloading, as mentioned in Section 4.1.1, changes cognitive processes such as thinking, memory retention, understanding, and engagement. Building on the SCT, downstream effects from the selected cognitive offloading approach can have both costs and benefits to a student's motivation to studying, academic performance, etc. (arrow B). Additionally, experience and opportunities from the executed offloading affect a student's perception on the tool's usefulness and reliability, as well as internal capabilities such as self-efficacy. Triggers, such as grades and free time, cause students to unknowingly evaluate their approach to studying. This forms a metacognitive shift in beliefs about GenAI's utility and one's internal capabilities (arrow C), and influences (unaware) future behavioral responses (arrow D) in the form of a preference for GenAI approach. However, the degree and pace of the metacognitive shift is different for each student, as their motivation and context behind GenAI usage is dynamically impacted by the factors outlined in the factor-centered framework (Figure 2), combined with external triggers. For instance, a students' metacognitive shift may be slow and gradual due to habituation (internal capacities), but occurs fast in situations that require immediate attention (deadlines).

Students approach a critical threshold. After this threshold, patterns of offloading become increasingly resistant to change. Once a student repeatedly and excessively offloads, they eventually skip all evaluation in GenAI usage and reflexively turn to using it as a subordinate, which is where overreliance lies (arrow E). This is especially reinforced by successful outcomes, as students may begin to develop a false sense of trust in the tool's reliability. The misplaced credibility makes students more likely to offload cognitive tasks to GenAI tools (Gerlich, 2025), further giving in into the cycle.

Even though using GenAI as a teacher still leads to offloading, it does so more strategically and consciously, often when the student evaluates the trade-offs. The danger is when shortcuts prioritizing efficiency become the default behavior, bypassing reflective reasoning altogether. That's where cognitive offloading evolves into overreliance.

Excessive offloading then becomes habitual, even when the situation might actually benefit more from engaging in learning or critical thinking. This results in the repeated decrease in cognitive thinking (Gerlich, 2025), which leads to disregarding the evaluation of future study approaches or academic engagement.

These mechanisms lay the foundation for understanding the dynamic components of the conceptual process-centered framework, and highlight the potential interactions between offloading and the responsible mechanisms.

5 Survey Setup

In this section, the design and structure of the survey is described. It outlines the survey questions and the distribution channel used to ensure a diverse and representative sample. Additionally, insights from a pilot study are discussed, which shaped the final version of the survey to improve clarity, relevance and reliability.

5.1 Survey Design

To empirically evaluate the conceptual frameworks, a survey was developed and distributed to assess the extent of GenAI overreliance among students in Dutch higher education institutions. The survey was primarily designed to measure GenAI overreliance as the dependent variable, and the cognitive, educational and technical factors as the independent variables contributing to this

Question	Label	Answer options
At what type of university are you currently registered?	`demograph_uni_type`	Dutch research university (wo), Dutch university of applied sciences (hbo)
What is your current study programme level?	`demograph_study_level`	Bachelor programme, Master programme, Pre-master programme, PhD programme
What is the name of your current study programme?	`demograph_study_name`	Open question
Which field of study best matches your current study programme?	`demograph_study_field`	Natural Sciences, Social Sciences, Humanities, Engineering and Technology, Business and Economics, Medicine and Health Sciences, Law and Governance, Arts and Design, Other

Table 1: Overview of demographic survey questions, corresponding variable labels, and answer options.

phenomenon. This was achieved by dividing the survey into multiple components, each empirically assessing the mentioned variables through mandatory questions.

The introductory section of the survey aimed to gather essential demographic information about the participants, as displayed in Table 1. This information enabled a deeper understanding of the relationships between academic background and the measure of GenAI overreliance.

Overreliance is measured by examining GenAI usage patterns in academic settings. Specifically, the survey differentiated between compensatory and constructive use, reflecting the approaches outlined in the process-centered framework respectively (see Section 4.4). Compensatory usage represented a reliance on GenAI to bypass learning efforts and academic responsibility, substituting critical thinking. Constructive usage, on the other hand, was characterized by leveraging GenAI to enhance understanding and drive academic growth, aligning with effective learning strategies.

To accurately capture these distinction, the survey first explored when students tend to use GenAI the most during their studies in order to determine where GenAI overreliance is most prevalent. Participants were presented with situational prompts (Table 2), which were measured on a Likert scale ranging from "Never" to "Every Day".

Then, the survey presented items in paired, generalized statements that reflect both constructive and compensatory behaviors. These items were rated on a Likert scale ranging from "Strongly Disagree" to "Strongly Agree", allowing for a quantitative assessment of dominant usage patterns that indicate varying levels of overreliance. Unlike the contextual items, these questions reflected the overall attitudes and habits related to GenAI usage in academia (Table 3).

This way, the survey drew a clear distinction between when GenAI is used most frequently, and how it is used as academic aid, which allowed for a nuanced analysis of overreliance.

To complement the analysis, the survey attempts to explain usage behavior through the influence of cognitive, educational and technical factors on overreliance. These items were measured through targeted Likert-scaled items, ranging from "Strongly Disagree" to "Strongly Agree", that assessed their degree of influence. These questions can be seen in Table 4, 5 and 6.

Question	Label
When I have plenty of time to study for an exam.	`frequency_plenty_of_time`
When I have a lot of deadlines.	`frequency_many_deadlines`
When I'm studying a new topic.	`frequency_new_topic`
When I feel confident in my understanding of the study material.	`frequency_confident_understanding`
When the study material is overwhelming.	`frequency_overwhelming_material`
When I am familiar with the study topic.	`frequency_familiar_topic`
When search engines provide ambiguous results for something I need for my study.	`frequency_ambiguous_search`
When I don't have any homework or assignments planned.	`frequency_no_homework`
When I have to study for an exam.	`frequency_exam_study`

Table 2: Overview of questions assessing the frequency of GenAI usage across various study situations

To explore the development of overreliance over time, the survey included items designed to capture shifts in usage through temporal comparisons (Table 7). This enabled an analysis of whether students' dependency on GenAI has changed as they progress through their studies.

By capturing contextual usage behaviors, dependency patterns and overreliance development, the survey aims to provide a holistic view of GenAI overreliance.

5.2 Participants and Distribution

To ensure a representative and diverse sample, participants for the survey were recruited from Dutch higher education institutions. These institutions, including research universities and universities of applied sciences, encompass a broad range of study fields and learning environments, and provided a comprehensive view of GenAI usage patterns. To allow for accurate assessment of these patterns and reliance behaviors, students from these universities required using GenAI tools regularly in an academic context to qualify for participation.

Distribution of the survey occurred through multiple channels to maximize reach and accessibility. These channels involve institutional communication channels and social media platforms such as LinkedIn and Instagram. With the help of my supervisors, the survey had been shared with teachers from other faculties and universities, with the survey being spread through announcements on study course pages. This multi-channel distribution strategy was intended to capture a diverse pool of respondents across different study fields and academic levels, in order to enhance the generalizability of the findings and allow for stratified analysis.

The survey was conducted anonymously to maintain ethical research standards. A privacy statement was included at the beginning of the survey, clearly explaining that all responses are confidential, and were used solely for research purposes.

Additionally, participants were given the option to receive a copy of the research findings after completion of the survey. However, to maintain anonymity, this process was entirely decoupled from the survey. This respects the rights and confidentiality of all participants, with no collection or storage of personal data that can be used to trace back to individual participants.

Question	**Label**	**Form of Usage**
I use GenAI tools to summarize texts with the goal of deepening my understanding, rather than bypassing the reading process.	`overreliance_summarize_for_understanding`	Constructive
I find myself using GenAI to quickly get through study materials, even if it means not fully understanding the content.	`overreliance_quick_through_material`	Compensatory
I accept GenAI-generated explanations without double-checking them against the original material.	`overreliance_accept_without_verification`	Compensatory
When I struggle with difficult problems, I use GenAI to understand the solution process, not just to get the correct answer quickly.	`overreliance_solve_vs_understand`	Constructive
GenAI assists me in improving my assignments by enhancing clarity, rather than completely relying on it to do the work for me.	`overreliance_clarity_not_dependence`	Constructive
I use GenAI-generated content directly in my assignments without verifying its accuracy.	`overreliance_direct_use_without_verification`	Compensatory
When GenAI's output is incorrect, I attempt to understand the mistake and correct it, rather than relying solely on the tool to fix it.	`overreliance_correct_mistakes_vs_rely`	Constructive
When GenAI suggests solutions, I critically analyze its suggestions before integrating them into my work.	`overreliance_critical_analysis_suggestions`	Constructive
I rely on GenAI more than I should for academic tasks, even when I am capable of completing them independently.	`overreliance_more_than_should`	Compensatory

Table 3: Items assessing compensatory and constructive usage patterns of GenAI tools in academic settings to evaluate overreliance.

Question	**Label**
When using GenAI tools, I feel like I save time by avoiding unnecessary effort in understanding complex topics.	`cognitive_effort_avoidance`
I prefer using GenAI for quick explanations rather than dissecting the study material myself.	`cognitive_quick_explanations_preference`
I rely on GenAI to fill knowledge gaps, even when I could attempt to understand the content myself.	`cognitive_fill_gaps_without_attempt`
I feel lazy when I use GenAI instead of manually working through study problems.	`cognitive_lazy_instead_of_manual`
I feel more confident in my work when GenAI confirms what I already think, even if I haven't fully verified the information.	`cognitive_confirmation_bias`
I find that I can solve study-related problems without heavily relying on GenAI tools.	`cognitive_solve_without_reliance`

Table 4: Cognitive factor items assessing internal motivations for using GenAI tools in academic contexts.

Question	**Label**
Deadlines influence my decision to use GenAI tools as a way to save time and complete tasks efficiently.	`educational_deadline_influence`
The university's pressure to perform well on exams sometimes leads me to use GenAI as a shortcut rather than fully engaging with the material.	`educational_performance_pressure`
I feel that the content and pace of my curriculum make GenAI a necessary part of my study routine to keep up.	`educational_curriculum_dependency`
My university provides clear guidelines on the appropriate use of GenAI tools in academic settings.	`educational_clear_guidelines`

Table 5: Educational factor items assessing the academic environment and study pressures related to GenAI usage.

Question	Label
I subscribed to a premium GenAI model because I am convinced it produces more reliable and accurate responses.	technical_premium_subscription
When GenAI tools experience technical issues or downtime, it significantly disrupts my study habits.	technical_downtime_disruption
I know how to write effective prompts for GenAI and consistently get useful responses in return.	technical_prompting_skill
GenAI tools seem to have sufficient knowledge of my study field, making me feel comfortable relying on them for study support.	technical_field_knowledge

Table 6: Technical factor items assessing user expertise and system reliability in the academic use of GenAI.

Question	Label
Improvements in GenAI's reasoning capabilities over the past years have led me to use it more frequently for academic tasks.	change_reasoning_improvement
Increases in the quality of GenAI-generated content have made me more confident in relying on its output for study purposes.	change_content_quality
Access to premium GenAI models has encouraged me to use these tools more often in my studies.	change_premium_access
As GenAI tools have become more efficient and accessible, my study habits have increasingly incorporated their use.	change_efficiency_access

Table 7: Overview of questions assessing the influence that improvements in GenAI capabilities and access have on GenAI usage over time.

5.3 Pilot Study

Prior to distributing the survey to the target group, a pilot study was conducted with a small sample of six participants from different study fields, universities and academic levels. After completing the draft survey, participants were asked to provide feedback through a feedback form. This form consisted of the following questions:

- **Were the questions clear and easy to understand? If not, which ones were unclear?** All participants found the questions clear and easy to understand. One participant noted that some questions seemed repetitive, and another mentioned they sometimes felt ambiguous and lacked a "depends" option.
- **Did you find the questions relevant to the topic? Are there any you would suggest adding or removing?** All participants agreed that the questions were relevant. However, one participant expressed the need to elaborate their responses. Another suggested adding questions about the institution's stance on AI usage.
- **Was the length of the survey appropriate? If not, do you feel it was too long or too short?** The average completion time of the survey was 5 minutes. Most participants felt the survey length was appropriate, with one participant suggesting it could be even longer before becoming "annoying".
- **Were there any technical issues or difficulties navigating through the survey?** All participants reported no technical issues.
- **Do you have any additional suggestions or feedback that could improve the survey?** Feedback was diverse. One participant suggested adding a brief explanation of GenAI at the beginning of the survey to enhance understanding. Another found it difficult to estimate weekly GenAI usage, proposing a "depends" option. Other suggestions involved making more diverse questions and offering a Dutch translation for broader accessibility.

The diversity of the group provided a broad range of perspectives, enabling the identification of potential ambiguities and areas for improvement. The obtained feedback led to several refinements, such as rephrasing questions and the addition of an explanation of GenAI, which substantially contributed to the assurance of the survey's quality and its results. Furthermore, the final survey questions, as presented in Section 5.1, are the product of iterative improvements based on the pilot study and feedback analysis. These final versions were presented to the target population.

6 Survey Results

This section presents the results of the survey analysis, structured into three parts:

- Descriptive analysis: establishes a baseline by examining how frequently students use GenAI, how this usage has changed over time, and the general patterns of constructive vs. compensatory reliance.

- Correlation analysis: investigates the strength and direction of associations between GenAI overreliance and various situations, as well as the cognitive, educational and technical factors.
- Regression analysis: identifies which of the researched factors significantly predict overreliance, providing empirical support for answering the research question.

This sequential approach facilitates a clearer interpretation of the gathered results, which allows for the grounded understanding of key relationships proposed in the conceptual frameworks (Figure 2 and 3).

In total, 99 responses were collected from the survey. Eight responses were excluded from the analysis, as respondents indicated not using GenAI for academic purposes under any of the suggested circumstances in Table 2. The final sample therefore consists of 91 valid responses (n = 91), which provides a solid basis for the analyses presented in this section.

6.1 Descriptive Analysis

To establish a foundational understanding of overreliance among Dutch university students, this section explores situations that foster GenAI usage, changes in usage over time and the nature of students' reliance on GenAI. These insights are drawn from survey responses to the questions listed in Tables 1, 2, 3 and 7. The distribution of responses to each question are highlighted in the figures shown in Appendix A.

6.1.1 Demographic Overview

Figure 4, corresponding to Table 1, presents an overview of the respondents' academic backgrounds in terms of study level (`demograph_study_level`), field `demograph_study_field` and university type (`demograph_uni_type`). The majority of participants are enrolled in Bachelor programs at research universities (wo) within the Engineering and Technology domain. This domain includes disciplines such as Computer Science, Engineering, Mathematics and Architecture.

6.1.2 Usage Frequency across Contexts

Figure 5, which is tied to the questions in Table 2, illustrates how GenAI usage frequency differs per academic scenario.

