On the relationship between students' overall satisfaction with higher education STEM courses, and university teachers' didactical practice. An empirical analysis

E. B. Abrahamsen[1*], D. Husebø[1], J. T. Kvaløy[1] and A. Alhourani[1]

[1] University of Stavanger, Norway

*Corresponding author. E-mail: eirik.b.abrahamsen@uis.no

Abstract: In this article, we discuss the relationship between students' overall satisfaction with higher education STEM courses (Science, Technology, Engineering, and Mathematics), and the teachers' didactical practice. We analyzed student feedback from 4,683 digital course evaluations which spanned the period from fall 2020 to fall 2021 at the Faculty of Science and Technology (TN-faculty), University of Stavanger (UiS). Additionally, we collected 1854 open-ended written responses in spring and fall 2022, and conducted two focus group interviews in spring 2023 with eight students from the same population. We raise the research questions: What is the connection between students´ overall satisfaction with STEM courses, found in an institutional student evaluation procedure at UiS, and the university teachers' didactical practice? And what can we learn from students´ evaluations and feedback that is relevant to how we perceive and strive for teaching and educational quality enhancement in higher education? Our analysis shows that there are strong correlations between students´ overall satisfaction, the evaluation questions, and how they understand what being prepared for teaching means.

Keywords: higher education, quality enhancement, student evaluations, AI-assisted analysis, didactical meetings

1 Introduction

The trend in higher education in recent years, both internationally and in Norway, has been to give students a prominent voice on institutional boards and committees, in quality enhancement systems, and in educational research (Hainline et al., 2010; Bishop et al., 2012; Hénard and Roseveare, 2012; Scott, 2015; Speight et al., 2020). This is about engaging the students as integrated members of the academic community and is acknowledged as playing a vital part in both sharing knowledge and strengthening the teaching and educational practices that are taking place (Barber et al., 2013; Fung, 2017; Loukkola and Dakovic, 2017; McCowan, 2017; Davey et al., 2018; Gaebel and Zhang, 2018; Knight-McKenna et al., 2018; Felten, 2020).

In higher education Norway, this trend has in many ways been rooted, enhanced and ensured through the establishment of the National Agency of Quality Assurance and Enhancement (NOKUT) twenty years ago. Since then, NOKUT has ensured that students in Norwegian higher education have had the chance to evaluate their educational programs and courses, through national and institutional surveys. At the national level, NOKUT has taken care of this through digital surveys aimed at both students (Studiebarometeret) and teachers (Underviserundersøkelsen) (NOKUT, 2024). At the local level, the institutions have been required by NOKUT to develop their own student evaluations, close to national understandings and requirements found in the educational law (Kunnskapsdepartementet, 2024). In applicable policy documents and from a quality enhancement perspective, the students are positioned to play a responsible and active part in developing higher education institutions from within (European Commission, 2013; ENQA, 2015; NOKUT, 2016; Meld. St. 16., 2016-2017; European Political Strategy Centre, 2017).

Although there has been a lot of focus on carrying out such evaluations at a national and institutional level, there have been fewer research studies on student evaluation data in Norway in the same period (Lauvås and Maurud, 2008; Aarstad, 2012; Bergfjord, 2014; Strømsø, 2016; Stokke et al., 2019; Worum et al., 2022). The reason for this is most likely related to findings from such studies, which show that there remains considerable reluctance to use student evaluations as grounds for research. This is because there are potential methodological biases related to student evaluation data as valid evidence, and statistical challenges when it comes to representativity, consistent reporting and sample sizes (Aarstad, 2012; Bergfjord, 2014; Strømsø, 2016). Other international studies show that student evaluations and feedback might lead to undesired consequences and decisions. The studies problematize what actually is measured in this kind of evaluations, that there are no grounds for linking student satisfaction with learning, and that, among other things, gender bias is found in research assessing student evaluations (Braga et al., 2014; Boring et al., 2016; Uttl et al., 2017; Feistauer and Richter, 2018; Stroebe, 2020).

But there is also research underscoring that student evaluations and feedback are highly beneficial for educational quality-development in higher education, providing findings and recommendations for this. One such recommendation is to involve students through feedback and collaboration more directly into the learning processes they are engaged in (Kahu, 2011; Dunne and Zandstra, 2011; Bishop, 2018; Fletcher,

2017, 2020; Felten, 2019; Speight et al., 2020). We follow Felten (2019), when he states that student engagement must go on at micro, meso and macro institutional levels. We argue, additionally, that students should also be listened to for the sake of generating new critical research-based knowledge relevant to teaching in higher education. Results from this kind of research is important for a broad audience of stakeholders to understand teaching in higher education better, and could lead to both more informed decisions at a systemic level, and for academics to develop their didactical practices.

