Student Perception on AI-Driven Assessment: Motivation, Engagement and Feedback Capabilities

Yasin Fahmy, University of Twente, The Netherlands

Motivation and engagement are essential for students to thrive in academia. With high student-to-teacher ratios, some students seek more feedback than they receive. To achieve on-demand feedback loops, employing AI-driven assessment is possible. Interviews with students have given insights into the perceived influence of AI assessment on different dimensions of motivation and engagement. Our results revealed that personalized and on-demand AI feedback could fulfill autonomy and competence needs, improving motivation. For engagement, it may reduce student attendance but increase cognitive effort put into studies. Lastly, students had mixed emotional responses to AI-driven assessment and how it would affect their relatedness to peers, professors, and the institute.

1 INTRODUCTION

The work of Braun et al. (2023) has shown that most students are open to the idea of AI-driven assessment in higher education. However, students were skeptical about an AI's ability to interpret written answers, and a balanced approach between human and AI assessment was the most prominent in students' opinions. Recognizing the student's perspective on how AI-driven assessment could impact their motivation and engagement is vital because developers of such assessment tools must consider their needs. AI assessment should optimally enhance student motivation and engagement, leading to better academic outcomes. This research encompasses a literature review and interviews with students. The perceived impact on motivation and engagement was each analyzed using three dimensions. The interview findings have shown that AI assessment could motivate students in two out of three aspects by fulfilling the need for autonomy and competence. Students had mixed views on its influence on the need for relatedness. Additionally, AI assessment could lower attendance, but increase effort exerted. Lastly, students had mixed emotions associated with AI-driven assessment.

2 PROBLEM STATEMENT

While it has been shown that students are generally open to AI-driven assessment tools (Braun et al., 2023), it has yet to be made clear how such tools would affect student motivation and engagement. Hence, this research aims to show the effects of AI-based assessment tools on student motivation and engagement. These AI-driven tools could be beneficial to outsource some of the teacher's workload, freeing up time that can be used for other educational purposes. Addressing this is important because AI assessment tools need to be developed considering students' opinions. We assume that not conducting this research could lead to neglecting students' needs in designing AI-driven assessment tools, which could negatively impact engagement and motivation.

Motivation and engagement are crucial for students' academic growth and development. These are necessary for students to reach their full potential despite their talent. Students who are motivated but do not have initial aptitude are more likely to succeed than students who are talented but not motivated. Goal setting and persistence are a vital trait for students. Furthermore, maintaining student engagement at a high level is crucial, as it determines a student's willingness to collaborate and "ability to function in and contribute to social institutions" (Education GPS, 2024).

Another critical aspect of students' academic development is the feedback they receive. There are several reasons why feedback is essential in higher education. First, it provides students with information on their performance and how to improve. Moreover, feedback can provide "meta-cognitive strategies" (Bick-har et al., 2017, p. 3), which help students set goals, reflect on their progress, and become autonomous learners. Feedback is, therefore, central to a student's motivation and engagement. (Bick-har et al., 2017)

2.1 Research Question

The main question of this research is: "How does AI-driven assessment influence the motivation and engagement of students by grading and providing feedback?".

Three sub-questions were developed that break down the main question into smaller parts:
1. What factors stimulate student motivation and engagement?
2. What are the feedback capabilities of AI-driven assessment tools?
3. What is the perceived impact of AI-driven assessment on student motivation and engagement?

First and foremost, this research aims to uncover the factors that drive student motivation and engagement, which is highly relevant for the education sector. Moreover, this research seeks to identify the feedback capabilities of AI-based tools. Lastly, the perceived impact of AI assessment on student motivation and engagement is investigated. While research on ethical implications of AI-driven assessment is essential, it is out of this project's scope. The insights gained from this exploration can be applied to developing AI-driven assessment tools that motivate and engage students.

3   RESEARCH METHODOLOGY

The research design consists of two methods. First, a literature review is conducted to give an overview of existing research and knowledge gaps. Second, students are interviewed to get new insights into their perspective on this matter.

3.1   Literature Review

This literature review aims to present existing findings from academic sources on motivation, engagement, and feedback and how they relate to AI-driven assessment tools in higher education. Moreover, attention is paid to specific knowledge gaps that can be covered in the interviews. Google Scholar and Elicit, an AI search tool, are used to find academic sources. Specific keywords are used for the search queries, which can be found in Appendix A1. The search results are evaluated for relevancy based on the content, publication year and the number of citations. All literature that is taken into consideration must be freely accessible. The literature review is structured based on each sub-research question. The main findings of the literature for each sub-topic are summarized and synthesized.

3.2   Interviews with Students

The other part of the research is interviewing students. Since this research is heavily focused on students' opinions and perceptions, the interviews can contribute to a better understanding of a student's view on the given topic. By conducting interviews, new insights that might not have been covered by literature before can be gathered. Talking with students gives the opportunity to fully understand their needs when designing AI-driven assessment tools that aim to influence their motivation, engagement, and feedback positively.

The research population consists of students studying at the University of Twente. Random students on campus and peers are asked if they would like to participate in this research. The sample size is 12. The semi-structured interviews give enough flexibility to deviate from the questions while providing a basic structure. The interviews take around 30 minutes and are audio recorded. During the interview, notes are taken. The data gathered from the interviews is anonymous, meaning no personal information is stored. Each interviewee receives an ID, and general information about the participant, such as gender and study program, is stored, which cannot be used to trace back an individual.