Unsurprisingly, students report lower usage during (most) periods with minimal academic pressure. These involve periods of no homework (`frequency_no_homework`), familiar study material (`frequency_familiar_topic`) and students' high confidence in their understanding (`frequency_confident_und`
In contrast, usage increases in more demanding situations such as pressure from deadlines (`frequency_many_deadlines`), unfamiliar topics (`frequency_new _topic`), ambiguous web results (`frequency_ambiguous_search`), and overwhelming study material (`frequency_overwhelming_material`). These patterns reinforce the notion that GenAI often functions as an academic coping mechanism when the demand for mental effort is high, time is short or when clarity is lacking.

Interestingly, some students also report frequent GenAI use when they have plenty of time to prepare for exams (`frequency_plenty_of_time`). While this may appear counterintuitive, a plausible interpretation is that the time leading up to exams is often used for preparatory tasks. Hence, GenAI may offer valuable assistance in assignments, projects, etc. Furthermore, some

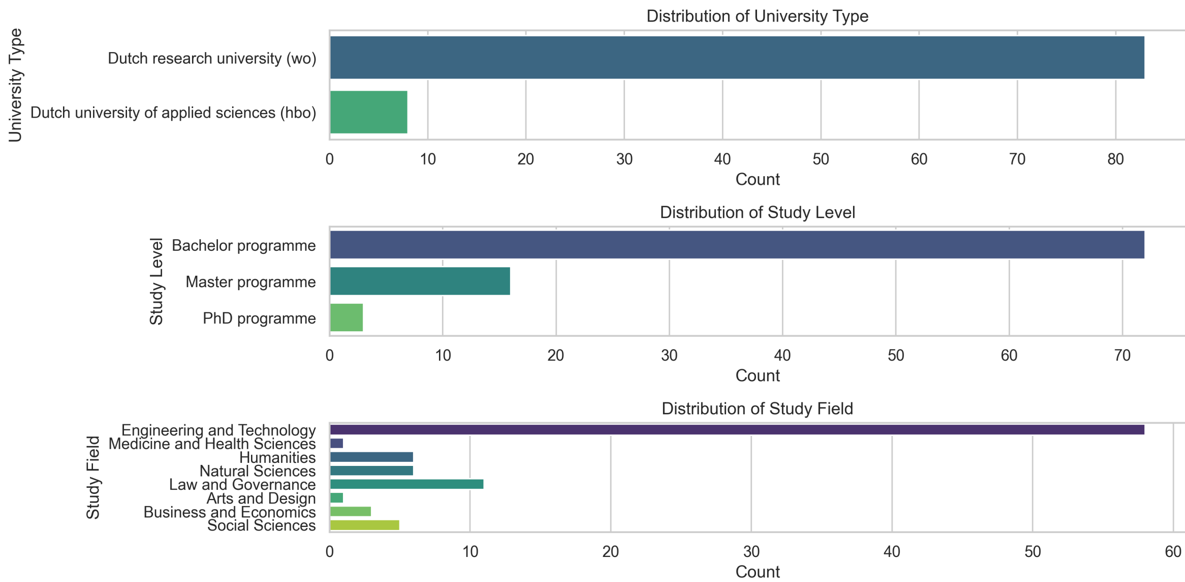

Figure 4: Distribution of respondent demographics, including university type, study level, and study field.

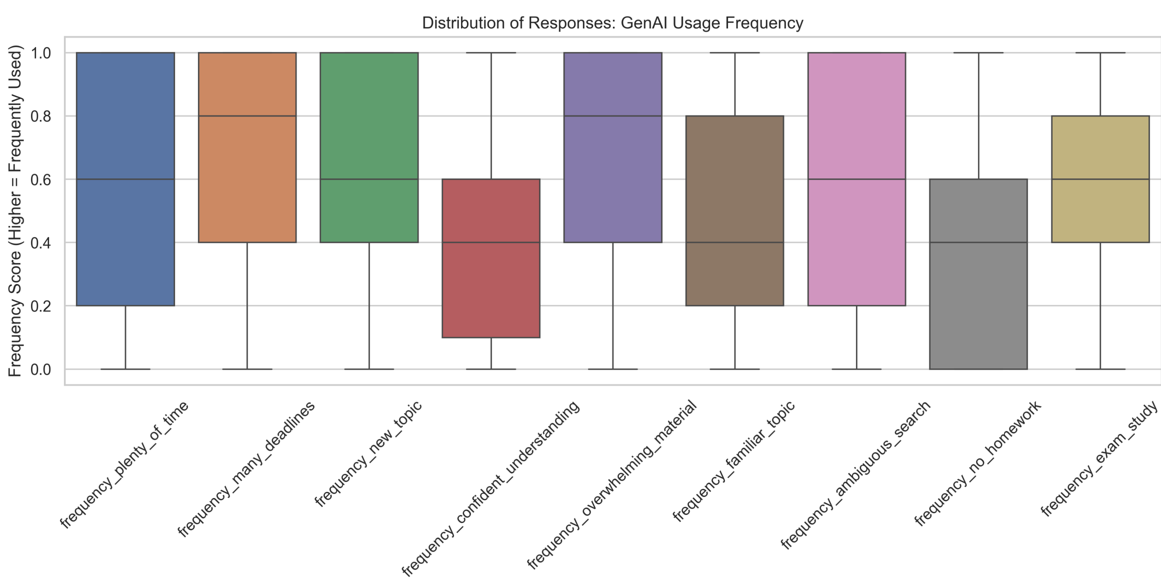

Figure 5: Distribution of GenAI usage frequency across academic contexts.

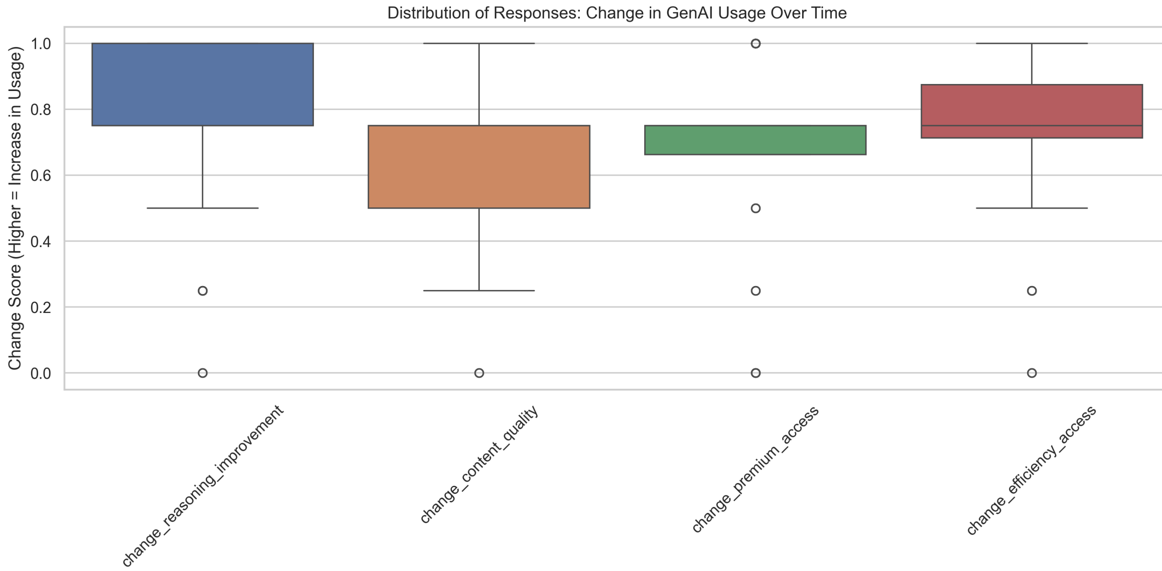

Figure 6: Reported changes in GenAI usage over time.

students may use GenAI to structure their study sessions or improve their study notes in advance. Either case, this finding suggests that GenAI is not merely a tool of last resort, but part of study habits for some students. Another plausible interpretation is that of students being genuinely interesting in certain courses they take. In those cases, students could resort to GenAI tools in order to explore different areas of the subject at hand, which weren't part of the course material.

Another noteworthy observation is the reduced frequency of GenAI usage when students prepare for exams (`frequency_exam_study`). This may suggest that students do not primarily view GenAI as a tool for active learning or long-term understanding. Instead, its use during exam preparation may be limited to occasional tasks such as simplifying complex concepts or generating summaries. This pattern aligns with the nature of exams, which typically assess a student's independent mastery and application of knowledge. Given that the use of GenAI is strictly prohibited in examination settings, students may deliberately reduce their reliance on it during preparation to avoid becoming dependent. In doing so, they consciously distance themselves from tools they cannot use during the actual assessment.

6.1.3 Usage Over Time

Figure 6 (referencing Table 7) shows that the overall usage of GenAI has increased over time. Based on responses, students primarily attribute this growth to improvements in the reasoning capabilities (`change_reasoning_improvement`) of GenAI models, and enhanced accessibility and efficiency (`change_efficiency_access`). Additionally, students with access to premium models indicated that the premium models encouraged them to use GenAI tools more often in their studies (`change_premium_access`). Perhaps this can be attributed to the belief that these models significantly contribute to academic performance, or simply indicate unconscious sunk cost fallacy.

While improvements in output quality (`change_content_quality`) were expected to be the most influential driver of adoption, respondents deemed it to be the least important. This raises

Table 8: Average scores per factor for compensatory vs. constructive GenAI users. T-tests are conducted on a 95% confidence interval.

Factor	Mean Comp.	Std Comp.	Mean Const.	Std Const.	Mean Diff.	p-value
`cognitive_lazy_instead_of_manual`	0.64	0.27	0.48	0.33	0.16	0.01
`educational_clear_guidelines`	0.60	0.27	0.48	0.37	0.11	0.09
`cognitive_fill_gaps_without_attempt`	0.58	0.28	0.47	0.33	0.11	0.09
`educational_deadline_influence`	0.70	0.27	0.60	0.33	0.10	0.11
`educational_curriculum_dependency`	0.50	0.33	0.41	0.32	0.09	0.20
`cognitive_effort_avoidance`	0.65	0.29	0.56	0.34	0.09	0.20
`cognitive_quick_explanations_preference`	0.60	0.29	0.53	0.30	0.07	0.26
`educational_performance_pressure`	0.51	0.33	0.44	0.34	0.07	0.35
`technical_downtime_disruption`	0.36	0.31	0.31	0.28	0.05	0.41
`technical_field_knowledge`	0.68	0.23	0.63	0.28	0.05	0.39
`technical_premium_subscription`	0.57	0.30	0.56	0.34	0.01	0.90
`cognitive_confirmation_bias`	0.60	0.30	0.62	0.26	-0.01	0.82
`technical_prompting_skill`	0.76	0.20	0.79	0.19	-0.04	0.40
`cognitive_solve_without_reliance`	0.74	0.20	0.84	0.17	-0.11	0.01

concerns to whether students are increasingly relying on GenAI because it produces accurate results of high quality, or because it is fast, convenient and always accessible.

The latter highlights the associated risks discussed in Section 4.1.3, where a growing preference for fast answers may gradually undermine students' willingness to engage in deeper learning processes. When efficiency and fluency become the dominant criteria for user satisfaction, resorting to GenAI may be less about meaningful academic progress and more about reducing cognitive effort.

6.1.4 Nature of Overreliance

To better understand the extent of GenAI overreliance, we turn to responses corresponding to Table 3 which is illustrated in Figure 8. In this analysis, positive statements reflecting constructive usage were inverted so that higher overall scores consistently indicate stronger overreliance.

Overall, students tend to (strongly) disagree with statements that signal overt misuse of GenAI (Figure 8), suggesting a basic awareness of its limitations and risks. Notably, the mean overreliance score across all students is approximately 0.297, suggesting a predominantly constructive use of GenAI in academic contexts. However, this general trend masks more subtle patterns of dependency.

To distinguish between different usage patterns, students were classified into constructive users, and compensatory users, based on whether their personal overreliance score falls below or above the total average overreliance score of 0.297, respectively. For each factor assessed in Tables 4, 5 and 6, we calculated the average response within each group. The resulting difference in means reveals which characteristics are more associated with either constructive or compensatory usage patterns, and is summarized in Table 8.

Based on the results, the largest differences are found in cognitive and educational factors, suggesting these may play a key role in determining whether GenAI use becomes compensatory or

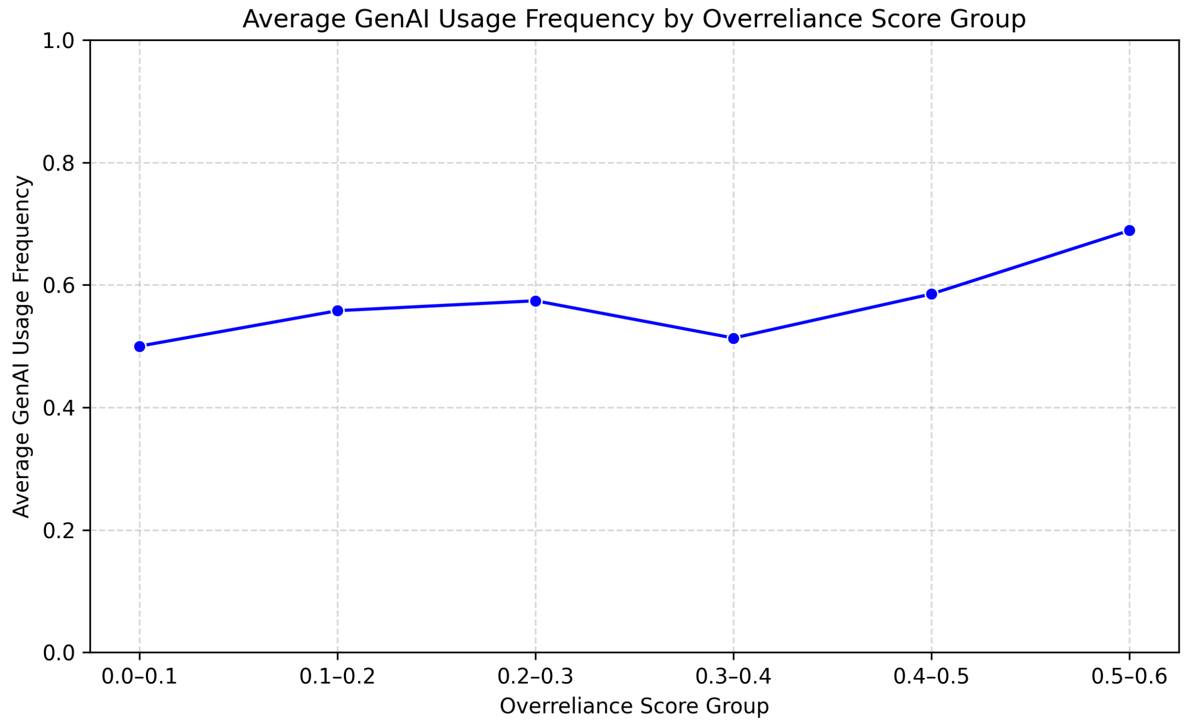

Figure 7: Average GenAI usage frequency usage per group of overreliance score.

remains constructive. Among compensatory users, the strongest indicator appears to be a tendency to use GenAI out of convenience, as this group scores significantly higher on statements assessing cognitive laziness (`cognitive_lazy_instead_of_manual`) and cognitive effort (`cognitive_fill_gaps_without_attempt`). In contrast, constructive users are more likely to report a high degree of self-efficacy when completing academic tasks (`cognitive_solve_without_reliance`).