With the discussions and potential biases in mind, we argue, once again in line with Felten (2019), that the answer is not to stop student evaluations and the engagement of students in developing the learning processes. The way forward is to do this even more thoroughly in line with established didactical theory, and to link the evaluations more closely to a holistic perspective on the didactical choices made for the teaching that goes on. We further claim in this article that using the data generated through student evaluations could be highly valuable, if interpreted and handled more carefully, scientifically speaking, by the institutions, and if they lay grounds for collaboration and shared commitment by the university teachers and students, developing the didactical practices as a shared endeavor. Finally, we argue that feedback received from the students should be a source of rigorous research, in which students' voices are analyzed into new knowledge for the teaching and research community (Austen, 2020; Bishop et al, 2012; Dunne and Zandstra, 2011; Bovill, 2017; Bishop, 2018; Davey et al., 2018; Felten, 2019; Neves, 2022).

Our aim in this study is to look at the relationship between quality enhancement systems and surveys in our university, and how students' evaluations being part of such systems can potentially influence teaching practices being planned, carried out and evaluated by university teachers. Searching for correlations and patterns here could strengthen the relevance of such evaluations and the findings from them, pointing towards important didactical implications and practice developments.

In this study, we perceive teaching as didactical meetings (Tiller, 1995; Hanssen and Østrem, 2013). The reason for this is that the didactical meetings could be seen as a concept catching the "moment of truth" in realizing quality in higher education. It is in the didactical meetings that the study program and course plans are operationalized and experienced by the students and teachers (Goodlad, 1979; Handal et al., 1990). This meeting must be seen as something that both parties have a shared responsibility and commitment to make as valuable as possible. Didactical meetings between students and teachers are about understanding chosen disciplinary content or carrying out a professional skill better. It represents the didactical triangle (Gundem, 2008; Künzli, 1997) and then the inner triangle of a didactical relational model (see Figure 1). Both students and teachers are learners in the meeting, with the teachers responsible for leading the learning process in which the meeting is taking part. The meetings must be planned, carried out and evaluated in relation to several important frames and structures, such as time, physical and/or digital learning environment, size of the class, and other resources allocated by the institution. Also, the students' prerequisites for learning and the teachers' prerequisites for leading learning are frames influencing the didactical meeting. In the article, we claim that such a didactical perspective should be given more attention in quality systems and discussions today.

We also argue for the use of a broad teaching concept that extends beyond what goes on when students and teachers meet in the classroom, anchored in sociocultural perspectives on learning (Säljö, 2001, 2006; Wittek, 2012). We use the term "didactical practice" in relation to all the things university teachers do as part of their teaching work. This makes it possible to take into consideration the planning, conducting, and evaluation work of the teaching carried out, as well as all the cultural and relational engagement university teachers are involved in today, together with the students and as part of their learning process (Hanssen et al., 2017). The didactical practice is what university teachers have in common as professionals leading research-based learning in higher education (Mausethagen and Mølstad, 2014; Hanssen and Husebø, 2022). Therefore, we can communicate across various scientific disciplines and study programs, sharing knowledge and practices through written and oral discussions. One important issue for this discussion and a broad didactical conception of teaching is that several didactical categories must be seen and considered by the university teacher in a relational and holistic way (Bjørndal and Lieberg, 1978; Hiim and Hippe, 2009). As Figure 1 shows, we need to make sure that what we select of scientific content is related to how we teach and assess and why students are to learn what they are supposed to learn. The students´ prerequisites for learning, and what frames and resources are given (time, rooms, digital tools, etc.), must also be taken into consideration by the university teacher, who is also seen as someone with prerequisites for leading learning and therefore as a didactical frame herself (Hanssen et al., 2012). In such a perspective, university teachers can learn and develop their prerequisites and didactical practices, through various mediating tools, such as reading a research article like this. This may strengthen their pedagogical and didactical language and, thereby, the ability to meta-communicate to students in an informed and qualified way; how these didactical categories have been taken into consideration (Bateson, 1972; Husebø, 2012).