4   LITERATURE REVIEW

In this section, the three concepts (engagement, motivation, and feedback) are defined and put in relation to each other. Moreover, theoretical frameworks and strategies are discussed to enhance student motivation and engagement. Additionally, the feedback capabilities of such tools are examined. Lastly, the impact of AI-driven assessment tools on student motivation and engagement is investigated.

4.1   Relationship of Motivation, Engagement and Feedback

Motivation and engagement are closely related topics but mean different things. In an academic context, motivation can be defined as the "inclination, energy, emotion, and drive relevant to learning, working effectively, and achieving" (Martin et al., 2017, p. 1). In contrast, engagement may be described as the "behaviors that reflect this [motivation]" (Martin et al., 2017, p. 1). On the one hand, motivation can act as a driving force for engagement. On the other hand, engagement can be a trigger for motivation as well, making them both "mutually reinforcing" (Martin et al., 2017, p. 10) (Fig. 1).

There are two types of feedback. Formative feedback can be described as an ongoing process during the learning phase of students. It is meant to improve a student's academic performance and understanding of a topic. Examples of this type of feedback are comments, peer reviews, and sample exams. Summative feedback, in contrast, refers to the "assessment of learning conducted at particular points of time" (Bick-har et al., 2017, p. 3), which can come in the form of assignments and exams. We assume that feedback can directly influence motivation and engagement, but not the other way around, because the quality of it depends on the capabilities of the AI-driven assessment tool itself.
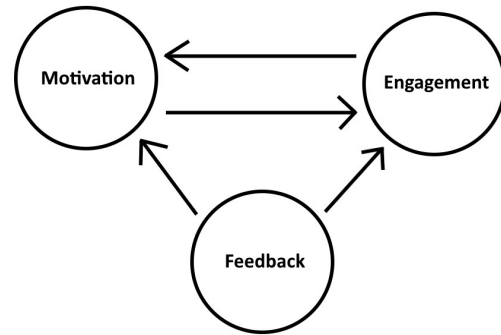

Fig. 1. Assumed relationship between main variables.

Summative feedback can positively and negatively affect student motivation and engagement. Students could become demotivated and disengaged if the feedback is perceived as personal and associated with a sense of failure. In contrast, if the feedback is constructive and proposes improvement, students feel supported, increasing their self-esteem and confidence (Hill et al., 2021). However, research on formative feedback has mainly highlighted its positive effects, which include reduced uncertainty in the learning process and increased student motivation and engagement. Those effects can be achieved by clarifying learning goals, monitoring progress, and promoting self-learning (Bick-har et al., 2017). We assume that bad formative feedback can also have adverse effects, similar to summative feedback.

4.2 Theoretical Frameworks and Factors Stimulating Student Motivation and Engagement

While there are many different theoretical models of motivation, the most common approach could be to differentiate between intrinsic and extrinsic motivation. Intrinsic motivation originates from the joy and satisfaction of a task, whereas extrinsic motivation stems from external rewards or punishments. For example, a student's interest in and enjoyment of a subject counts as intrinsic, whereas grades are an extrinsic source of motivation. The relationship between intrinsic and extrinsic motivation is rather complex since both variables could affect the other in various ways. Study results have shown that the best combination may be a moderate amount of extrinsic and a high amount of intrinsic motivation that leads to the highest average grades and lowest levels of test anxiety among students (Lin et al., 2003).

The self-determination theory (SDT) suggests that a student's motivation can be improved by fulfilling three psychological needs (autonomy, competence, and relatedness), which boosts academic performance. The need for autonomy can be fulfilled by creating an environment that supports students having control over their learning journey to some degree and recognizing the student's perspective. The feeling of competence refers to academic confidence and success and can be invoked by providing purposeful and demanding tasks, recognizing their work, and giving feedback. The third need for relatedness can be achieved through promoting collaborative tasks, establishing a sense of community, and offering student support. (Ryan & Deci, 2020)

Researchers have agreed that engagement consists of multiple facets, but the exact number and types of dimensions that constitute engagement are still under discussion. The widely adopted and comprehensive model of Fredricks et al. (2004) has been chosen to conceptualize student engagement in this research. It considers three different dimensions. Behavioral engagement encompasses various aspects, such as positive conduct and participation in learning, class, and extracurricular activities. Cognitive engagement refers to how much effort is put in by the student to master a specific skill or to obtain domain knowledge. Indicators include the number of questions asked and the level of perseverance and creativity shown when solving problems. Emotional engagement relates to a student's associated emotions (positive and negative) with all aspects of education, for example the level of identification with the institute. (Alrashidi et al., 2016)

Stanford Teaching Commons (n.d.) has composed various strategies to boost student engagement. Firstly, the concerns and worries of students about the course should be addressed to enable class participation. Student engagement can be achieved through asking open-ended questions and providing practice assignments to reduce the fear of failure. Moreover, encouraging students to become more active can be facilitated through group discussions and peer review. Lastly, increasing the diversity of task types may lead to an increase in student engagement. One way to achieve this is by offering different assignments and learning activities.