Interestingly, the factor with one of the smallest differences is confirmation bias (`cognitive_confirmation_bias`), suggesting that both groups are equally susceptible to accepting GenAI outputs that align with their own views or expectations, rather than questioning it. This may point to a general lack of awareness about the tendency of GenAI tools to reinforce user assumptions, regardless of usage.

In terms of educational factors, compensatory users seem to agree more strongly that their institution offers clear guidelines on how to use GenAI appropriately (`educational_clear_guidelines`). These students also feel more strongly influenced by deadlines (`educational_deadline_influence`) and therefore resort to GenAI. Technical factors however, show relatively minor differences between the two groups. While modest, the largest gap is observed in perceptions of whether GenAI tools posses sufficient field knowledge (`technical_field_knowledge`).

Another notable pattern is observed when comparing GenAI usage frequency to scores from the overreliance indicators (Figure 7). We notice that an increase in GenAI usage frequency corresponds to a higher overreliance score. This relationship suggests that repeated use of GenAI may gradually foster dependency, regardless of constructive or compensatory usage. As discussed by Jo and Bang (2023), when students find GenAI tools efficient and satisfying, they are more likely to use them frequently, reinforcing habits that may tilt toward overreliance.

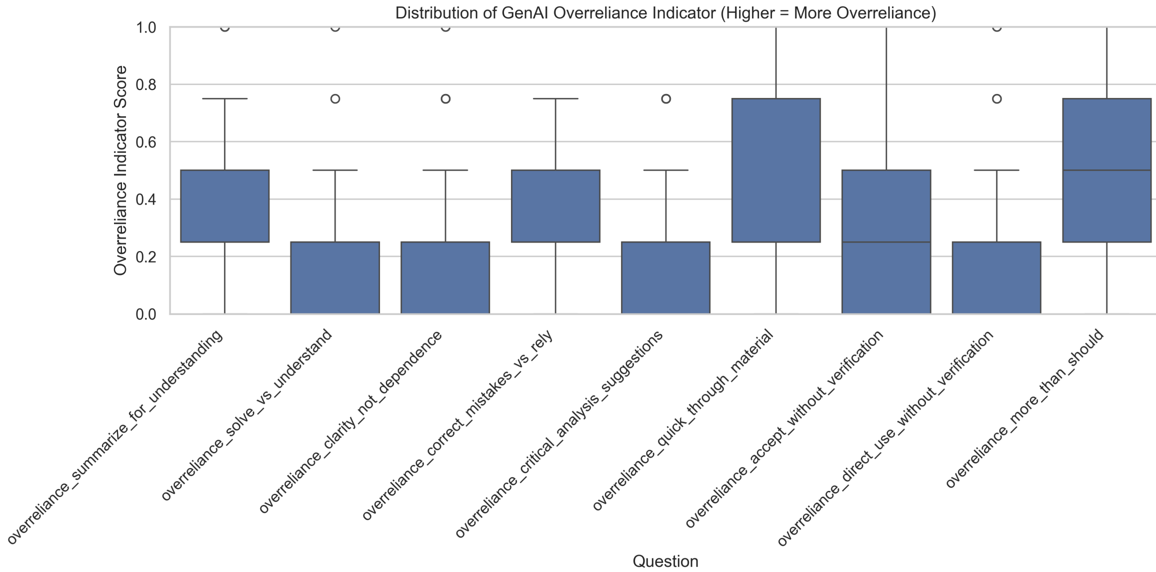

Figure 8: Distribution of responses related to GenAI overreliance. Higher scores reflect compensatory usage, and therefore overreliance.

When looking at Figure 8, we gain a deeper understanding of which indicators contribute the most to overreliance. Many students admit to using GenAI to quickly get through study materials at the expense of understanding (`overreliance_quick_through_material`). This behavior reflects the need for cognitive closure, to which GenAI precisely caters by offering fast, confident, and often authoritative outputs that reduce ambiguity. Yet, these outputs are often accepted at face value rather than encouraging students to reflect or verify them. As a result, GenAI may be reinforcing rote-learning rather than supporting meaningful learning as discussed in Section 4.1.2. This idea is further supported by our findings showcasing (high) responses to questions about deepening understanding through summarizing study material (`overreliance_summarize_for_understanding`, this question has been inverted, meaning students summarize study materials to bypass the reading process), correcting mistakes in GenAI output (`overreliance_correct_mistakes_vs_rely`) and uncritically accepting outputs (`overreliance_accept_without_verification`), which also indicate signs of overreliance. While rote-learning doesn't develop transferable knowledge, they are effective in the short-term for passing exams with minimal cognitive effort. When exams reward recall over reasoning, students may see little incentive to invest in deeper understanding, viewing it as a burden. Hence, the style and structure of university assessments could be another contributing factor in determining GenAI usage among students for future research.

Another notable finding is that, despite the apparent constructive usage of GenAI among university students, a high degree of students report relying on GenAI even when they are fully capable of completing assignments independently (`overreliance_more_than_should`). This signals another emerging behavioral pattern, in which students develop a habitual reliance on GenAI. The habitual reliance causes students to integrate GenAI so deeply into their study, that it becomes their default resource, or substitute, for critical thinking. This aligns with the process-centered framework introduced in Section 4.4, which conceptualizes overreliance as a gradual development. Students may not perceive their usage as problematic since they frame it as constructive. Yet, it is the same students that simultaneously acknowledge a growing dependence.

This reinforces the idea that overreliance can coexist with intentional and strategic GenAI use. Students are learning to learn with GenAI, and over time, they optimize their behavior to extract maximum benefit (e.g., reduced study time and passing grades by only "remembering") while minimizing costs (e.g., cognitive effort spent understanding the material). While efficient, these behaviors raise questions about long-term learning outcomes.

While this descriptive analysis reveals clear distinctions between user type characteristics, it does not uncover how these factors interact, nor does it establish which variables are most important for predicting overreliance.

6.2 Correlation Analysis

This section delves into the strength and direction of the relationships between GenAI overreliance and various influencing factors. By examining these correlations, we aim to understand the relationship between several factors and overreliance in advance of the regression analysis in Section 6.3.

For each correlation matrix presented in the following sections, the responses for positive statements reflecting constructive usage were not inverted. Additionally, an extra column is added which displays the average absolute correlation between overreliance and each assessed question. This provides a summarized indicator of relevance, regardless of whether the correlation is positive or negative. The relevant survey questions are listed in Tables 2, 3, 4, 5 and 6. Finally, to preserve model performance for this analysis, empty values for certain questions have been replaced with the mean of the corresponding filled values from other responses, instead of being left out.

6.2.1 GenAI Usage Frequency and Overreliance

Figure 9 displays the correlation matrix between students' reported GenAI usage frequency across different academic situations (Table 2) and their self degree of reliance on the tool (Table 3), revealing several trends.

One of the most consistent patterns is that students demonstrate both constructive and compensatory GenAI usage, regardless of study situations. Some usage indicators do however highlight the difference based on the situation. For example, students who frequently use GenAI in low pressure academic situations, such as having no homework (`frequency_no_homework`) or being confident in their understanding (`frequency_confident_understanding`), also tend to engage with the tool in slightly more constructive ways compared to high pressure situations. In these situations, students often report using GenAI to deepen their understanding of summarized content (`overreliance_summarize_for_understanding`) or to clarify complex concepts (`overreliance_clarity_not_dependence`) rather than simply bypassing the learning process.

Contrarily, the constructive usage slightly diminishes under pressure. When faced with high academic demands such as deadlines (`frequency_many_deadlines`) and ambiguous search engine results (`frequency_ambiguous_search`), students' usage of GenAI gains compensatory intent. Instead of using GenAI effectively to improve their learning, they prefer GenAI for its speed and ease.

This same duality is mirrored in responses to questions assessing whether students rely too heavily on GenAI during difficult periods. Students admit to leaning more on GenAI than they

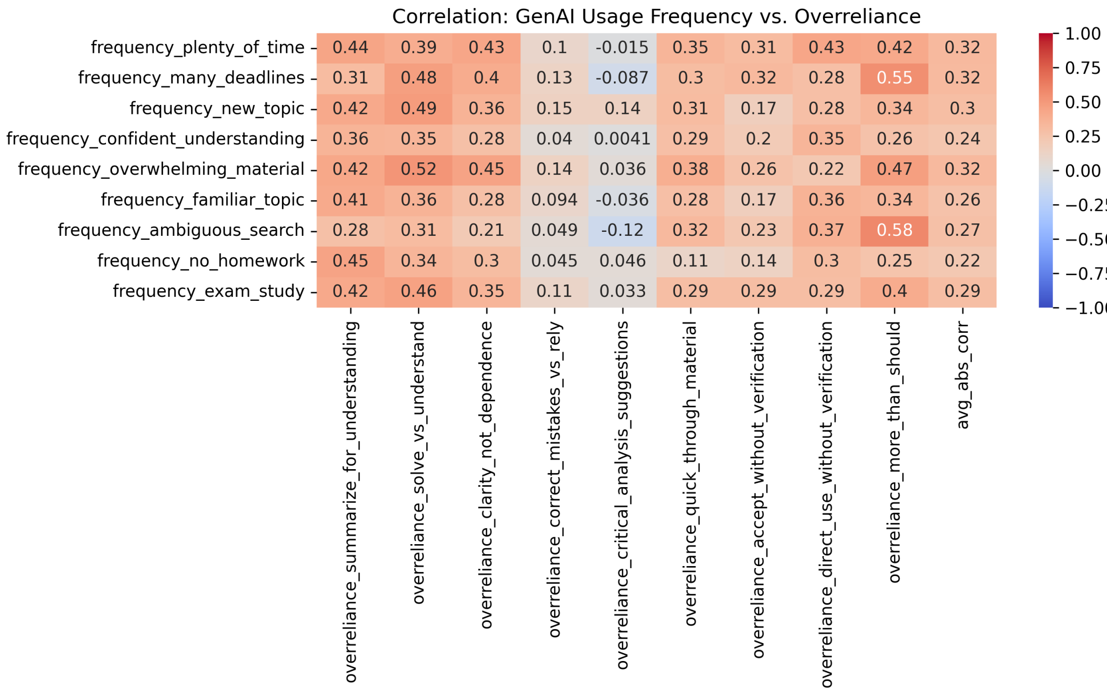

Figure 9: Correlation matrix between questions assessing the frequency of GenAI usage and questions assessing the degree of overreliance.

believe is appropriate (`overreliance_more_than_should`) during stressful periods, while reporting less reliance during more relaxing periods. These behavioral patterns are in line with the theoretical discussion in Section 4.1.1 regarding the need for cognitive closure. The theory states that individuals experiencing stress are more likely to favor fast answers over deeper cognitive engagement. In such cases, GenAI becomes a tool for alleviating immediate pressure.

Interestingly, the data reveal an absence of correlation between frequency-oriented situational questions and students' tendencies to correct GenAI output (`overreliance_correct_mistakes_vs_rely`) or critically reflect on its suggestions before use (`overreliance_critical_analysis_suggestions`). This surprising neutrality might suggest the presence of a difference in the individual's characteristics. Students may fall into either the group that routinely critically engages with GenAI, or those who don't, independent of circumstance. Another explanation could be based on prior interactions students may have had with GenAI tools, where successful interactions influence their decisions to neglect verifying outputs. This behavior is especially reinforced by external triggers in the shape of rewards such as good grades, as described in the process-centered framework (Section 4.4. Over time, such habits could become resistant to change even when context shifts.

Another notable finding lies in the correlation patterns associated with the exam preparation period. Contrary to what might be assumed, students who report having plenty of time to study for an exam (`frequency_plenty_of_time`) show relatively high levels of both constructive and compensatory usage. This suggests a contradiction in the expectation that GenAI usage in lower pressure situations is not uniform across students. Some students appear to use their time to deepen understanding (`overreliance_summarize_for_understanding`), while others use GenAI's output

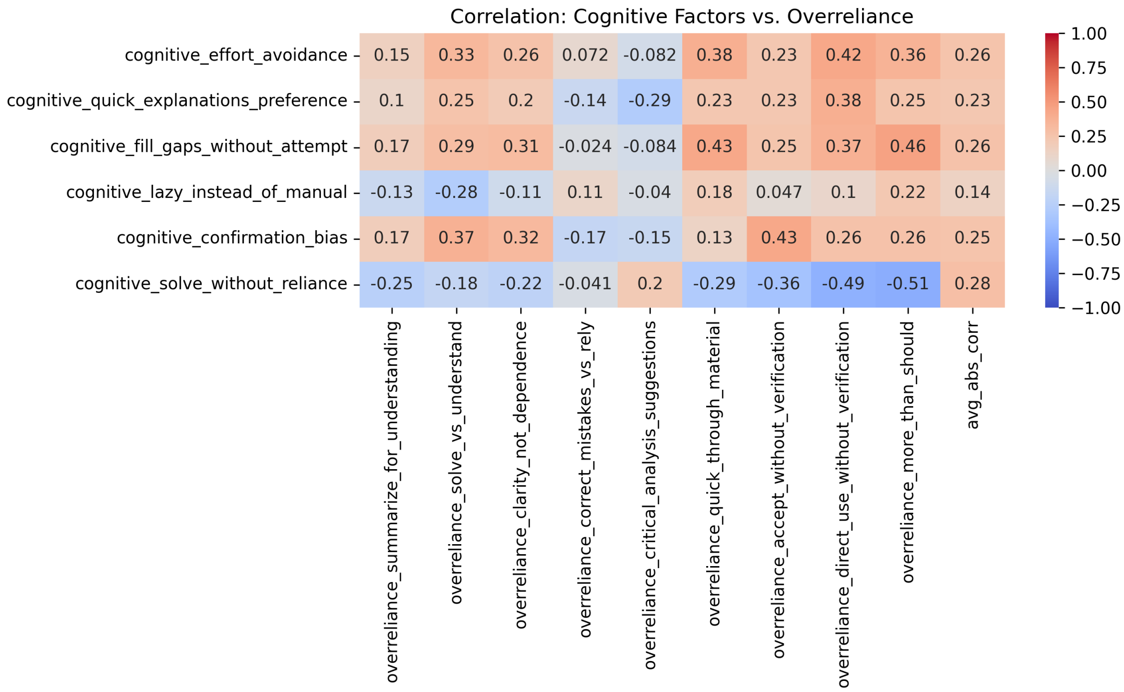

Figure 10: Correlation matrix between questions assessing cognitive factors and questions assessing the degree of overreliance.

directly in their work without verification (`overreliance_direct_use_without_verification`). This variety reflects a combination of factors not captured by the survey, such as course difficulty, interest in the subject matter and existing knowledge, academic confidence and the weight of the course in terms of European Credits (EC).