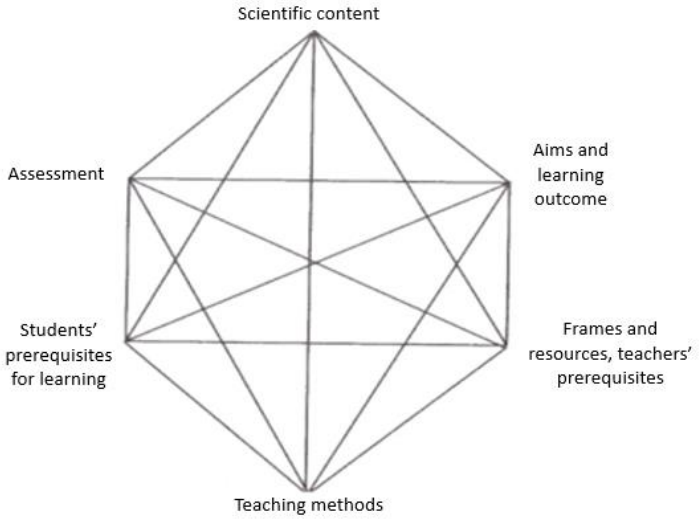

Figure 1. *Teaching as didactical meetings (Tiller, 1995; Østrem and Hanssen, 2013), embedded as the core triangle (Gundem, 2008; Künzli, 1997) in a didactical relational model. Our translation and modification of Bjørndal and Lieberg (1978) and Hiim and Hippe (2009).*

The aim of this article is to bridge the quality enhancement discussions and didactical theory mentioned above, and through the text as a mediating tool, strengthen university teachers' knowledge and skills to consciously plan and lead learning in articulate and explicit ways together with their students. Based on our study, in which we analyze students' evaluations, we try to come to the core of where the quality in education is made real and raise the following two research questions:

> What is the connection between students' reports of overall satisfaction in STEM courses (Science, Technology, Engineering, and Mathematics) and the university teachers' didactical practices?
>
> What can we learn from students' evaluations and feedback that is relevant to how we perceive and strive for teaching and educational quality enhancement in higher education?

2 Data collection and analysis

The data that forms the basis of our study comes in the following two stages from the Faculty of Science and Technology at the University of Stavanger (UiS): The two first datasets stem from a standardized institutional questionnaire established as part of the quality system of UiS. The third part of the dataset is based on this study's own design, inviting 20 randomly chosen students in their last semester of a STEM-oriented bachelor at UiS to take part in a focus group interview and this study.

1. Quantitative data
   Established digital course evaluations at UiS as part of the educational quality system from three semesters, spanning from autumn 2020 to autumn 2021.
2. Qualitative data
   a. 1854 written texts in which the teknisk-naturvitskaplege (TN)-students commented freely on the numeric scoring they gave in the digital course evaluations in spring and fall 2022.
   b. Two focus-group meetings, held with eight students invited as third year STEM bachelor students in spring 2023 (three students attended the first meeting, and five students the second meeting). These students were part of the same bachelor population answering the student evaluation fall 2000
   c.

2.1 Quantitative data (digital student course evaluations)

UiS sends out an institutional survey to evaluate student satisfaction as part of the educational quality system required by NOKUT. The survey is distributed to all students, independent of faculty, after they have completed the examinations in a course (10 – 15 ECTS) embedded in a study program. Consequently, during a semester, a student may receive multiple surveys, with the number of surveys equaling the number of completed courses. The questions in the survey have been very similar in recent years, with just

minor adjustments. In the period from 2020-2021, the following eight questions were asked in the institutional course evaluation:

1. To what extent do you think that the course's teaching and learning methods contribute to your learning?
2. To what extent do you think the teaching in this course conveys the subject content in an understandable way?
3. To what extent are you satisfied with the subject teacher's use of digital tools in teaching?
4. How many hours do you work on the course per week (including preparation, lectures, practice hours, seminars, and lab)?
5. To what extent do you think that the subject teacher(s) are well prepared for teaching?
6. To what extent do you receive feedback and guidance from the subject teacher(s)?
7. To what extent are you satisfied with the learning environment of the course?
8. How satisfied are you overall with the course?

For all questions, except question 4, the students replied on a five-point Likert scale with alternatives: 1="not at all", 2="to a low degree", 3="neither nor", 4="to a large degree", 5="to a very large degree". There is also a "do not know" answer option. For question 4, the scale is 1="0-5 hours", 2="6-10 hours", 3="11-15 hours", 4="16-20 hours", 5="21-25 hours" and 6="more than 25 hours."