4.3 Feedback Capabilities of AI-driven Assessment Tools

Becerra-Alonso et al. (2020) discuss the educational tool EduZinc, which aims at improving student engagement in online education. On the one hand, it can grade students; on the other hand, it can provide tailored feedback and adapt to a student's learning journey. Students can submit their results on a website, which is automatically evaluated. Answer types include multiple-choice, code, text, and numeric values. Daily feedback is enabled through the automatic evaluation of personalized exercises and the display of progress reports. The study results have shown that employing the tool positively correlated with student satisfaction and performance.

Currently, only multiple-choice-type questions can reliably be graded by an automatic system due to the complexity of grading criteria for open-ended tasks. A mature AI-driven assessment tool that can evaluate essays based on several grading criteria, such as content, relevance, structure, grammar, and spelling, is yet to be developed. Some assessment tools only consider parts of the assessment criteria of essays, which utilize various machine learning (ML) techniques. Based on the grading criteria, features can be extracted and used by neural networks, regression, and classification models to determine the score of an essay. Semantic features, which capture the meaning of texts and sentences, are more difficult to extract than statistical features. (Ramesh and Sanampudi, 2022)

Perronoski and Bixby (2024) have emphasized the limitations in using generative AI models for grading and giving feedback. Large Language Models (LLMs) were given identical assignments to grade and give feedback based on a grading rubric. The study results have shown that they could not consistently grade or generate feedback. LLM creativity might be beneficial for producing original texts but is counterproductive for delivering consistent grading and feedback. For instance, ChatGPT variations in grading the same assignment ranged from an 8% to 24% difference. This fallacy of LLMs does not contribute to an accurate and fair assessment of students. Such inconsistencies impede students from achieving their learning goals. We therefore assume that irregularities in grading and providing feedback could also hurt student motivation and engagement.

Keuning et al. (2018) have conducted a systematic literature review on how AI-driven assessment tools could provide feedback for programming exercises. Feedback on those exercises can be informative about task constraints, requirements, concepts, mistakes, and how to overcome them. Depending on the type of feedback, specific techniques can be chosen to generate it. The three most common techniques include automated testing, which checks if the given code yields the expected results; program transformations to simplify the code structure without changing its behavior; and static analysis, which hints at fixing mistakes without prior execution. Furthermore, instructors can customize the assessment and feedback of such tools by giving solution templates, specifying correct answers, and providing test cases and common mistakes.

4.4 Impact of AI-driven Assessment Tools on Student Motivation and Engagement

There might be several ways AI-driven assessment tools could enhance student motivation through the fulfillment of needs, according to the SDT. A Delphi study has concluded which generative AI-based activities can fulfill those needs. Out of all the activities that could be part of formative assessment, getting more examples and insights into a problem was linked to a sense of autonomy. The ability to verify one's answers and get feedback on improvement was associated with a feeling of competence. The least amount of activities were associated with the need for relatedness due to a lack of human connection (Chiu, 2024). However, automatic feedback can free up time teachers can use to foster relationships with students and provide personalized assistance if needed (Neji et al., 2023).

Strategies to improve engagement might be similar to motivation because they are highly interlinked. Gamification uses game elements and mechanics in non-game contexts, such as education. It could help keep students engaged by incentivizing learning and participation. For instance, classroom attendance and the completion of exercises could give a student digital points, which can be used to make up a digital leaderboard (Duggal et al., 2021). Another way to boost engagement is by personalizing learning materials and activities. Students could be more interested in learning if activities are tailored to their needs. Exercises could be customized based on a student's aptitude and learning style. Lastly, the individual feedback of AI-driven assessment tools can engage students, showing them their mistakes and how to improve (Nguyen et al., 2024).

Georgouli (2002) developed a prototype for an adaptive student assessment tool that considers a student's current level of achievement, motivation, and overall performance. The level of motivation is determined by a student's effort to solve a problem, confidence in their ability, and how often hints are asked for. While the system has a standard succession of "learning units", the sequence is altered depending on aptitude and motivation. For instance, if the student feels unmotivated or unconfident, the system proposes a problem type that the student is good at to induce a sense of achievement. This system looks very promising regarding motivating students whenever necessary and curating a personalized learning journey.

No literature could be found on the students' perspective on AI-driven grading and how they perceive it would influence their motivation and engagement. This research closes this knowledge gap by conducting interviews with students. Nonetheless, research has been carried out that focused on students' levels of acceptance and trust regarding AI-driven assessment tools. Trust and acceptance could be indicators of the general attitude of students toward such tools. The level of trust students have towards AI-driven assessment tools has been identified as adequate but could be better due to concerns about the ability of such automated systems to assess essay-like answers (Van De Leur, 2022). Lastly, student acceptance is the highest when teachers are still involved in the assessment process (Balickis, 2022).

5 INTERVIEWS WITH STUDENTS

The literature review has answered the first two sub-research questions. However, a knowledge gap on the perceived impact of AI-driven assessment on student motivation and engagement has been identified. Hence, the interview questions focus on answering the third sub-research question.

5.1 Hypotheses

Before formulating the interview questions, hypotheses were formed to determine what insights the qualitative research should bring. The previous literature review has given a broad overview of many dimensions of this research topic. Nine hypotheses were formulated to cover a broad range of aspects to consider.