6.2.2 Cognitive Factors and Overreliance

To better understand the internal motivations behind GenAI usage, we turn to the questions asked in Table 4. These include variables such as cognitive effort, cognitive laziness, self-efficacy and confirmation bias. Figure 10 illustrates the strength and direction of the correlations between these variables and the measured indicators of overreliance in Table 3.

When examining the absolute correlations with the average overreliance score, we observe that the cognitive factors are weak to moderately associated with overreliant behaviors. Notably, some cognitive variables show almost consistent positive correlations across both constructive and compensatory GenAI usage, which seems contradictory.

This could be explained by the confident and unambiguous answers GenAI provides, which may cause students to feel as though they are learning, while in reality, their engagement is minimal. A cognitive bias is created, where convenience is mistaken for comprehension. Therefore, students feel as if they use GenAI constructively but the underlying motivation and depth of engagement might reflect compensatory strategies. This duality raises concerns about the long-term risks of GenAI usage on learning quality, especially when students favor efficiency in their work.

The opposite of this pattern occurs in the assessment of self-efficacy (`cognitive_solve_without`

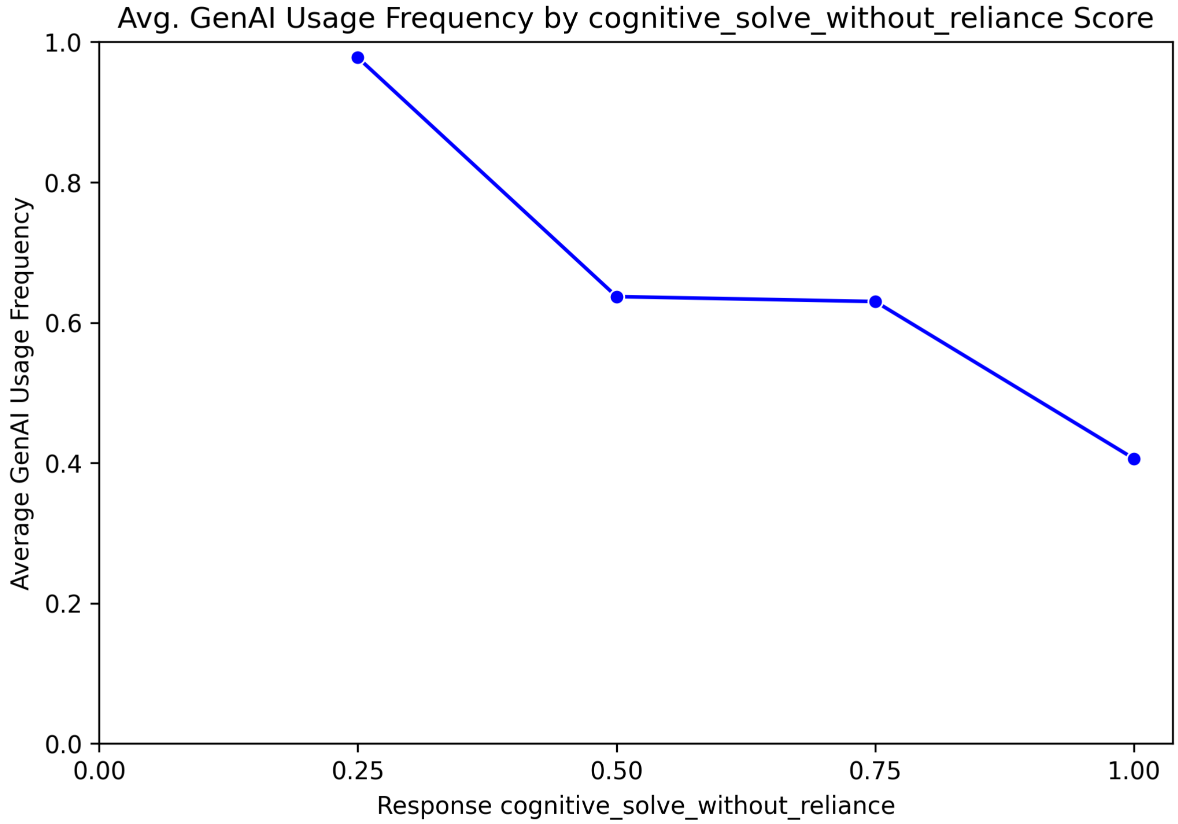

Figure 11: Average GenAI Usage Frequency for `cognitive_solve_without_reliance`

`_reliance`), which is measured here by whether students believe they can solve academic problems on their own. This belief correlates negatively with both constructive and compensatory usage, and suggests that students who are confident in their own abilities are less likely to use GenAI altogether. This is confirmed by Figure 11, which showcases GenAI usage frequency decreasing as users are more confident in their abilities to solve problems on their own.

Despite using GenAI less frequently, they exhibit more critical engagement when considering implementing GenAI suggestions in their work (`overreliance_critical_analysis_suggestions`). This reflects a basic awareness of the tool's limitations.

In line with our expectations, students who report that GenAI helps them save time (`cognitive_effort_avoidance`), prefer quick explanations over detailed study (`cognitive_quick_explanations_preference`), and rely on GenAI to bridge knowledge gaps without attempting to understand the material themselves (`cognitive_fill_gaps_without_attempt`) exhibit predominantly positive correlations with compensatory overreliance indicators. This suggests that these cognitive tendencies may reflect a general inclination toward surface-level processing, and therefore rote-learning, where speed and convenience are prioritized over deeper learning and engagement. In such cases, GenAI becomes a tool for bypassing cognitive effort rather than enhancing comprehension.

One particularly interesting finding emerges from the item assessing confirmation bias (`cognitive_confirmation_bias`). This factor shows consistent positive correlations with nearly all overreliance indicators, regardless of whether the item reflects constructive or compensatory use. This dual association may be explained by the idea that students who seek affirmation from GenAI are not necessarily concerned with the quality or depth of the information they receive, but rather with validating their existing beliefs or assumptions. Consequently, both constructive and compensatory behaviors can be shaped by this need for affirmation. However, the two exceptions `overreliance_correct_mistakes_vs_rely` and `overreliance_critical_analysis_suggestions` show negative correlations. This may be due to the fact that students with strong confirmation bias are less likely to scrutinize or question GenAI output, since doing so could challenge their existing views. As a result, these students may be less inclined to identify errors or critically evaluate suggestions, reinforcing passive and potentially misguided use of GenAI tools.

The item measuring cognitive laziness (`cognitive_lazy_instead_of_manual`) reveals another interesting pattern. Overall, it positively correlates (or nearly at all) with compensatory behavior while negatively correlating with constructive usage. The most negative correlation is observed with the statement about using GenAI to grasp the solution process (`overreliance_solve_vs_understand`), which aligns well with the theoretical understanding of cognitive laziness explained in Section 4.1.1. Students with this trait are less likely to invest the required cognitive effort and are thus assumed to increase the offloading of cognitive tasks to external tools such as GenAI. In line with this, cognitive laziness also correlates positively with admitting of relying on GenAI more than one should (`overreliance_more_than_should`), and thus plays a large role in shaping GenAI usage patterns.

6.2.3 Educational Factors and Overreliance

In Figure 12, we present the correlation matrix between all educational factors (Table 5) and various overreliance indicators (Table 3).

Overall, the average absolute correlations are generally low, suggesting that educational factors do not strongly predict overreliance tendencies. This implies that while these factors may shape students' attitudes towards GenAI, they are not the primary drivers of overreliance.

The variable `overreliance_critical_analysis_suggestions` however, is negatively correlated with nearly all educational factors. This suggests that the more students report being influenced by their educational institute, particularly their curriculum (`educational_curriculum_dependency`) and the guidelines of appropriate GenAI usage (`educational_clear_guidelines`), the less likely they are to critically assess GenAI outputs before integrating them into their academic work.

While one might expect that the presence of clear institutional guidelines on GenAI use would promote more responsible and critical engagement with AI tools, the data reveals a more nuanced picture. The overall correlations between `educational_clear_guidelines` and overreliance indicators are largely neutral, suggesting that the mere existence of such guidelines does not strongly influence students' general patterns of GenAI usage.

However, a closer look reveals weak but noticeable correlations with three specific behaviors: a negative correlation with `overreliance_critical_analysis_suggestions`, and positive correlations with `overreliance_accept_without_verification` and `overreliance_direct_use_without_verification`. This suggests that students who perceive their university as offering clear guidance on GenAI use may, in some cases, interpret this as a form of institutional endorsement or safety net. Rather than promoting critical evaluation, the guidelines may create a false sense of security, implying that if GenAI usage is institutionally regulated, its output can be trusted without further scrutiny. In other words, institutional clarity may unintentionally reduce

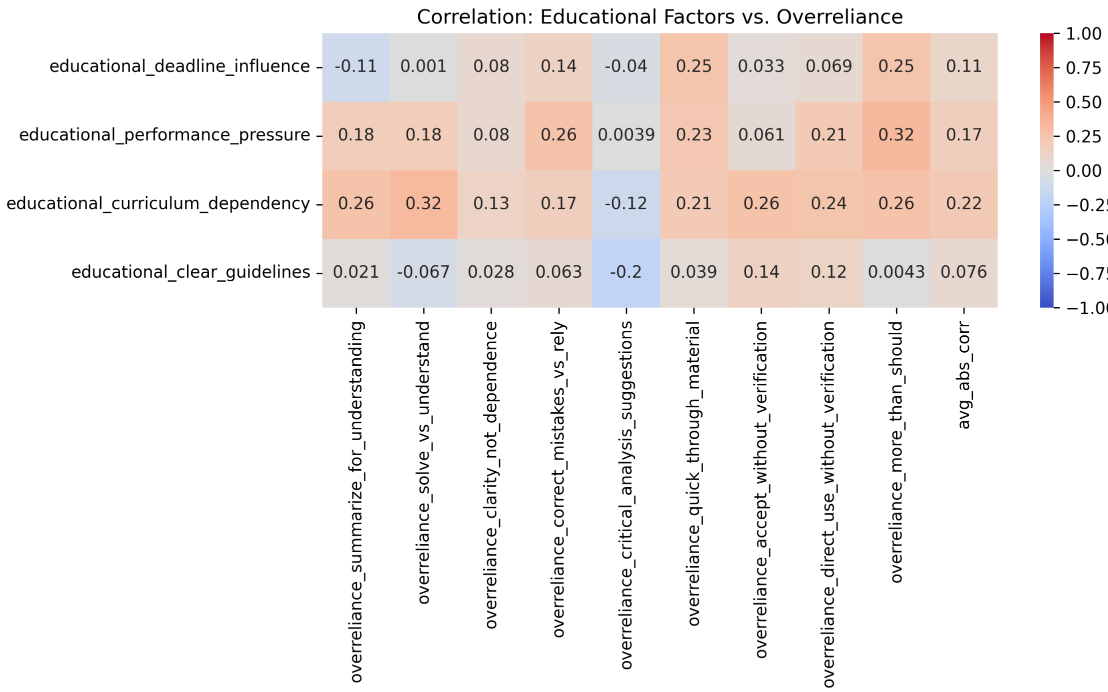

Figure 12: Correlation matrix between questions assessing educational factors and questions assessing the degree of overreliance

students' perceived need for personal responsibility in evaluating GenAI-generated content. Hence guidelines alone are not sufficient if they do not also emphasize the importance of critical thinking and independent verification of AI-generated outputs.

The factor `educational_curriculum_dependency` show consistent positive correlations with most overreliance indicators, regardless of constructive or compensatory use. This trend suggests that students who experience high pressure from their fast paced curricula to perform well are more likely to rely on GenAI overall, both constructively and compensatorily. These students may use GenAI to simplify complex concepts in order to meet performance expectations, even when they feel it might compromise their own understanding. This contradiction between the two usage behaviors explains why students simultaneously acknowledge high reliance (`cognitive_solve_without_reliance`) and use GenAI for academic purposes in both constructive and compensatory ways.

Instead of being strongly correlated with compensatory usage, correlations between the overreliance indicators and the influence deadlines have on using GenAI tools to save time (`educational_deadline_influence`) are weaker than anticipated, except for `overreliance_quick_through_material` and `overreliance_more_than_should`. This suggests that students may turn to GenAI during intensive deadline periods to rapidly process study materials at the expense of their understanding. However, this behavior is not consistent with the other overreliance dimensions. It also does not align well with students' reported GenAI usage during academic situations involving deadlines (Section 6.2.1, Figure 9). Two possible explanations arise. First, students might use GenAI differently depending on the type or timing of deadlines, or the content and EC of the course. For instance, students may use GenAI more constructively when faced with major assignments for courses that interest them, but less with minor weekly tasks for courses they don't enjoy. Second,

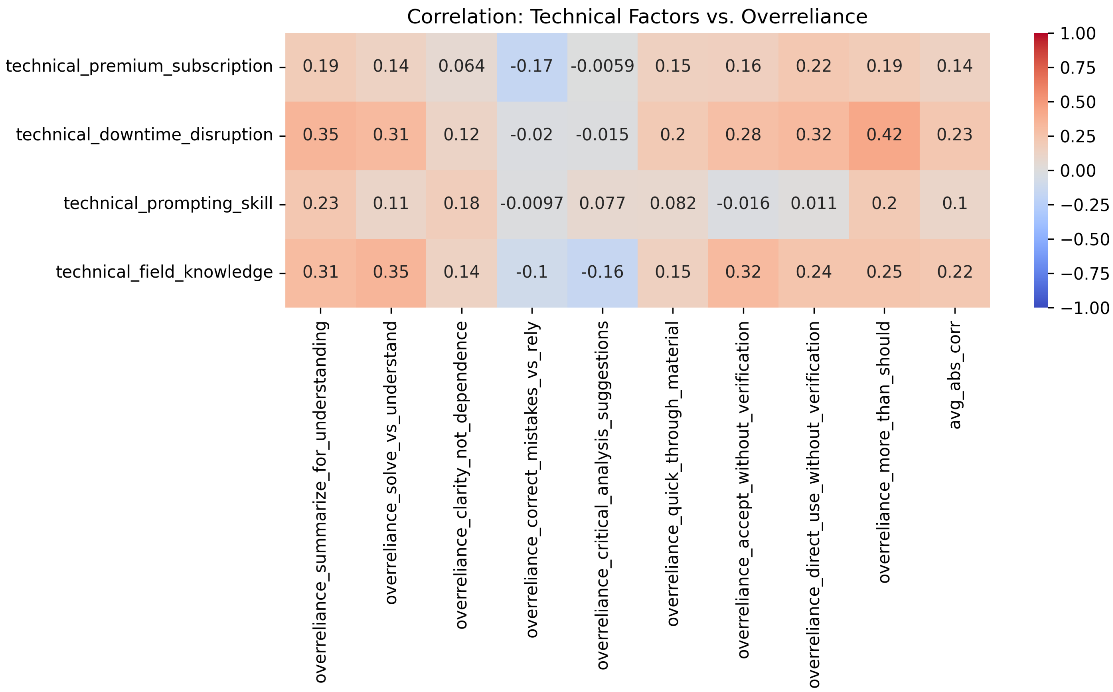

Figure 13: Correlation matrix between questions assessing technical factors and questions assessing the degree of overreliance

the volume of deadlines could also be a factor in driving reliance behavior. It is assumed that students may begin to prioritize completion over comprehension and quality once deadlines pile up.