To examine the relationships between the students' responses to each of the questions, we set up cross-tables and calculated Spearman correlations. Spearman correlation is a nonparametric correlation measure, which can also be used on ordinal data, as we have here. The strength and direction of the association between the variables is, as for other correlation measures, quantified on a scale from [-1,1], where 0 means no monotonic relationship, and larger positive/negative numbers mean a stronger positive/negative monotonic association. To test whether the relationships are statistically significant, we use chi-squared tests or tests based on the Spearman correlation. Statistical tests with a p-value<0.05 are considered significant. The analyses were performed in Excel and R version 4.3.3.

The Faculty of Science and Technology at UiS received a total of 1933, 1055, and 1693 course evaluations for autumn 2020, spring 2021, and autumn 2021, respectively. The corresponding response rates were 17%, 14% and 17%.

2.2 Qualitative data (free text answers and focus group interviews)

Due not only to low response rates, but also to get a deeper understanding of the students' impressions, we supplement the quantitative data with qualitative insights, to obtain more and different kinds of data in this study. The qualitative data stemmed from 1845 free-text contributions in the digital course evaluation in spring and fall 2022, and two focus-group interviews organized in spring 2023. The written responses are found in two open free-text opportunities to comment upon their scoring of the eight questions in UiS standardized digital student evaluation to allow the students to provide further comments on their overall satisfaction.

Thematic analysis of the free-text student feedback was processed by artificial intelligence. Utilizing large language models in a multi-stage process was implemented to ensure participant privacy, anonymity, costs and quality of data. First, a secure, locally hosted Llama 3-based language model (Dubey et al., 2024) was used to rewrite each piece of feedback concisely in English. The model was prompted to remove all identifying information such as names, specific course titles, dates, and locations while maintaining the original sentiment and main points. A random sample of 10% of the rewritten feedback was then reviewed to ensure successful removal of identifying information and preservation of original meaning.

The anonymized data was then analyzed using GPT-4T (specifically gpt-4-turbo-2024-04-09), an advanced, capable language model developed by OpenAI. A set of 13 themes was developed based on an initial review of the feedback. This was done in incremental prompts with token overlap to ensure that all the data fit into the model's context window of 128,000 tokens**.** Unlike traditional keyword-based methods, this approach uses the model's natural language understanding capabilities to identify and categorize themes within the context of the entire feedback, allowing for a more nuanced understanding of the themes present. The produced themes were then qualitatively validated. GPT-4T was then used to categorize each piece of feedback into one of these themes, and classify the overall sentiment as positive, negative, or neutral/mixed. The prompt for this analysis was structured to ensure consistency across all feedback entries using a chain-of-thought method containing examples of appropriate and validated classifications, and the outputs were closely monitored. The final dataset included an anonymized response identifier, assigned theme, concise summary, and sentiment classification for each entry. Further analysis and visualizations were carried out using PowerBi. And theme-specific summaries including key findings were generated using GPT-4T by analyzing pooled feedback from one theme at a time.

The objective of the two focus-group interviews was to gather additional information on some students' reasoning behind their scoring in previous course evaluations, as well as their interpretations of the quantitative correlations that we identified in the first part of our study.

The interviews were carried out for one hour, with three students participating in the first focus-group interview and five students in the second. In total, 20 third year STEM-students were randomly invited to take part. The focus-group interviews were audio-recorded and later transcribed and analyzed. As a basis for the focus-group interviews, the students were first asked to answer the eight digital survey questions orally and score their satisfaction for the whole bachelor program. Then, the students were asked to reflect upon their own overall satisfaction with the study program and share their main arguments for this. They were also asked to share their interpretations of the main findings from the quantitative study presented above.

After the interviews were transcribed, each of the authors of this article wrote down key words based on the students' feedback to each of the eight questions in the survey.

A summary from each of the authors was then compiled, and consensus was reached upon the meaning of the interviews1.

3 Results

3.1 Quantitative insights from 2020-2021

To address our research question about the connection between students' overall satisfaction with a higher education STEM course and university teachers' didactical practice, we will examine the initial quantitative study, conducted based on the 4,683 course evaluations received during 2020-2021, covering 315 courses. From Table 1 and Figure 2, we see the relationship between students' overall satisfaction with a course, Q8, and the dimensions covered by all the other questions in the course evaluation.

Based on the results from Table 1 and Figure 2, we see a clear and significant positive relationship between students' overall satisfaction with the course and the dimensions represented by all the other questions from the student evaluation, except for Q4 (time students spend on the subject).