Hypotheses based on SDT (Motivation):
1. AI assessment fulfills students' need for autonomy
2. AI assessment fulfills students' need for competence
3. AI assessment fulfills students' need for relatedness

Hypotheses based on Fredricks et al. (2004) Model (Engagement):
4. AI assessment affects behavioral engagement
5. AI assessment affects cognitive engagement
6. AI assessment affects emotional engagement

Other hypotheses:
7. The impact of AI evaluation on student motivation and engagement is perceived differently for formative and summative assessment.
8. The impact on student motivation and engagement is perceived differently for specific assignments or question types evaluated by AI.
9. Personalized and on-demand AI feedback increases the intrinsic motivation of students.

5.2 Interview Design

After the participant has signed the informed consent sheet, the purpose and topic of the interview are introduced. Before conducting the interview, the participant is asked what their study background is and what they imagine by AI-driven assessment. After that, we gave our definition so that each participant could imagine the same thing under this term. By AI-driven assessment, we understand a formative or summative assessment tool that utilizes AI to grade and give feedback. It is a broad definition that does not restrict the participants' imagination. They are also encouraged to deviate towards non-AI-related digital components that could be integrated into an AI assessment tool.

The interview questions are based on the previously stated hypotheses, which contain terminology students might need help understanding. Therefore, the questions are formulated in simple language, and necessary concepts are explained. The first six questions are designed to be answered with a Likert scale, which helps discover patterns in the responses. After each answer, the participant is prompted to justify it so that the reasoning behind their answers can be understood. The last three open-ended questions intend to ask for students' opinions. The list of all interview questions can be found in Appendix A2.

6   RESULTS OF INTERVIEWS

This section showcases the results of all the interview questions. First, the students' answers to the motivation questions are described using lollipop charts (Fig. 2-4). Second, the answers on student engagement are summarized and visualized (Table 1, Fig. 5). Lastly, the responses to the other three hypotheses are outlined.

6.1   Student Motivation

Based on the SDT, the questions aim to determine to what extent AI-driven assessment fulfills the three needs (autonomy, competence, relatedness) that contribute to the level of motivation.
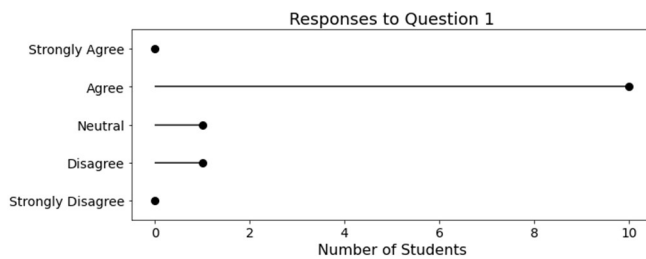

Fig. 2. Results of Q1 about the fulfillment of the need for autonomy (SDT)

The overwhelming majority of students agreed that an AI-driven assessment tool capable of giving feedback would make them feel more autonomous. The students recognized that "teachers are busy" and that "there are [many] students", which leads to teachers or teaching assistants (TAs) needing much more time to formulate feedback than an AI would, which is favorable for students, as they do not need to wait for personalized feedback. Students felt they could work more independently instead of waiting for someone else. A student also believed that AI assessment is more objective than a teacher's evaluation, as "personal attitude", "mood," and relationship with the teacher can influence the grading outcome. Hence, this would make students less dependent on a connection with the teacher. Students have seen formative AI assessment as an opportunity to ask questions and optimize their work independently from asking a teacher or TA. Most students felt that AI-driven assessment would give them more control over their learning and "speed up the learning process". However, students have pinpointed potential weaknesses that AI-generated feedback could have, including missing real-life experience, lack of sources, and low relatability to the student. These aspects may decrease the trust of some students, which would not make them feel more autonomous with AI grading.
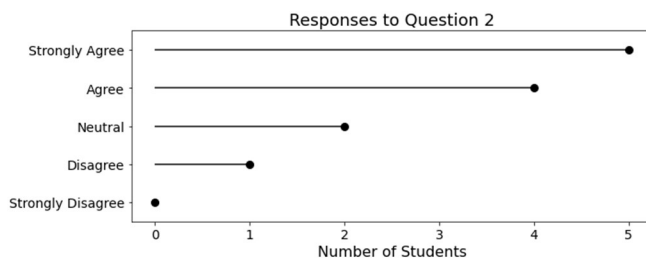

Fig. 3. Results of Q2 about the fulfillment of the need competence (SDT)

For the second question, there was a high level of agreement that AI-driven assessment would make them feel more capable of achieving their academic goals. Students thought they could learn more efficiently and achieve learning goals better due to on-demand and personalized AI feedback. AI was perceived to be good at identifying a student's weaknesses and strengths and showing how to improve academically. Other positive points associated with AI assessment were becoming more consistent in learning due to iterative feedback and an improved understanding of the topic. However, some students thought that it would even have a negative influence on academic performance. Criticism was expressed on AI being prone to hallucinations, which would give students false information. A student expressed concerns that students could become reliant on AI explanations and take its information for granted without checking sources. Another student said that feedback from professors is more suitable for improving academic performance, especially in creative fields.
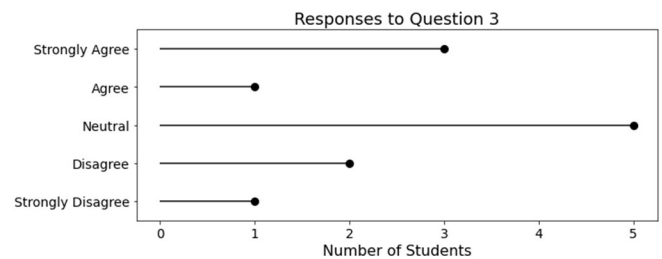