Overall, educational pressures appear to minimally interact with overreliance. It is expected that the underlying cognitive effects they have, such as psychological stresses, institutional trust and time management challenges, contribute more directly to the degree of overreliance.

6.2.4 Technical Factors and Overreliance

Figure 13 visualizes the correlations between the technical factors listed in Table 6 and the overreliance indicators (Table 3). When averaged across all overreliance items, the factors `technical_downtime_disruption` and `technical_field_knowledge` exhibit the strongest absolute correlations.

Students who have subscribed to a premium GenAI model, believing it delivers more reliable and accurate responses (`technical_premium_subscription`), exhibit a generally moderate positive correlation with compensatory overreliance behaviors. Although there is also a positive association with constructive usage indicators (e.g., `overreliance_summarize_for_understanding`, `overreliance_solve_vs_understand`), the correlations with the compensatory items are slightly stronger, confirming the hypothesis in Section 4.1.3 that increased technical performance from advanced models is associated with a higher degree of overreliance. This pattern suggests two possible explanations:

1. Students may overestimate the accuracy of premium outputs, leading them to disregard critical verification steps and accept responses at face value.

2. Even if premium models make fewer errors overall, students with access may simply encounter fewer mistakes to correct. Consequently, they report lower rates of manual verification (`overreliance_critical_analysis_suggestions`) because fewer corrections are needed.

In both scenarios, premium models can foster a false sense of credibility. Regardless of the model's actual accuracy, students must remain conscious about the limitations and biases in GenAI usage.

Additionally, students report a positive correlation with both constructive and compensatory GenAI usage when GenAI tools experience technical disruptions (`technical_downtime_disruption`). The most notable finding is its particularly high correlation with `overreliance_more_than_should`, which highlights that students who rely heavily on GenAI find their study routines most disrupted. This finding emphasizes the degree to which GenAI has become integrated into daily study habits. In other words, heavy GenAI dependence, regardless of intent, amplifies the negative impact of technical issues.

The factor assessing the knowledge GenAI tools possesses of a student's field of study (`technical_field_knowledge`) has contradictory correlations with overreliance indicators. Positive correlations with constructive indicators suggest that a high accuracy in specific subjects can facilitate deeper engagement. Yet, compensatory usage is dominant as observed by the negative correlation with `overreliance_critical_analysis_suggestions` and `overreliance_correct_mistakes_vs_rely`, and the positive correlations with compensatory statements. In practice, this indicates a trend in which students are less inclined to engage in deeper thinking when utilizing GenAI tools that have high domain knowledge. The perception of high GenAI expertise is associated with increased credibility and satisfaction in the tool's utility, which is in line with our hypothesis in Section 4.1.3. This dynamic highlights the risk of placing confidence in the GenAI's field proficiency, as it can erode student's critical thinking and problem solving abilities.

Interestingly, students who report strong prompting skills (`technical_prompting_skill`) tend to correlate more with constructive usage than compensatory usage. Hence, prompting expertise aligns with a tendency to use GenAI for meaningful learning instead of bypassing engagement and thinking. Nonetheless, these same students still show moderate positive correlations with `overreliance_more_than_should`, indicating some underlying dependency. The overall neutrality of correlations between `technical_prompting_skill` and most overreliance indicators explain why it has the lowest average absolute correlation with overreliance compared to other technical factors. This neutrality also highlights that constructive students might occasionally use GenAI in compensatory ways.

Together, these insights underscore the importance of cultivating GenAI while remaining conscious of its ability to strike a balance between constructive and compensatory GenAI usage.

6.3 Regression Analysis

To better understand the relative influence of the cognitive, educational and technical factors (Tables 4, 5 and 6), a regression analysis was conducted using a machine learning approach. Given the nonlinearity of the data in student behavior (visible in Figure 22 in Appendix B) and the likelihood of interactions between the independent variables, a Random Forest (RF) regression model was selected in order to allow for capturing complex interdependencies, instead of using simpler methods such as linear regression. This model has been imported from the `scikit-learn, RandomForestRegressor` library in Python.

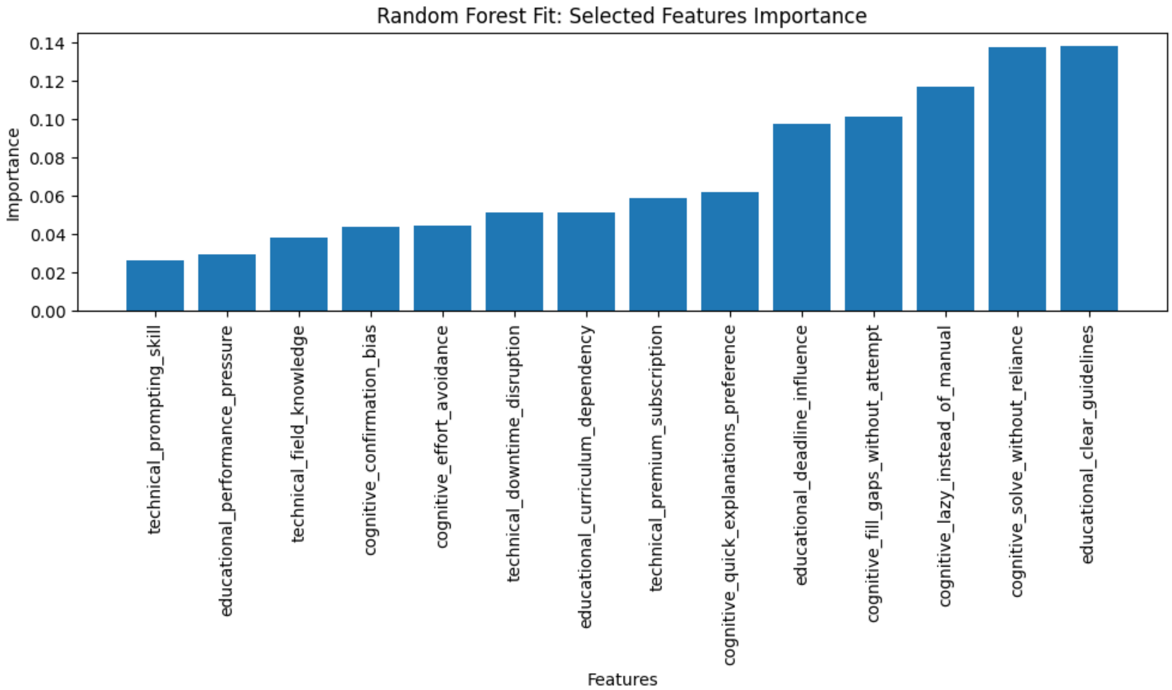

Figure 14: Feature importance from the Random Forest regression model predicting GenAI overreliance

Before fitting the RF model, all columns associated with constructive GenAI usage were inverted so that higher values consistently represented more overreliant compensatory behavior. This behavior reflects stronger dependence on GenAI to compensate for gaps in knowledge, time and motivation. To prevent the model from overfitting, we limited the tree's complexity by setting `min_sample_leaf` to "4", as it seemed appropriate to the sample size we are working with. Additionally, `random_state` was set to "42". Other model parameters weren't adjusted. Similar to the correlation analysis, empty values in each instance were replaced with the mean of the corresponding filled values from other instances.

The trained RF model assigned relative importance scores to each feature based on its contribution to predicting overreliance (Figure 14). The first decision tree that lead to the prediction is illustrated in Figure 23 (Appendix C).

The three most important predictors of overreliance were `educational_clear_guidelines`, `cognitive_lazy_instead_of_manual`, and `cognitive_solve_without_reliance`. Interestingly, technical factors such as `technical_prompting_skill` were assigned negligible importance, suggesting that overreliance is shaped less by students' technical familiarity with GenAI tools and more by their internal capacities and academic context.

To better understand these results, SHAP values were computed using TreeExplainer (from the `shap` and `scikit-learn` library) on the RF model. SHAP values enable an intuitive understanding of each feature's marginal impact on individual predictions based on its input values, accounting for feature interaction and multicollinearity more rigorously than standard feature importance metrics.

The SHAP summary bar chart (Figure 15) confirms the RF's findings. The features with the largest average impact on the model's output this time were `cognitive_solve_without_reliance`,

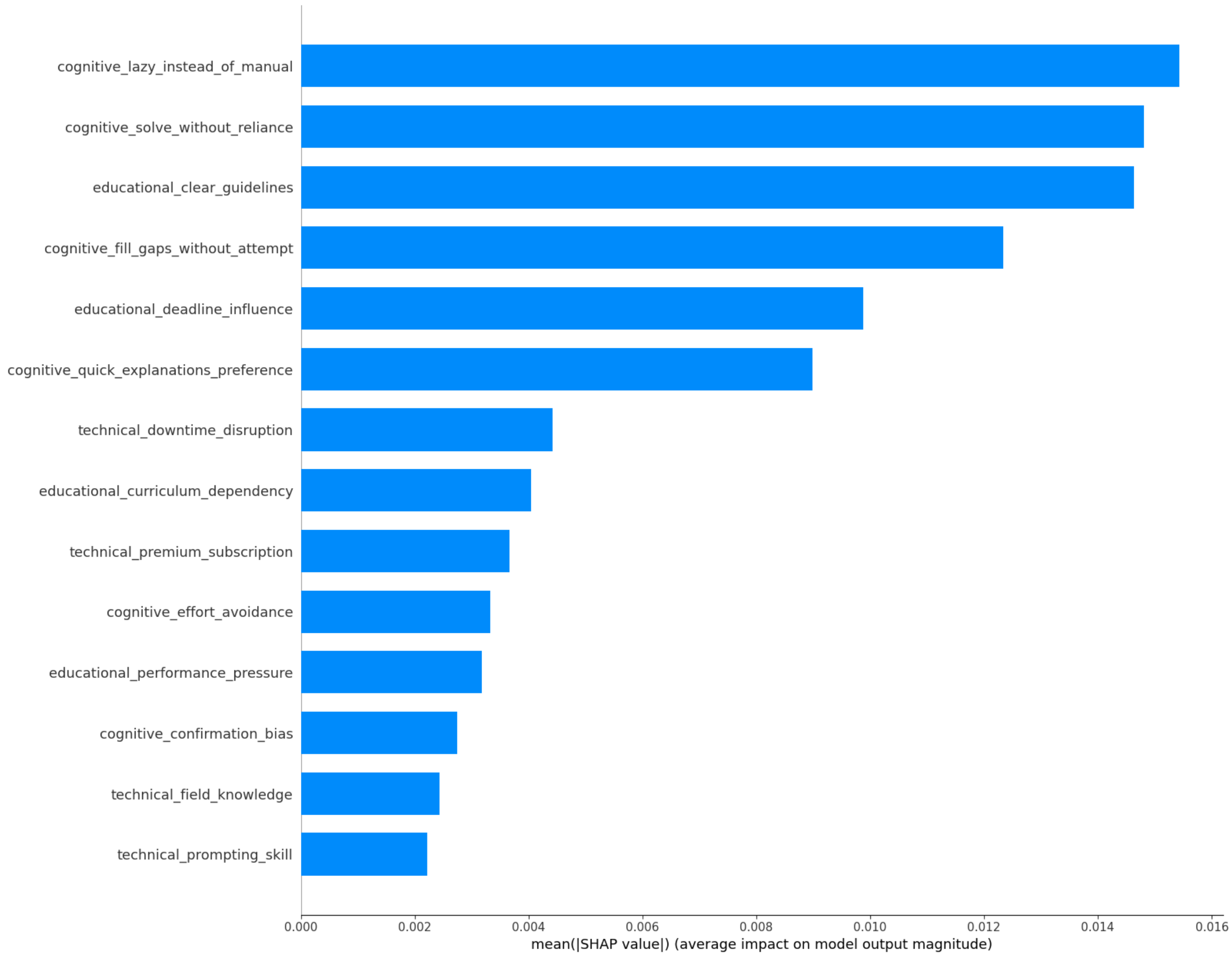

Figure 15: Mean absolute SHAP values showing feature impact on overreliance prediction

cognitive_lazy_instead_of_manual, and educational_clear_guidelines, with technical_prompting_skill having the lowest impact. The values are different as SHAP values offer a local explanation for each instance of how each feature impacts predictions individually.

To better visualize how individual feature values influence the model's prediction of overreliance, a SHAP beeswarm plot was also created (Figure 16), which reveals nuanced patterns in the relationship between feature's input values and their corresponding SHAP values.

According to the SHAP beeswarm analysis (Figure 16), the feature with the greatest impact on overreliance prediction is cognitive_lazy_instead_of_manual, which reflects students' tendency to avoid manual cognitive effort. Higher agreement with this item consistently pushes the model toward greater compensatory overreliance predictions, while lower agreement indicates a higher degree of constructive usage. This finding strongly supports the theoretical expectation introduced in Section 4.1.1, where motivational traits such as cognitive laziness drive reliance on external aids like GenAI. The pronounced effect in the model highlights that cognitive effort avoidance is a core driver of compensatory GenAI usage.