Table 1. *Relationship between the students' overall satisfaction with a course (Q8) and their response to each of the questions, Q1-Q7. For each level of reported course satisfaction (Q8), the table shows the proportions (absolute numbers in parentheses) of students in each of the answer categories on the other question (Q1-Q7). The p-values reported are for the relationship between Q8 and each of the other questions, Q1-Q7.*

		Q1. To what extent do you think that the teaching and learning methods of the course contribute to your learning?				
	P<0.001	1	2	3	4	5
Q8. How satisfied are you overall with the course?	1	50.4% (194)	40.3% (155)	6.8% (26)	1.0% (4)	1.6% (6)
	2	10.8% (71)	52.1% (342)	28.2% (185)	7.8% (51)	1.1% (7)
	3	2.2% (16)	18.1% (133)	48.5% (356)	27.9% (205)	3.3% (24)
	4	0.2% (3)	2.7% (44)	16.5% (267)	63.5%(1030)	17.1% (278)
	5	0.1% (1)	0.3% (4)	1.4% (16)	27.1% (313)	71.1% (820)

		Q2. To what extent do you think the teaching in this course conveys the teaching material in an understandable way?				
	P<0.001	1	2	3	4	5
Q8. How satisfied are you overall with the course?	1	52.1% (199)	38.7% (148)	6.5% (25)	2.6% (10)	0.0% (0)
	2	11.4% (75)	58.5% (384)	23.0% (151)	6.4% (42)	0.6% (4)
	3	2.0% (15)	24.1% (177)	49.0% (359)	21.8% (160)	3.0% (22)
	4	0.4% (7)	4.6% (74)	18.4% (293)	60.7% (968)	15.9% (254)
	5	0.3% (3)	0.3% (4)	1.9% (22)	28.2% (323)	69.3% (793)

[1] We have submitted and received approval for the processing of personal data from the Norwegian Agency for Shared Services in Education and Research (SIKT), with reference number 712088. All students engaged in the study are anonymized. When we refer to names among the eight students taking part in the focus group interviews, we use nicknames in order to create more lively language.

		Q3. To what extent are you satisfied with the subject teacher's use of digital tools in teaching?				
	P<0.001	1	2	3	4	5
Q8. How satisfied are you overall with the course?	1	51.7.% (193)	24.9% (93)	12.9% (48)	7.5% (28)	2.9% (11)
	2	14.5% (94)	36.3% (236)	31.7% (206)	13.7% (89)	3.8% (25)
	3	4.3% (31)	15.8% (115)	43.6% (317)	27.9% (203)	8.4% (61)
	4	1.2% (19)	5.4% (86)	20.1% (322)	48.4% (776)	25.0% (401)
	5	0.3% (3)	1.6% (18)	5.6% (64)	24.8% (281)	67.8% 769)

		Q4. How many hours do you work on the course per week (including preparation, lectures, practice hours, seminars and lab)?				
	P<0.001	1	2	3	4	5
Q8. How satisfied are you overall with the course?	1	22.2% (86)	26.1% (101)	19.4% (75)	16.0% (62)	5.9% (23)
	2	15.0% (99)	32.0% (211)	24.3% (160)	15.9% (105)	6.1%(40)
	3	20.5% (152)	32.5% (241)	26.1% (194)	12.0% (89)	4.3% (32)
	4	12.9% (210)	32.5% (530)	29.1% (475)	15.1% (247)	6.3% (103)
	5	10.0% (115)	31.2% (360)	32.5% (375)	15.2% (175)	6.5% (75)

		Q5. To what extent do you think that the subject teacher(s) are well prepared for teaching?				
	P<0.001	1	2	3	4	5
Q8. How satisfied are you overall with the course?	1	30.1% (107)	25.3% (90)	29.8% (106)	11.5% (41)	3.4% (12)
	2	5.6% (34)	21.3% (129)	33.3% (202)	30.0% (182)	9.9% (60)
	3	0.6% (4)	6.1% (42)	30.9% (214)	45.9% (318)	16.6% (115)
	4	0.0% (0)	1.7% (26)	8.7% (137)	49.1% (769)	40.5% (635)
	5	0.0% (0)	0.2% (2)	1.2% (14)	17.0% (193)	81.5% (923)

		Q6. To what extent do you receive feedback and guidance from the subject teacher(s)?				
	P<0.001	1	2	3	4	5
Q8. How satisfied are you overall with the course?	1	47.6% (177)	29.0% (108)	17.2% (64)	4.3% (16)	1.9% (7)
	2	20.7% (128)	32.7% (202)	30.1% (186)	13.6% (84)	2.8% (17)
	3	10.5% (73)	24.9% (172)	37.1% (257)	21.8% (151)	5.6% (39)
	4	5.6% (86)	11.9% (181)	29.5% (449)	39.7% (605)	13.3% (203)
	5	1.4% (15)	3.9% (43)	13.7% (150)	31.8% (348)	49.1% (537)