Fig. 4. Results of Q3 about the fulfillment of the need relatedness (SDT)

The answers were more balanced for the third question on whether AI-driven assessment would make students feel less connected to peers, professors, or the institute. On the one hand, some students believed that AI-driven assessment would make them lose touch with peers and teachers. Exchanging feedback was described as a "bonding activity" which would get lost. Furthermore, it was imagined that AI-generated feedback could be very "impersonal, emotionless and blunt", which could make students feel disconnected. Some students said they would only lose touch with their teachers, but not with peers, as they would socialize outside of university. Moreover, the students who answered "neutral" often would still seek feedback from peers or professors, and their relationships would remain unaffected. On the other hand, some students said that AI-driven assessment would make them feel more connected to others. "AI feedback could be a discussion topic" for fellow students, and time saved from formative AI assessment could be spent socializing. Lastly, one student said they would not feel less connected to the institute because "it's cool to work with cutting-edge technology".

6.2   Student Engagement

Compared to the interview statements about motivation, the questions about engagement asked the students to what degree AI assessment would impact their engagement. On the one hand, this left room for the expression of both positive and negative influences; on the other hand, based on the Likert scale numbers, it is impossible to derive a positive or negative trend. Hence, student answers were classified as mentioning only positive or negative, both types or no influences at all (Table 1).

Table 1. Share of Answers for Engagement Influences

Answers in %	Question 4	Question 5	Question 6
Only Positive	8.3%	50%	16.7%
Only Negative	50%	16.7%	25%
Both	0%	25%	25%
None	41.7%	8.3%	33.3%

Unlike Fig. 2 – 4, which aim to show the direction of agreement, Fig. 5 showcases the distribution of perceived impact levels from very low (strongly disagree) to very high (strongly agree). A high variance in the students' judgments on the extent to which AI assessment would influence their engagement has been noticed.
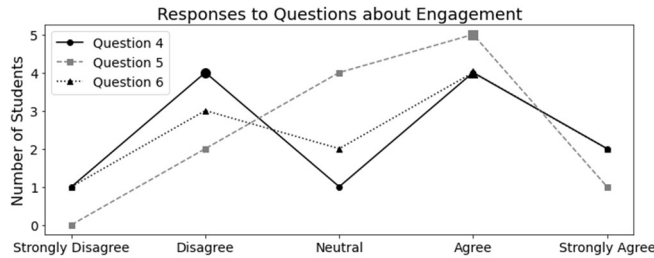

Fig. 5. Perceived degree of impact on different dimensions of engagement

The responses to the fourth question were divided. One-half of the students agreed that AI-driven assessment would affect their behavioral engagement, and the other half disagreed. Some students viewed formative AI-driven assessment as a guidance tool but not as a "replacement of teachers or the people around you" and would still attend all on-campus activities. AI-driven assessment was found to influence attendance and participation in both positive and negative ways. One student said that formative AI assessment would help them catch up on previous lecture materials so that the following lectures can be attended with sufficient base knowledge. The remaining students expressed that AI-driven assessment would make them disengage from school activities, especially those that could be done from home, which include question sessions, tutorials, exam reviews, and study evenings with peers. Of those students, some say they would still attend lectures, while others would not. Lecture attendance was seen as the least affected type of activity. One student perceived AI assessment as a way for teachers to put in less effort and would hence also attend fewer on-campus activities.

Compared to the last question, the students perceived the impact of AI-driven assessment on cognitive engagement more. Most students said it would influence how much effort they put into their studies. Many reasons for a positive influence have been mentioned. Most of them were associated with the feedback capabilities of the tool. Iterative feedback loops were linked to putting more effort because students would know their strengths and weaknesses, striving to improve themselves. The extra time gained from shortening the learning process due to an improved understanding could also be used to dive deeper into a subject. Students also mentioned that merely anticipating good AI-generated feedback would make them put more effort into their work. Other reasons for more effort included checking AI feedback

for correctness and perceiving AI assessment fairer than regular assessment. Some students have mentioned that they would not put extra time gained from more efficient learning into studying when they are extrinsically motivated for that subject. Other reasons for putting less effort included intentionally not trying one's best to get answers from the AI, not caring too much about the "structure, look and feel of a report," and yet again because the reduced effort on the teacher side was seen as unfavorable.

The responses were again mixed for the sixth question on whether AI assessment would change students' feelings about various aspects of education. Some students said their emotions would not change if the AI were unbiased and made for its purpose. One student said, "AI is already part of this world," showing indifference over how work is evaluated. Positive emotions associated with AI assessment were improved satisfaction with the quality and quantity of feedback and trust towards the grading process, as AI was perceived to be more objective and fairer than a human evaluator. On the contrary, some students expressed frustration, fear, and anger if AI would provide irrelevant feedback, teach them wrong things, or misunderstands them. Some students perceived a higher degree of irrelevance from the institute and professors if AI could educate, provide feedback, and assess them. Lastly, learning with an AI-driven assessment in place was seen as less fun because of losing the human touch.