The second most influential factor is cognitive_solve_without_reliance, a measure of

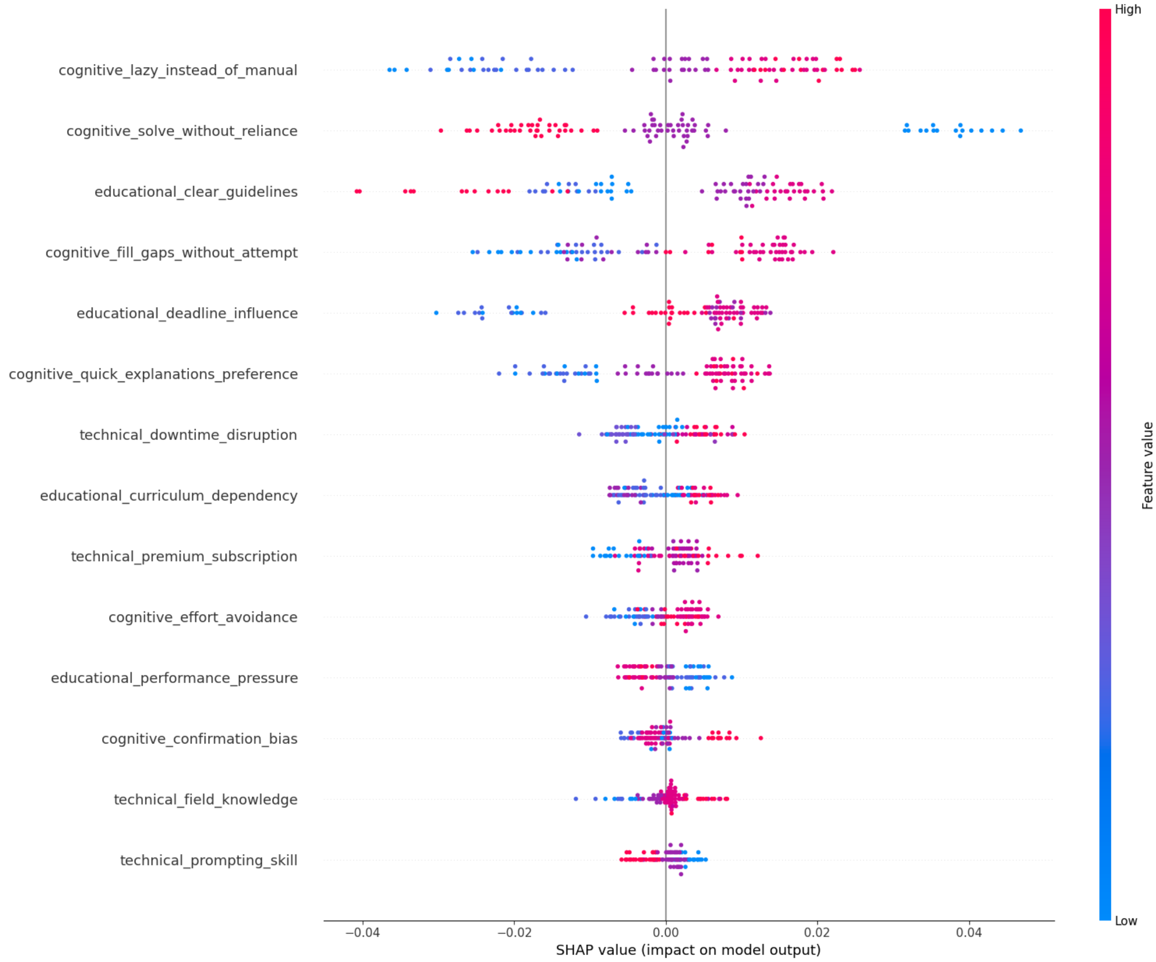

Figure 16: SHAP beeswarm plot illustrating how feature values influence overreliance predictions

self-efficacy in academic problem solving. The SHAP beeswarm plot shows that higher values on this variable predict lower overreliance. This pattern reinforces the idea that students who feel confident in their ability to complete tasks independently are less likely to offload thinking to GenAI. Those with less self-efficacy however, are more likely to use GenAI in a compensatory manner. This finding echoes the importance of internal capacities discussed in Section 4.1.1, especially the protective role of self-efficacy in reducing AI dependency.

Interestingly, the third most important feature was `educational_clear_guidelines`, which reflects students' perception that their institution offers clear rules on GenAI usage. While one might expect such clarity to reduce risky behavior, SHAP values indicate that slightly high agreement with this item increases the likelihood of overreliance. This result supports the earlier interpretation from Section 6.2.3, which suggested that clearly communicated guidelines may inadvertently foster a false sense of security. Students may interpret institutional guidance not as a cautionary framework, but as a green light reducing their perceived need to verify AI outputs or question their use of GenAI in learning contexts. Surprisingly, students that very much agree or disagree with the clarity of these guidelines appear less likely to be overreliant. While these instances can be outliers, it raises an interesting point. Students who are aware of AI-related guidelines may be more deliberate and cautious in their use, regardless of whether they agree or disagree with the statement. In contrast, those with neutral responses likely lack awareness or understanding of the guidelines altogether. As a result, they might use GenAI more freely simply because they haven't been made aware of any boundaries to its use.

Other features, including `cognitive_fill_gaps_without_attempt`, `cognitive_quick_explanations_preference`, and `educational_deadline_influence` showed moderate influence. The SHAP values reveal that students who agreed with these statements were more likely to exhibit compensatory usage patterns. These traits suggest a broader behavioral orientation toward efficiency and limited engagement with material, further reinforcing findings from Section 6.1.

In contrast, features such as `cognitive_confirmation_bias`, `technical_field_knowledge` and `technical_prompting_skill` had the lowest impact in the model, with SHAP values tightly clustered near zero. This suggests that while these variables may intuitively seem important for GenAI interaction, they do not meaningfully differentiate between constructive and compensatory usage in practice. It appears that students rather prefer validating their existing beliefs instead of engaging with the information they receive from GenAI output. One possible explanation for the technical factors is that once a baseline of technical skill or tool familiarity is reached, these competencies no longer significantly influence offloading tendencies. Moreover, it appears that having a technically knowledgeable GenAI or strong prompting abilities may not directly shape the student's cognitive engagement, but rather reflect surface-level interaction.

Based on the distribution of the overreliance score (Figure 17), and the resulting MAE of 0.093 based on a 5-fold cross-validation, we can conclude that the RF model performs well on the range and distribution of the dependent variable (overreliance score). The low MAE indicates that the predictions are close to the actual scores for the students. Therefore, the RF model is solid for behavioral pattern detection, which is very uncommon in psychological scoring, putting further emphasis on the scientific significance of our survey data.

Taken together, these results reinforce that students' internal motivations and behaviors, and environmental factors are the most important in determining overreliance. Particularly, individual characteristics such as cognitive laziness and self-efficacy, and the educational factor on clear guidelines appear to be most influential. Technical proficiency in contrast appears less decisive, suggesting that understanding students' internal reasoning and behavior is more critical to predicting and addressing overreliance.

7 Discussion

This section provides an overview of the thesis' contributions while comparing it with existing literature. It also discusses the significance and limitations of both the conceptual frameworks and the survey, and provides directions for future research.

7.1 Reflection on the Conceptual Frameworks

This thesis introduces two conceptual frameworks: the factor-centered framework (Figure 2) and process-centered framework (Figure 3). When combined, they analyze the effect of cognitive, educational and technical factors on GenAI usage behavior during learning. These factors structure and explain student overreliance on GenAI in academic contexts, and aim to move beyond existing taxonomies of cognitive offloading by integrating environmental, situational and internal capacities with behavioral outcomes. The frameworks do not only serve to structure the mentioned influence in the survey data, but also to interpret and conceptualize how these factors interact dynamically and shape student behavior.

Inspired by the Social Cognitive Theory, the factor-centered framework serves as a visual taxonomy, representing a cascading model where foundational environmental factors create situational contexts, which are in turn interpreted through internal cognitive capacities, leading to a behavioral response. This classification builds and expands on prior work. Furthermore, the model's structural organization allows us to understand how students arrive at different patterns of engagement with GenAI tools. It further emphasizes that overreliance is not the result of a single cause but rather the product of interacting influences filtered through an user's cognitive framework.

Building on the factor-centered framework, the process-centered framework explores how these influences dynamically shape learning behavior. Drawing inspiration from the Dual Process Theory, Social Cognitive Theory and Metacognitive Model of Cognitive Offloading, it discusses that students either use GenAI as a subordinate (prioritize speed and efficiency) or as a teacher (prioritize deeper learning). The framework visualizes how feedback loops such as academic success or tool satisfaction reinforce behavioral preferences towards GenAI usage, and eventually lead to habitual offloading over time. The speed and degree to which this occurs is determined on a personal level, and varies per student. Altogether, the framework suggests that overreliance emerges when students increasingly bypass reflection and default to using GenAI as a shortcut, especially when used as a subordinate.

Both frameworks significantly contribute to the literature by emphasizing the progressive and interactive nature of GenAI usage. Rather than viewing overreliance as a binary state, they position it as the dynamic result of a gradually reinforced behavioral loop. Furthermore, they extend and integrate existing theories and findings in GenAI and cognitive offloading research. While prior work has examined isolated contributors to offloading, few studies attempt to map these influences in a layered or interactive structure. Therefore, by layering these influences and framing cognition as the behavioral response trigger, this thesis introduces a novel hierarchical perspective that clarifies the role of cognition in shaping the progression of habitual offloading that complements, but goes beyond the static typologies in prior studies.

7.2 Reflection on the Survey and Results

Building on the conceptual frameworks, we have designed and distributed a structured survey to 99 Dutch university students who regularly use GenAI tools in their studies. As outlined in Section 5.1, the survey consisted of four parts: demographic and academic background (Table 1, situational frequency of GenAI use (Table 2), overreliance indicators (Table 3), and cognitive, educational and technical drivers (Tables 4, 5 and 6). We also included temporal questions assessing shifts in usage over time (Table 7). Based on the survey responses, we have conducted three different analyses.

Section 6.1 covered the descriptive analysis of the responses. Findings confirmed that students most often turn to GenAI under academic pressures such as deadlines, upcoming exams and overwhelming material. Furthermore, students report a substantial increase in GenAI usage due to tools becoming more efficient and accessible. This posses a danger for students' learning as students could prioritize efficiency driven approaches over critical thinking. These findings provide empirical backing for the first two cascade levels in the frameworks and the triggers, in which external pressures and affordances shape the likelihood of GenAI usage approach.

In the same analysis are the findings on the overall overreliance score, which tie into the interpretive cognitive layer of the factor-centered framework. Despite its low average value of 0.297, students tend to primarily use GenAI tools constructively. However, despite these findings, students frequently admit to relying on GenAI tools more than they should, as well as have a preference for efficiency over understanding and verification of GenAI output. This highlighted complex dependence, despite overall constructive use, substantiates the process-centered framework's claim that habitual GenAI use is the result of diminished cognitive deliberation in repeated academic behavior.

Our correlational analysis in Section 6.2 expanded on the overreliance indicators and revealed modest, but meaningful links, particularly with several cognitive traits. A low self-efficacy had the strongest overall association with overreliance, showing that those that were more confident in their abilities do not use GenAI altogether. This confirms the relevance of internal capacities of the frameworks as a mediator between environment and behavior leading to GenAI usage.

Educational and technical factors on the other hand, showed weaker associations overall, underscoring that external factors alone are insufficient explanations to determining GenAI overreliance. Out of these factors, the content and high pace of certain curricula was associated the strongest with both constructive and compensatory uses, while tool literacy does not seem to immunize against compensatory use. This emphasizes the created GenAI dependency in order to keep up.

For our regression analysis in Section 6.3, we employed a Random Forest model and SHAP values to define the feature's importance in predicting overreliance. This analysis identified three standout predictors of overreliance: perceived clarity of institutional guidelines (`educational_clear_guidelines`), cognitive laziness (`cognitive_lazy_instead_of_manual`), and self-efficacy in problem solving (`cognitive_solve_without_reliance`). Particularly, the predictive importance of `cognitive_lazy_instead_of_manual` supports the tendency to bypass reflective engagement as described in the frameworks, ultimately reinforcing habitual offloading. These emphasize the significance of our frameworks in highlighting the importance of internal capacities, which lead to a behavioral response based on academic, environmental scaffolding rather than technical.

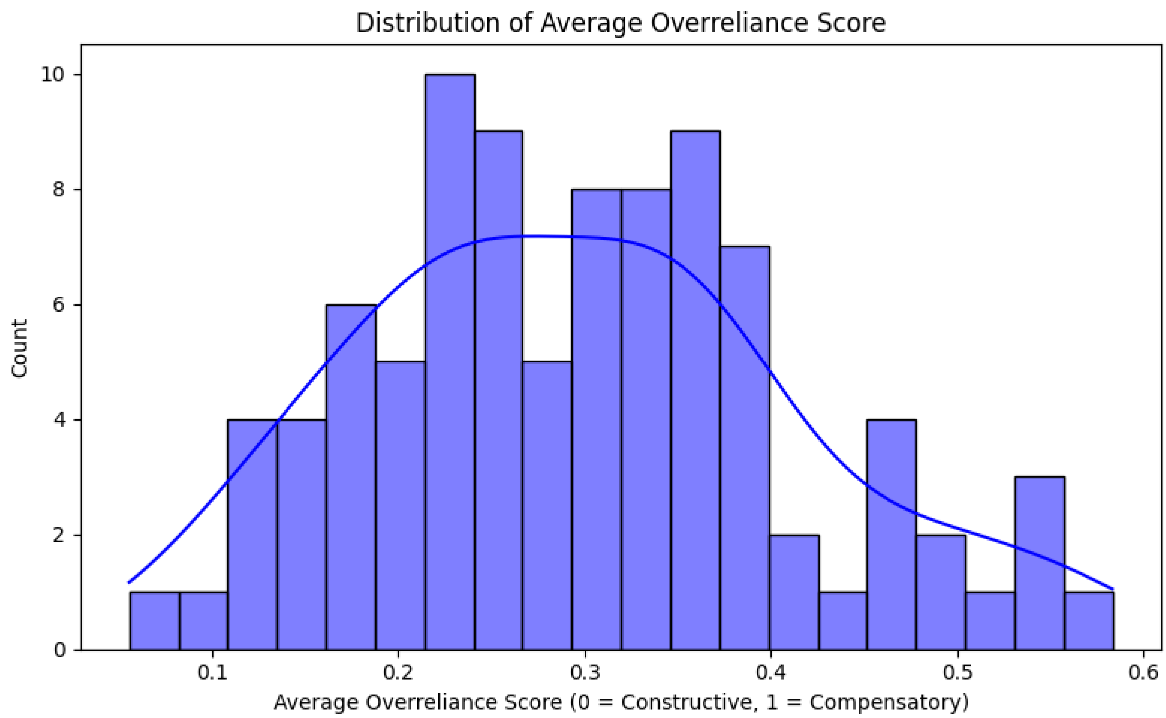

Figure 17: Frequency distribution of average overreliance scores.

The survey results serve as a robust empirical validation of both conceptual frameworks. They confirm that GenAI overreliance is not a simple function of tool availability or academic demand, but the product of a progressive interaction between environmental pressures, internal cognitive traits and repeated behavioral responses.

7.3 Limitations and Future Work

While both the factor-centered and process-centered framework offer a comprehensive and structured foundation for understanding GenAI overreliance, several limitations must be acknowledged to pave the way for future inquiries.

First, the frameworks are conceptual in nature and are not empirically validated. While they are grounded in the survey findings and supported by literature, the models assume a largely linear or unidirectional progression from influencing factors to behaviors. This simplification overlooks the potential for bidirectional or recursive interactions, particularly the feedback mechanisms proposed in the process-centered framework. The actual causal relationships and feedback mechanisms remain theoretical and require longitudinal data to verify. As a result, this thesis is unable to confirm or refute the hypothesis that repeated reliance on GenAI tools shifts cognitive engagement from deliberate to automatic processes (see Section 4.1.1). Similarly, we cannot draw conclusive evidence regarding the hypothesis that overreliance is primarily moderated by students' coping mechanisms, self-regulation abilities, and study habits (see Section 4.1.2). These hypotheses remain open for future empirical investigation.