		Q7. To what extent are you satisfied with the learning environment of the course?				
	P<0.001	1	2	3	4	5
Q8. How satisfied are you overall with the course?	1	51.5% (183)	22.0% (78)	23.1% (82)	2.8% (10)	0.6% (2)
	2	9.6% (55)	33.0% (189)	43.5% (249)	11.5% (66)	2.4% (14)
	3	3.0% (19)	15.6% (99)	55.2% (350)	21.1% (134)	5.0% (32)
	4	0.7% (10)	4.6% (66)	29.7% (423)	50.5% (718)	14.5% (206)
	5	0.2% (2)	0.9% (10)	7.5% (80)	29.7% (317)	61.6% (657)

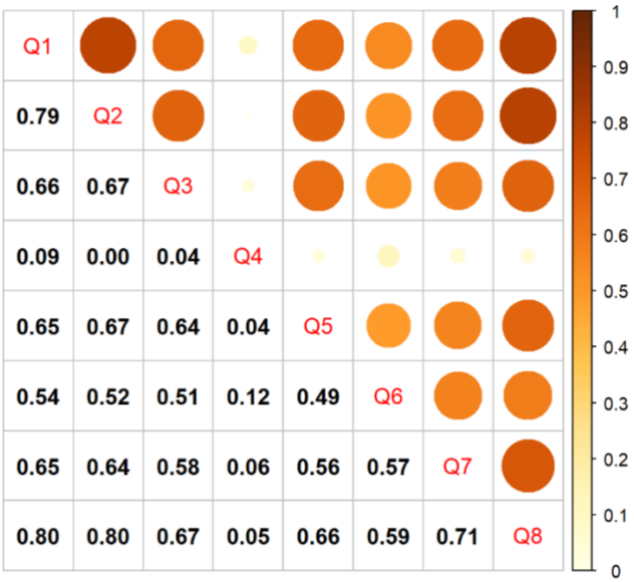

Figure 2*. Spearman correlations among all the questions in the course evaluations. The numerical values of the correlations are reported in the lower left triangle, and the strength of the correlations are visually illustrated in the upper right triangle, with darker colors meaning higher correlations.*

The calculated Spearman correlations, as a measure of the direction and degree of connection between the variables, are shown in Figure 2. Figure 2 shows the strongest correlations observed between Q1 and Q8, Q2 and Q8 and between Q1 and Q2. This indicates that the teaching and learning methods (Q1) and the dissemination of the teaching material (Q2) are the factors most strongly associated with how satisfied students are with the course. These two factors (Q1 and Q2) are also strongly correlated.

From the results, we also see that there is a clear tendency for students who give a high score on their overall satisfaction with the course to give a high score on their perception of how prepared the teacher is for teaching, Q5. There is also a clear relationship between Q3 (use of digital tools in teaching) and Q8 and between Q7 (learning environment) and Q8.

To gain more insight into the relationship between students' overall satisfaction with higher education STEM courses and their perception of the teachers' didactical practice, we look more closely at what the students say when they are commenting upon their own scoring, in both written and oral manner.

3.2 Qualitative information from open-ended comments, spring and fall 2022

The open-ended comments we have analyzed in this study consist of 1854 responses. These responses were provided by the students in relation to two open free-text opportunities embedded in the institutional student evaluation questionnaire. The first opportunity the students are given to express themselves freely relates to commenting on the first eight questions raised in the student evaluation (described above). The second opportunity was to express further reflections concerning their overall satisfaction after finishing up the whole questionnaire.

Thematic analysis of these responses revealed 13 distinct themes related to various aspects of their academic experience. Table 2 and Figure 3 present an overview of these themes, their descriptions, and key messages derived from the feedback. Using AI-assisted analysis for getting more and deeper insight out of student evaluations is an interesting and emerging research field (Rybinski & Kopciuszewska, 2020; Gallaro et al., 2023).

Teaching and Learning emerged as the most prominent theme, accounting for 37.9% (732) of all responses. Students expressed a strong preference for interactive, engaging teaching methods that promote active learning and critical thinking. They show appreciation when various teaching methods are made use of to accommodate their learning, and efforts to connect theoretical concepts to practical applications.

Course Structure was the second most discussed theme, representing 11.5% (222) of responses. In the texts students valued well-organized courses with clear progression of topics and effective integration of various learning components. On the other hand, serious concerns raised about poorly organized or disjointed course structures.