6.3 Other Hypotheses

Formative assessment was seen as engaging and motivating due to the on-demand personalized feedback. However, it could also have a reverse effect if the feedback is considered flawed. Students had more trust in fully automating this type of assessment because the AI would not make any final decisions on grades. For summative assessment, skepticism was higher among students because AI was not considered reliable enough to completely take over the grading process. Strict and unfair AI grading was seen as a driver for demotivation and disengagement.

Generally, technical assignment types that require low interpretability were seen as appropriate, including code, final answer questions, multiple-choice, physics, and mathematics exercises. Small text-based open questions with predetermined answers were also seen as acceptable. Creative, personal, or philosophical assignments were seen as unsuitable for AI assessment. Examples include presentation slides, posters, essays, and reflection reports. The reasons were that the AI had no personal connection to the student and might not fully understand the design choices or ideas a student wants to express.

If AI-generated feedback was detailed, fast, and personalized, students perceived it as a positive influence on intrinsic motivation. Additionally, AI explanations were perceived to help understand a topic better and make learning more fun. Some students said it would not change their intrinsic motivation but merely "speed up" the learning process. In contrast, AI-driven assessment was seen as demotivating when it exceeded personal boundaries related to data privacy, the feedback was too shallow, or the assessment was too strict, unreliable, or unfair.

7   DISCUSSION

This chapter considers the interview results in the context of the literature review. The following discusses the significance of the findings for each hypothesis and gives an interpretation. The interview answers provided many valuable insights.

7.1   Student Motivation

The first three hypotheses aimed to determine if AI-driven assessment could motivate students based on SDT. The interview responses have shown that AI-driven assessment could satisfy the need for autonomy due to the automatic evaluation of one's work. Students would need to wait less time until they receive feedback because of the independence from Teachers or TAs. However, inadequate AI feedback could lessen the positive effect on autonomy because students would still need to seek input from professors and TAs. Ultimately, the extent of increased autonomy depends on the capabilities of the assessment tool. Most students also felt that AI-driven assessment would increase their sense of competence. Such an assessment tool could pinpoint weaknesses and strengths by providing personalized feedback, enabling students to improve. On the contrary, some students believed AI could have the potential to be misleading by providing false information. Learning wrong things would be counterproductive to acquiring academic competencies. Chiu (2024) supports the claim that AI-generated feedback satisfies the need for competence. The interviewed students had mixed views on whether AI-driven assessment would make them feel less connected to others. While some believed they would lose touch with peers and professors due to fewer interactions, others saw AI-generated feedback as an opportunity for discussion rounds. Almost half of students said it would not affect their relationships because they would still socialize outside of class-related activities. Whether and how much AI-driven assessment would fulfill the need for relatedness may come down to the individual students' socialization habits.

7.2   Student Engagement

For behavioral engagement, AI-driven assessment could decrease attendance and participation. Some students said their attendance would stay the same because they saw formative AI assessment more as guidance than a replacement for anything. Others believed their behavioral engagement would decrease because formative AI assessment could replace parts of traditional education. Student attendance is universally seen as desirable, yet tutorials and exam reviews in person might become redundant in the future. Conversely, students were more inclined to believe that AI-driven assessment could increase their cognitive engagement due to iterative feedback loops and identifying strengths and weaknesses. Students could use such a tool to dive deeper into a topic. However, cognitive engagement may decrease if the tool discloses solutions too quickly. Such a tool should be designed to give enough hints when a student is stuck but does not provide shortcuts. Lastly, the responses on emotional engagement were mixed. On the one hand, AI-driven assessment can alleviate stress, worry, and anxiety because it gives you more time to implement feedback before a deadline and know whether you are on the right path. It can also improve students' understanding of a topic and thus make learning more fun. On the other hand, there are many reasons why it can trigger negative emotions, such as, if personal data is used for the

assessment. The effects on emotional engagement may be challenging to assess due to complex factors affecting it.

7.3   Other Hypotheses

Students have considered AI not mature enough to reliably grade students, which matches the study results of Perronoski and Bixby (2024). Hence, due to the significant risks of mis-evaluating students, AI would be more suitable for formative than summative assessment. Employing such AI tools for formative assessment is an excellent opportunity to provide on-demand and personalized student feedback and to conduct pilot projects for AI-driven assessment. The results showed that students trusted AI to evaluate technical more than creative assignment types. This preference was mainly due to the perceived inability of AI to interpret the work as a whole and evaluate new things the AI has never seen before. Ramesh and Sanampudi (2022) support this perspective, stating that the meaning of texts is more difficult to capture than pure statistical features. For some students, good formative AI assessment could increase their intrinsic motivation to study, which boosts academic performance and confidence (Lin et al., 2003). For others, it may not change or even be negatively impacted. All in all, the interview results have shown a diverse range of student opinions on AI-driven assessment. Some interview statements had more consensus among students than others.

8   LIMITATIONS

Firstly, as with any research, there can be unintentional bias, for example, in selecting sources for the literature review, data collection, and interpretation of results. Referenced literature has been selected based on relevancy, but there may be more relevant papers connected to this topic that were not included. Not all participants were selected randomly, which means 25% of interviewed students were already known by the research before. It could also be that students' self-assessments do not accurately reflect their dispositions if AI assessments were implemented in reality. Additionally, since the interviews were based on a hypothetical AI tool, many assumptions had to be made to answer the interview questions, which may have impacted the precision of the results. Besides that, the interview questions were open and some students only answered a specific part of the question. For instance, the concept of autonomy is quite expansive and can take shape in different forms (see interview questions in Appendix A2). Lastly, all participants were students from the University of Twente, for which the study results may not apply to other institutions. The sample size of twelve students may need to be more significant to generalize the study findings.