Second, the frameworks assume a relatively rational and self-aware decision making process on the part of students, especially in the transitions between offloading styles. However, such reflective awareness and metacognitive evaluation are difficult to quantify, and may differ substantially across individuals. While the dual process framing is useful for distinguishing between reflective and reflexive usage, it risks oversimplifying the nuanced strategies students employ, as the idea of a binary "AI as subordinate vs. teacher" role may not capture dynamic use cases where students combine efficiency with learning intent.

Third, the frameworks are focused primarily on the individual, and do not account for broader social or cultural dynamics that may mediate GenAI use. Therefore it underrepresents influential factors such as peer behavior, social norms, student competition, or instructor guidance. The omission of social influence is notable, especially given the role of social learning in technology adoption and academic decision-making. As GenAI becomes embedded in group-based tasks and collaborative learning environments, social influences may become increasingly important.

Finally, the frameworks do not account for the role of emotional dimensions and affective states, such as guilt and shame, as well as other emotional reactions to GenAI usage. For example, students might experience guilt when bypassing learning through GenAI, or conversely feel relieved. These may significantly influence offloading behaviors as either inhibitors or reinforcers.

Despite these limitations, the proposed frameworks offer a structured and flexible foundation for understanding how GenAI overreliance develops. By linking internal cognitive tendencies, situational contexts, and usage behaviors, they help bridge conceptual gaps in existing literature through an interactive perspective. Despite their scope limitations, they provide a valuable starting point for future empirical and theoretical explorations, especially in an evolving educational landscape where AI technologies continue to reshape learning processes. Their significance lies not only in explaining current behaviors but in guiding the design of interventions, policies, and educational strategies aimed at promoting constructive GenAI use.

While our survey yielded rich results, the data remains cross-sectional and self-reported, preventing (direct) causal inference. This suggests the need for more nuanced analyses of GenAI behavior that account for individual's differences and motivational factors, which could be an important direction for future research. In addition, the sample skews toward engineering students at research universities (wo) following a bachelor degree. This demographic imbalance limits the generalizability of the findings and prevents stratified analysis across study field, study level and university type from being conducted. Therefore, we can't answer the hypothesis on whether students in analytical or text-heavy domains (e.g., law, humanities) use GenAI tools differently than students in STEM or arts faculties, and whether it induces GenAI overreliance differently.

Future research should employ longitudinal designs to track how offloading trajectories evolve over time. Expanding the sample to include students from more diverse disciplines and institutions will increase generalizability. Furthermore, future research should explore additional influencing factors such as perceived importance of study credits, academic interests and social norms. Additional data from experiments could also test targeted interventions to disrupt or prevent emerging overreliance patterns.

Despite these constraints, the survey delivers a comprehensive empirical test of the proposed frameworks. Not only does the thesis confirm the interplay of situational pressures, environmental scaffolds, and individual characteristics in determining GenAI overreliance, it also pinpoints which specific factors most strongly predict compensatory usage. These insights are actionable, and equip educators and policymakers with targets promoting mindful GenAI engagement while safeguarding the integrity of meaningful learning.

Altogether, these limitations point to promising directions for future research. Summed up, future research should:

- Explore the proposed feedback loops and behavioral reinforcement mechanisms of overreliance development in more detail through longitudinal or experimental studies.
- Incorporate emotional, cultural and social dimensions to better account for peer effects and behavioral responses leading to GenAI usage.
- Assess the frameworks' applicability across different institutions contexts, given the variability in GenAI integration worldwide.
- Expand the binary mindset model to accommodate mixed or transitional usage strategies.
- Developing instruments to better assess students' awareness and reflectivity in their GenAI approach.
- Differentiate between GenAI tools (e.g., ChatGPT vs. Perplexity) to see if specific tools are more likely to induce overreliance.

8 Conclusion

This thesis explored the growing phenomenon of overreliance on GenAI tools in academic settings. Through the introduction of two novel conceptual frameworks, and analyzing survey data from Dutch university students, it aimed to understand why certain patterns of GenAI usage emerge.

The findings point to a complex relationship between GenAI tools and student behavior. Students often turn to GenAI under situational pressures such as deadlines, overwhelming material and exam preparations. While most students use GenAI in a constructive manner, many admit to relying on it more than they should. However, the decision to rely on GenAI is shaped by a combination of internal cognitive traits, educational structures and the technical aspects around GenAI tools. Our results show that internal factors such as low self-efficacy and cognitive laziness are particularly important. When paired with a strong understanding of institutional guidelines, they often lead to overreliance.

The proposed frameworks in this thesis help structure these relationships. They show that overreliance is the result of a gradual shift from conscious to automatic use, shaped by a continuous loop of convenience, triggers and reduced reflection. This interpretation complements but extends existing work on cognitive offloading by accounting for the influence of both external scaffolds and internal capacities, taking inspiration from several behavioral models and theories.

Naturally, this work comes with its limitations. The frameworks remain partly theoretical, and the survey data is skewed, cross-sectional and limited in scope. Future research should deepen this work with longitudinal designs, larger and more diverse samples, and the inclusion of novel factors covering social, emotional and cultural influences.

Despite its shortcomings, the thesis provides a strong starting point for understanding the drivers of GenAI overreliance as a behavioral outcome shaped by individual characteristics, academic environments and learning habits. This is particularly relevant in an era where GenAI is becoming more integrated into educational systems. The thesis contributes to that understanding by offering both conceptual clarity through the introduction of two frameworks and empirical grounding of them based on survey results. The insights are meant not only to describe present situations, but to inform future policies and tools with the idea of helping students use GenAI as a companion in their learning journey, instead of an escape.

References

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*, *21*(1), 10.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual review of psychology*, *52*(1), 1–26.

Bandura, A., & Wessels, S. (1997). *Self-efficacy*. Cambridge University Press Cambridge.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, O., & Mariman, R. (2024). Generative ai can harm learning. *Available at SSRN*, *4895486*.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, qjae044.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1), 43.

Chen, H.-L., Vicki Widarso, G., & Sutrisno, H. (2020). A chatbot for learning Chinese: Learning achievement and technology acceptance. *Journal of Educational Computing Research*, *58*(6), 1161–1189.

Choi, J. H., & Schwarcz, D. (2023). Ai assistance in legal analysis: An empirical study.

Cleiren, M., van der Kraats, J., van den Herik, J., & van der Putten, P. (2025, February). AI and the education of Gen Z: Transforming tools into companions [Poster session presented at the 17th International Conference on Agents and Artificial Intelligence (ICAART 2025), Porto, Portugal].

Dai, Y., Liu, A., & Lim, C. P. (2023). Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *Procedia CIRP*, *119*, 84–90.

Dell'Acqua, F. (2022). Falling asleep at the wheel: Human/AI collaboration in a field experiment on HR recruiters. *Work. Pap.*

Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (24-013).

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2024). Does ChatGPT enhance student learning? a systematic review and meta-analysis of experimental studies. *Computers & Education*, 105224.

Doshi, A. R., & Hauser, O. P. (2023). Generative artificial intelligence enhances creativity but reduces the diversity of novel content. *arXiv preprint arXiv:2312.00506*.

Duff, A., Boyle, E., Dunleavy, K., & Ferguson, J. (2004). The relationship between personality, approach to learning and academic performance. *36*.

Eckhardt, S., Kühl, N., Dolata, M., & Schwabe, G. (2024). A survey of AI reliance. *arXiv preprint arXiv:2408.03948*.

Ferguson, A. M., McLean, D., & Risko, E. F. (2015). Answers at your fingertips: Access to the internet influences willingness to answer questions. *Consciousness and cognition*, *37*, 91–102.

Fong, R. W.-t., Lee, J. C.-k., Chang, C.-y., Zhang, Z., Ngai, A. C.-y., & Lim, C. P. (2014). Digital teaching portfolio in higher education: Examining colleagues' perceptions to inform implementation strategies. *The Internet and Higher Education*, *20*, 60–68.

Garbarino, E. C., & Edell, J. A. (1997). Cognitive effort, affect, and choice. *Journal of consumer research*, *24*(2), 147–158.

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15*(1), 6.

Gu, S. (2024). Exploring the role of AI in business decision-making and process automation. *International Journal of High School Research*, *6*(3).

Hiel, A. V., & Mervielde, I. (2002). Effects of ambiguity and need for closure on the acquisition of information. *Social Cognition*, *20*(5), 380–408.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European journal of education*, *57*(4), 542–570.

Hu, K. (2023). CHATGPT sets record for fastest-growing user base - analyst note — reuters. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—are they really useful? a systematic review of chatbot-supported language learning. *Journal of computer assisted learning*, *38*(1), 237–257.

Jo, H., & Bang, Y. (2023). Analyzing ChatGPT adoption drivers with the TOEK framework. *Scientific Reports*, *13*(1), 22606.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus; Giroux.

Klarin, J., Hoff, E. V., Larsson, A., & Daukantaitė, D. (2024). Adolescents' use and perceived usefulness of generative AI for schoolwork: Exploring their relationships with executive functioning and academic achievement. *Frontiers in Artificial Intelligence*, *7*, 1415782.

Klingbeil, A., Grützner, C., & Schreck, P. (2024). Trust and reliance on AI—an experimental study on the extent and costs of overreliance on AI. *Computers in Human Behavior*, *160*, 108352.

Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. sustainability, 15 (7), 5614.

Kriger, B. (2024). Mental laziness and erosion of critical thinking. https://medium.com/business-expert-news/mental-laziness-and-erosion-of-critical-thinking-6818c31073e0

Kruglanski, A. W., Orehek, E., Dechesne, M., & Pierro, A. (2010). Lay epistemic theory: The motivational, cognitive, and social aspects of knowledge formation. *Social and Personality Psychology Compass*, *4*(10), 939–950.

Lavin, D. E. (1965). The prediction of academic performance.

Lee, E., & Hannafin, M. J. (2016). A design framework for enhancing engagement in student-centered learning: Own it, learn it, and share it. *Educational technology research and development*, *64*(4), 707–734.

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–22.

Li, B., Zhang, Y., Bubeck, S., Pathuri, J., & Menache, I. (2024). Small language models for application interactions: A case study. *arXiv preprint arXiv:2405.20347*.

Li, Y., Choi, D., Chung, J., Kushman, N., Schrittwieser, J., Leblond, R., Eccles, T., Keeling, J., Gimeno, F., Dal Lago, A., et al. (2022). Competition-level code generation with alphacode. *Science*, *378*(6624), 1092–1097.

Luszczynska, A., & Schwarzer, R. (2015). Social cognitive theory. *Fac Health Sci Publ*, *2015*, 225–251.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into practice*, *41*(4), 226–232.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, *27*(4), 12–12.

Mollick, E. (2023, September). Everyone is above average. https://www.oneusefulthing.org/p/everyone-is-above-average

Mukhtar, M., Firdos, S. S., Zaka, I., & Naeem, S. (2025). Impact of AI dependence on procrastination among university students. *Research Journal of Psychology*, *3*(1), 246–257.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187–192.

Passi, S., & Vorvoreanu, M. (2022). Overreliance on AI literature review. *Microsoft Research*.

Patil, D., Rane, N. L., & Rane, J. (2024a). *The future impact of ChatGPT on several business sectors*. Deep Science Publishing.

Patil, D., Rane, N. L., & Rane, J. (2024b). *The future impact of ChatGPT on several business sectors*. Deep Science Publishing.

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & mass communication educator*, *78*(1), 84–93.

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied sciences*, *13*(9), 5783.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in cognitive sciences*, *20*(9), 676–688.

Rosbach, E., Ammeling, J., Krügel, S., Kießig, A., Fritz, A., Ganz, J., Puget, C., Donovan, T., Klang, A., Köller, M. C., et al. (2024). When two wrongs don't make a right"–examining confirmation bias and the role of time pressure during human-AI collaboration in computational pathology. *arXiv preprint arXiv:2411.01007*.

Schreuter, D., van der Putten, P., & Lamers, M. H. (2021). Trust me on this one: Conforming to conversational assistants. *Minds and Machines*, *31*, 535–562.

Skjuve, M., Brandtzaeg, P. B., & Følstad, A. (2024). Why do people use ChatGPT? exploring user motivations for generative conversational AI. *First Monday*.

Spector, J. M., & Ma, S. (2019). Inquiry and critical thinking skills for the next generation: From artificial intelligence back to human intelligence. *Smart Learning Environments*, *6*(1), 1–11.

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, *160*, 108386.

Swaroop, S., Buçinca, Z., Gajos, K. Z., & Doshi-Velez, F. (2024). Accuracy-time tradeoffs in ai-assisted decision making under time pressure. *Proceedings of the 29th International Conference on Intelligent User Interfaces*, 138–154.

Tay, S. W., Ryan, P., & Ryan, C. A. (2016). Systems 1 and 2 thinking processes and cognitive reflection testing in medical students. *Canadian medical education journal*, *7*(2), e97.

Terwiesch, C. (2023). Would chat GPT3 get a Wharton MBA. *A prediction based on its performance in the operations management course*.

Tuomi, I. (2019). The impact of artificial intelligence on learning, teaching, and education: Policies for the future. JRC science for policy report. *European Commission*.

Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences*, *13*(11), 6716.

Wang, Y. F., Petrina, S., & Feng, F. (2017). VILLAGE—v irtual i mmersive l anguage l earning and g aming e nvironment: Immersion and presence. *British Journal of Educational Technology*, *48*(2), 431–450.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of personality and social psychology*, *67*(6), 1049.

Yin, J., Goh, T.-T., Yang, B., & Xiaobin, Y. (2021). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance. *Journal of Educational Computing Research*, *59*(1), 154–177.

Zhai, X. (2022). ChatGPT user experience: Implications for education. *Available at SSRN 4312418*.

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, *2021*(1), 8812542.

Zhang, S., Shan, C., Lee, J. S. Y., Che, S., & Kim, J. H. (2023). Effect of chatbot-assisted language learning: A meta-analysis. *Education and Information Technologies*, *28*(11), 15223–15243.