Lecture Quality comprised 10.5% (203) of responses, focusing on the effectiveness and engagement level of lectures. Students highly value well-structured, engaging lectures that make complex topics accessible and relate to real-world applications. Criticism was directed at disorganized, monotonous, or overly fast-paced lectures that fail to explain difficult subject matter.

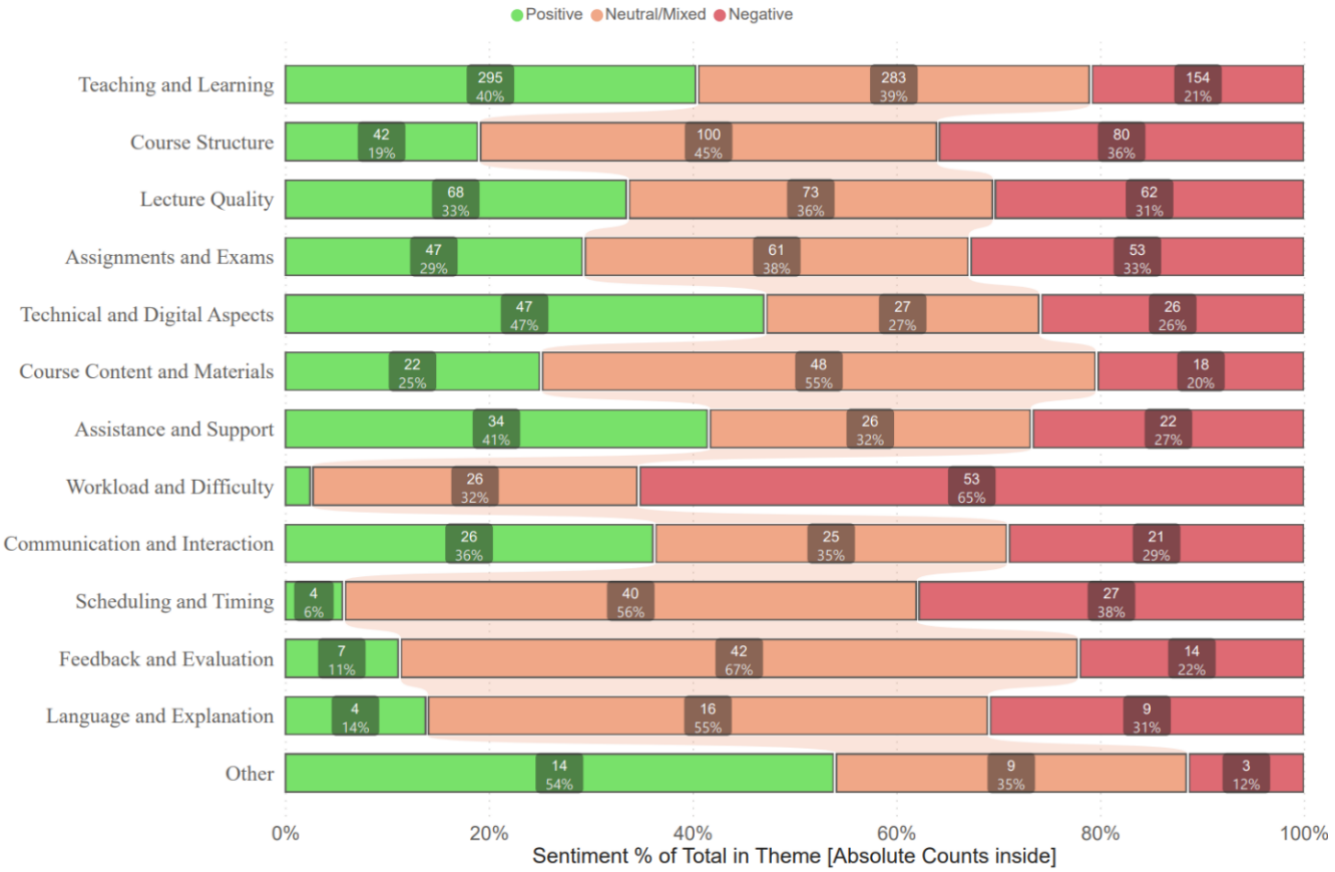

Figure 3. *Sentiment Distribution Across Identified Themes in Student Feedback. Bars represent the percentage of positive, neutral/mixed, and negative sentiments within each theme. Themes are ordered by total response frequency.*

Table 2. *Thematic Analysis of Student Feedback. This table presents the main themes identified from student feedback responses, including theme descriptions and key messages derived from the analysis.*