9   FUTURE RESEARCH

There are many ways to expand the research on AI-driven assessment and its influence on students' motivation and engagement. Future research may build upon the limitations mentioned in the previous chapter. For instance, it could be interesting to conduct experiments to measure student motivation and engagement when they can use a formative AI assessment tool. Then, these study results could be compared to the findings of this research to see if students' hypothetical perceptions would match their behavior in reality. Moreover, due to the exploratory nature

of this research, complementary quantitative research could be beneficial in accepting or refuting the nine formulated hypotheses and strengthening the generalizability of the findings. A survey might be suitable for asking more granular questions, such as differentiating the perceived influences between summative and formative AI-driven assessment for specific dimensions of motivation and engagement. It also could reach more students, making it possible to compare participants' answers for different regions, universities, and educational levels. Lastly, research can be conducted on the ethical implications of AI-driven assessment. With AI starting to be employed in education, further research will be relevant to ensure the best outcomes.

10   CONCLUSION

The main research question was how AI-driven assessment influences the motivation and engagement of students. To answer this question, theoretical considerations of motivation and engagement, the feedback capabilities of AI tools, and the perceived impact on students were examined. The literature review answered the first two sub-questions. First, the three variables, motivation, engagement, and feedback, were defined and put in relation to each other. For all variables, theoretical frameworks were discussed, i.e., summative vs formative feedback. Strategies have been outlined for improving student motivation and engagement. Because the literature review could not answer the student's perceived impact on motivation and engagement, interviews were conducted. Nine hypotheses were formulated, which shaped the interview questions. The results have shown that AI-driven assessment could motivate students in two out of three dimensions. Intrinsic motivation could be influenced positively or negatively, depending on the features of the AI tool. Students were disunited on how much AI-driven assessment would impact their engagement. The analysis has shown that AI-based assessment could lower student attendance while making students put more effort into their studies. Students had mixed emotional responses to AI-driven assessment. Moreover, formative AI assessment was considered more suitable to fully automize due to the less apparent risks of misgrading students. AI was perceived to be more suitable for assessing technical than creative assignments. Ultimately, AI-driven assessment has the potential to revolutionize education with personalized feedback and automated assessment. Considering the student's perspective is crucial to uphold their engagement and motivation.

ACKNOWLEDGMENTS

AI-based grading is a sustained effort by my supervisors and their students. Many papers have been written under the overarching topic of AI-driven assessment in higher education. I am pleased to contribute to this research area with my bachelor's thesis. I want to thank my supervisors, Patricia Rogetzer and Daniel Braun, for their valuable feedback and guidance during the development of this research. Lastly, thank you to all the interviewed students who took the time to share their views on AI-driven assessment.

TOOLS

"During the preparation of this work the author used:
- Word in order to write the proposal.
- Elicit in order to search for relevant literature.
- Google(Scholar) in order to search for relevant literature.
- PDF.ai in order to interact with documents.
- ChatGPT in order to brainstorm, ask questions and generate python scripts for the visualizations.
- Grammarly in order to check spelling and grammar.
- Mendeley in order to manage references.

After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the work."

REFERENCES

[1]   Alrashidi, O., Phan, H. P., & Ngu, B. H. (2016). Academic Engagement: An Overview of Its Definitions, Dimensions, and Major Conceptualisations. *International Education Studies*, *9*(12), 41. https://doi.org/10.5539/ies.v9n12p41

[2]   Balickis, K. (2022). *Influence of human-in-the-loop on the acceptance of AI-driven evaluation of essay questions by students. CC-BY-NC.* https://essay.utwente.nl/91297/

[3]   Becerra-Alonso, D., Lopez-Cobo, I., Gómez-Rey, P., Fernández-Navarro, F., & Barbera, E. (2020). EduZinc: A tool for the creation and assessment of student learning activities in complex open, online and flexible learning environments. *Distance Education*, *41*(1), 86–105. https://doi.org/10.1080/01587919.2020.1724769

[4]   Bick-har, L., Wing-yi Cheng, R., & Yang, M. (2017). Formative feedback as a global facilitator: Impact on intrinsic and extrinsic motivation and positive affect. In *Emerging Practices in Scholarship of Learning and Teaching in a Digital Era* (pp. 265–288). Springer Singapore. https://doi.org/10.1007/978-981-10-3344-5_17

[5]   Braun, D., Rogetzer, P., Stoica, E., & Kurzhals, H. (2023). Students' Perspective on AI-Supported Assessment of Open-Ended Questions in Higher Education. *International Conference on Computer Supported Education, CSEDU - Proceedings*, *2*, 73–79. https://doi.org/10.5220/0011648900003470

[6]   Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: a case of ChatGPT. *Educational Technology Research and Development.* https://doi.org/10.1007/s11423-024-10366-w

[7]   Duggal, K., Gupta, L. R., & Singh, P. (2021). Gamification and Machine Learning Inspired Approach for Classroom Engagement and Learning. *Mathematical Problems in Engineering*, *2021*. https://doi.org/10.1155/2021/9922775