A Appendix: Visualized Survey Answer Distribution

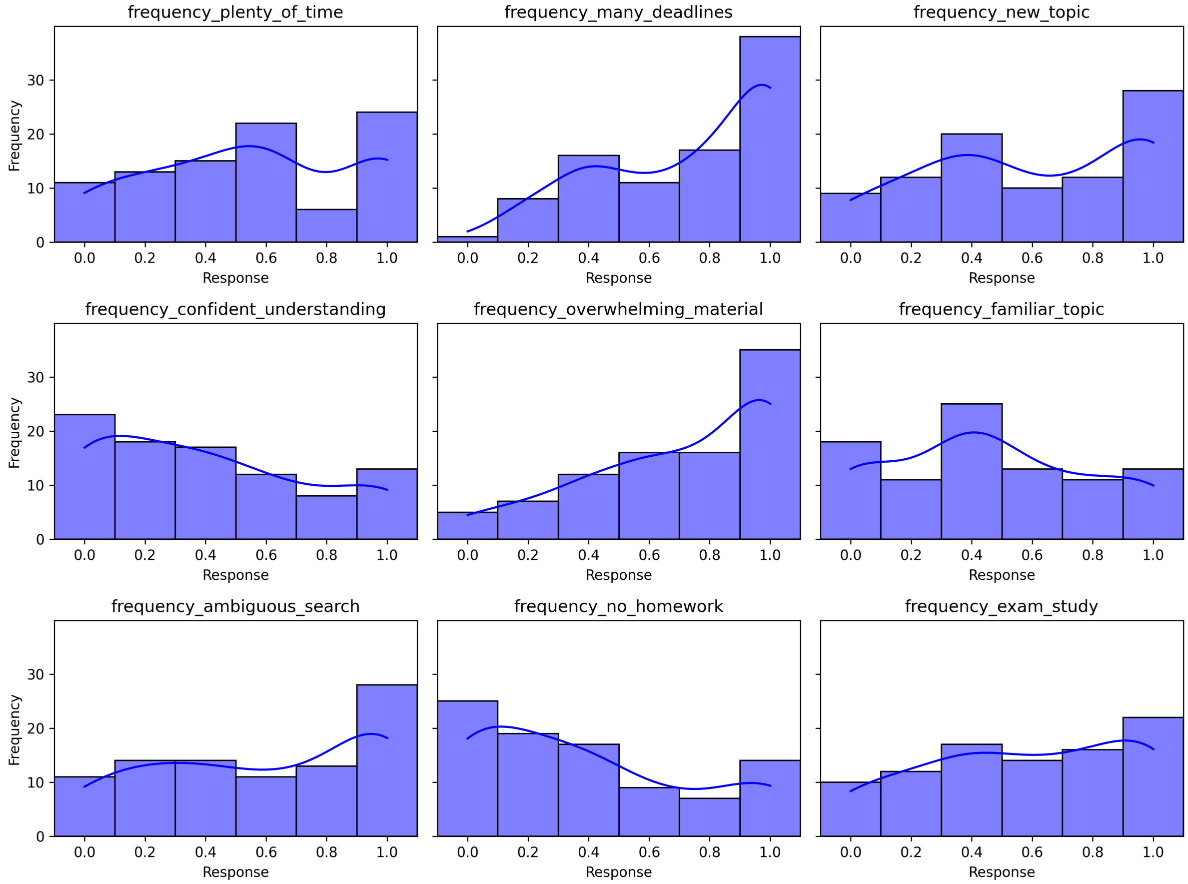

Figure 18: Frequency distribution of responses to frequency-related questions.

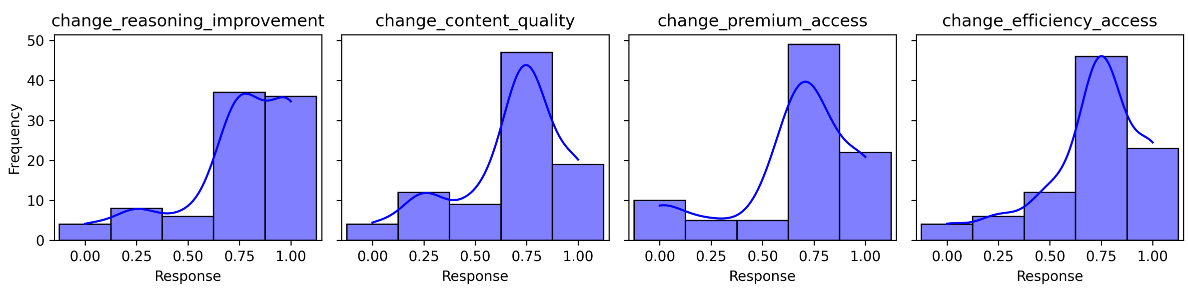

Figure 19: Frequency distribution of responses to questions about change over time.

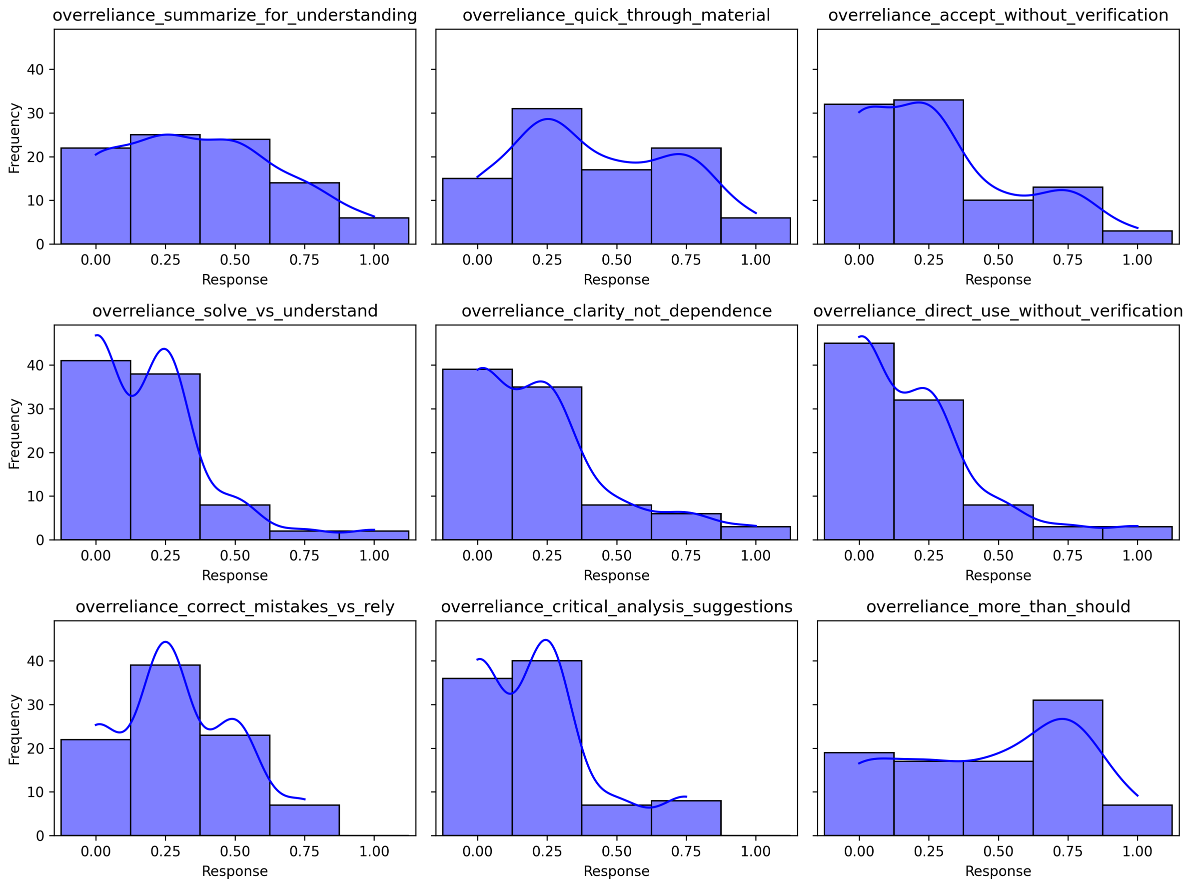

Figure 20: Frequency distribution of responses to overreliance indicator questions.

Figure 21: Frequency distribution of responses to factor-centered questions.

B Appendix: Visualization of Overreliance Score per Factor

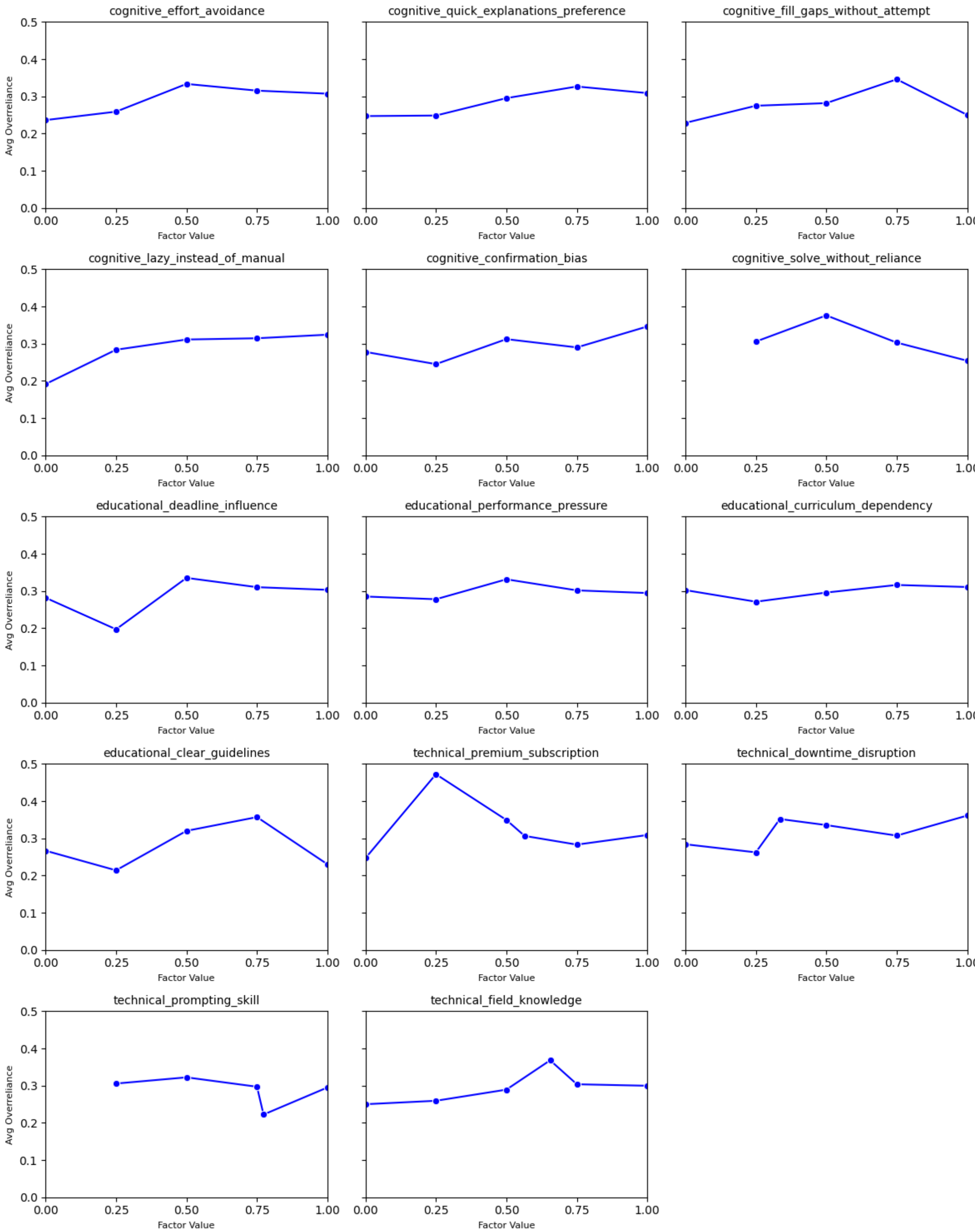

Figure 22: Average overreliance scores across different investigated factors.

C Appendix: Visualized Decision Tree

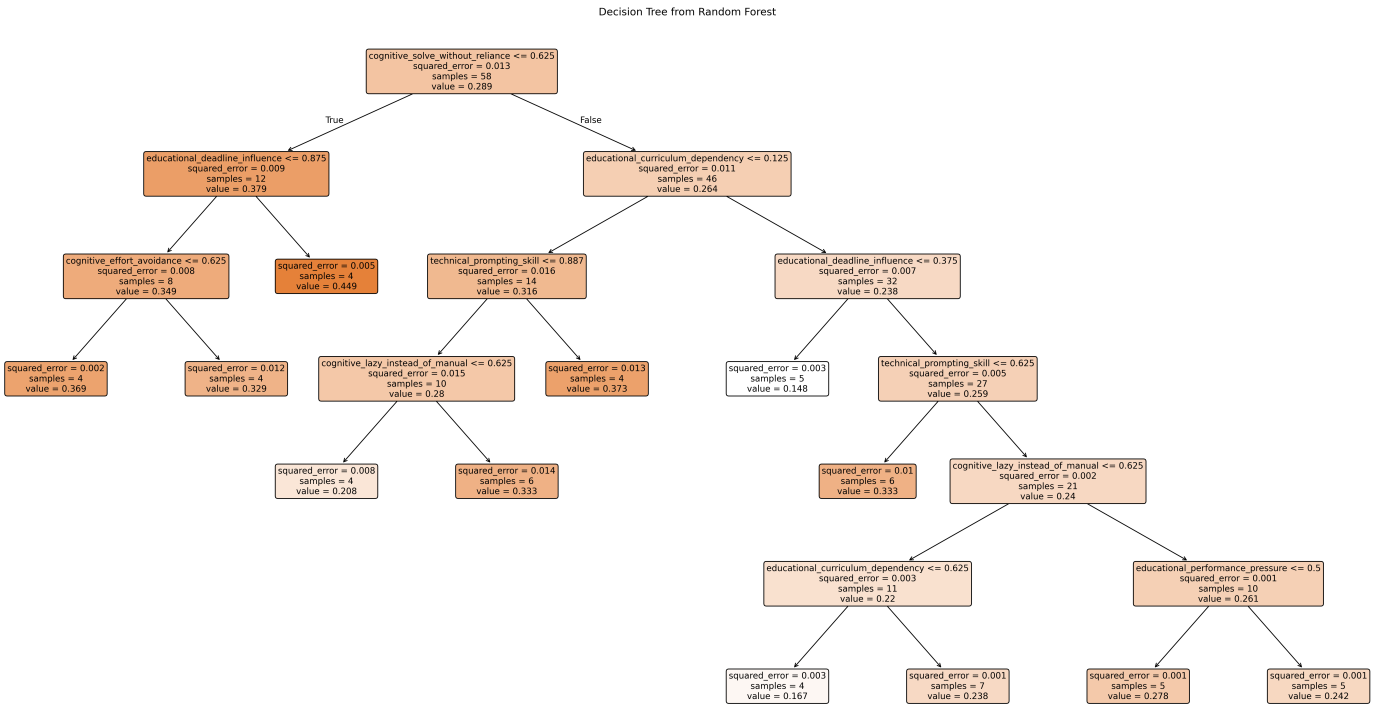

Figure 23: Visualization of the first decision tree extracted from the random forest regression model.