[8]   Education GPS. (2024, April 26). *Student engagement and motivation.* https://gpseducation.oecd.org/revieweducationpolicies/#!node=44131&filter=all

[9]   Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. In *Source: Review of Educational Research* (Vol. 74, Issue 1). http://www.jstor.orgURL:http://www.jstor.org/stable/3516061

[10] Georgouli, K. (2002). The design of a 'motivating' intelligent assessment system. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *2363*, 811–820. https://doi.org/10.1007/3-540-47987-2_81

[11] Hill, J., Berlin, K., Choate, J., Cravens-Brown, L., McKendrick-Calder, L., & Smith, S. (2021). Exploring the emotional responses of undergraduate students to assessment feedback: Implications for instructors. *Teaching and Learning Inquiry*, *9*(1), 294–318. https://doi.org/10.20343/TEACHLEARNINQU.9.1.20

[12] Keuning, H., Jeuring, J., & Heeren, B. (2018). A systematic literature review of automated feedback generation for programming exercises. *ACM Transactions on Computing Education*, *19*(1). https://doi.org/10.1145/3231711

[13] Lin, Y. G., McKeachie, W. J., & Kim, Y. C. (2003). College student intrinsic and/or extrinsic motivation and learning. *Learning and Individual Differences*, *13*(3), 251–258. https://doi.org/10.1016/S1041-6080(02)00092-4

[14] Martin, A. J., Ginns, P., & Papworth, B. (2017). Motivation and engagement: Same or different? Does it matter? *Learning and Individual Differences*, *55*, 150–162. https://doi.org/10.1016/j.lindif.2017.03.013

[15] Neji, W., Boughattas, N., & Ziadi, F. (2023). Exploring New AI-Based Technologies to Enhance Students' Motivation. *Issues in Informing Science and Information Technology*, *20*, 095–110. https://doi.org/10.28945/5149

[16] Nguyen, A., Kremantzis, M., Essien, A., Petrounias, I., & Hosseini, S. (2024). Enhancing Student Engagement Through Artificial Intelligence (AI): Understanding the Basics, Opportunities, and Challenges. In *Journal of University Teaching and Learning Practice* (Issue 6).

[17] Perronoski, F. L., & Bixby, C. (2024). The Theoretical and Practical Implications of OpenAI System Rubric Assessment and Feedback on Higher Education Written Assignments. *American Journal of Educational Research*, *12*(4), 147–158. https://doi.org/10.12691/education-12-4-4

[18] Ramesh, D., & Sanampudi, S. K. (2022). An automated essay scoring systems: a systematic literature review. *Artificial Intelligence Review*, *55*(3), 2495–2527. https://doi.org/10.1007/s10462-021-10068-2

[19] Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*. https://doi.org/10.1016/j.cedpsych.2020.101860

[20] Stanford Teaching Commons. (n.d.). *Increasing Student Engagement.* Retrieved May 17, 2024, from https://teachingcommons.stanford.edu/teaching-guides/foundations-course-design/learning-activities/increasing-student-engagement

[21] Van De Leur, R. (2022). *Challenges and approaches related to AI-driven grading in higher education: the procedural trust of students.* CC-BY-NC. https://essay.utwente.nl/91313/

A APPENDIX

A.1 Search Keywords

Keyword Groups:
1) AI, Machine Learning, Large Language Model, Chatbot
2) Assessment, Summative, Formative, Evaluation, Grading, Essay, Questions, Rubric, Exercises
3) Digital Tool, Software, Program, Application, Platform, Intelligent System
4) Academia, Education, School, Classroom, Student, Opinion, Learning, Acceptance, Trust, Performance, Satisfaction, Emotions,
5) Motivation, Intrinsic, Extrinsic, Strategies, Self-Determination Theory
6) Engagement, Strategies, Gamification, Dimensions
7) Feedback, Summative, Formative, Self-Regulated Learning, Automation, Personalization, Capabilities

A.2 Interview Questions

Questions 1 – 6 are statements for which the participant was prompted to rate them on a Likert scale (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree) and justify his response. Questions 7 – 9 are open-ended, for which no level of agreement was sought after, only the participant's opinion.

A.2.1 Motivation Questions based on Self-determination theory (SDT)
1) I believe that an AI-driven assessment tool that is capable of giving feedback would make me feel more autonomous (i.e. independence, own choices, control over learning)
2) I believe that an AI-driven assessment tool that is capable of giving feedback would make me feel more capable of achieving my academic goals (i.e. confidence, mastery, accomplishments)
3) I believe that AI-driven assessment would make me feel less connected to peers, professors or the institute (i.e. sense of belonging, relationships)

A.2.2 Engagement Questions based on Fredricks et al. (2004) Model
4) AI-driven assessment would change my behavior and participation in school-related activities (i.e. attendance, level of engagement).
5) AI-driven assessment would change how much effort I put into my studies or learning new things (i.e. persistence, study time, discipline).
6) AI-driven assessment would change how I feel about various aspects in education (i.e. teachers, students, class, academic work, grading, institute).

A.2.3 Other Open Questions
7) Do you believe that AI-driven assessment would impact your motivation or engagement differently for practice and graded assignments / assessments?
8) Do you believe that specific assignments or question types evaluated by AI would impact your motivation or engagement differently?
9) Do you believe that personalized and on-demand feedback received by an AI-driven assessment tool would increase your joy of learning?