Theme	**Description**	**Key Messages**
Teaching and Learning	Feedback on overall pedagogical approaches and learning experiences	- Preference for active, engaging teaching methods - Appreciation for diverse teaching strategies catering to different learning styles - Desire for more practical, hands-on learning experiences
Course Structure	Comments on the design and organization of courses	- Appreciation for logical progression of topics - Need for better integration of various course components - Desire for flexibility to accommodate different learning paces
Lecture Quality	Feedback on the effectiveness and engagement of lectures	- High value placed on well-structured, engaging lectures - Appreciation for lecturers who relate content to real-world applications - Need for more interactive elements in lectures
Assignments and Exams	Feedback on course assessments, their difficulty, frequency, and relevance	- Desire for alignment between course content and assessments - Need for clear instructions and timely feedback - Appreciation for well-distributed workload
Technical and Digital Aspects	Comments on the use of technology in education	- Appreciation for well-implemented digital learning platforms - Frustration with technical issues that disrupt learning - Need for consistent and effective use of technology across courses
Course Content and Materials	Feedback on the substance of courses and learning resources	- Preference for up-to-date, relevant course materials - Appreciation for well-structured, comprehensive content - Desire for more diverse and engaging content delivery methods
Assistance and Support	Feedback on help received from faculty, TAs, and support services	- Appreciation for accessible and responsive faculty - Need for consistent quality in TA support - Desire for more specialized assistance in complex topics
Workload and Difficulty	Perceptions of course demands and academic challenge	- Desire for challenging but manageable workloads - Need for better workload distribution and coordination across courses - Appreciation for clear communication of course expectations

Theme	Description	Key Messages
Communication and Interaction	Comments on the quality of interactions within the academic environment	- Value placed on clear, timely communication from instructors - Appreciation for interactive learning environments - Need for improved communication channels and responsiveness
Scheduling and Timing	Comments on course scheduling and academic calendar	- Desire for balanced workload distribution throughout the semester - Need for better coordination of deadlines across courses - Appreciation for flexible scheduling options
Feedback and Evaluation	Perceptions of the quality and usefulness of feedback received	- High value placed on timely, detailed, and constructive feedback - Need for clearer grading criteria and more consistent evaluation methods - Desire for more personalized feedback, especially in larger classes
Language and Explanation	Comments on clarity of communication in teaching	- Appreciation for clear, concise explanations of complex concepts - Frustration with overly technical language without proper clarification - Need for diverse explanation methods to suit different learning styles
Others (Combined Minor Themes)	Diverse feedback on aspects not covered by main themes, including course relevance, general academic experience, attendance, and course evaluation processes	- Strong desire for courses with clear connections to future careers or academic pursuits - Need for addressing unique challenges faced by specific student groups (e.g., international students, part-time students) - Mixed views on mandatory attendance policies and their impact on learning - Interest in more comprehensive and transparent course evaluation processes - Suggestions for improving the overall academic experience, considering both academic and non-academic factors - Desire for more flexible policies to accommodate diverse student needs and circumstances

3.3 Focus group interviews, spring 2023

During April and May 2023, we conducted two focus-group interviews with eight students from a variety of STEM programs. The students were all third-year bachelor students. Initially, 20 students were invited to take part, based on the wish to obtain a diverse student population involving different study programs and gender. Besides these considerations, they were randomly selected. Six males and two females, representing

four different study programs, took part in the interviews. The students came from study programs in data science, electrical engineering, building construction and biochemistry. Each interview was planned to involve four students, but one student had to change the date of being interviewed.

Several aspects were highlighted as particularly important to the students in the interviews. One of these aspects was *structure and organization* as a didactical frame. The students stated that "The structure of the course and how things are aligned is very important." They also highlighted that "Canvas as a learning management system proved to be an effective platform for organizing information in a structured manner." Another aspect considered important was the *social learning environment*. As one student said, "For me, the social aspect is important", while another student pointed out that "Lack of satisfaction has to do with a poor social environment." Kharim, a 31-year-old male studying building and construction, followed up on other students' positive claims and stated:

> I would like to say that I'm very satisfied with the Canvas use. I can say five (full score 5). The Canvas use is very structured and organized. For almost all courses, we get the course description in the first week, with all the deadlines and expectations for us students. Most lecturers stream, record and upload almost all content. This suits me fine. (Focus-group interview 2nd May 2023)

Another aspect considered important among the students was *variability* in the university teachers' use of different learning methods in their didactical practice. Even so, the students in the interviews expressed a quite narrow understanding of teaching, understood as lecture-like activity going on in auditoriums. In addition, they valued a clear structure but, at the same time, more student activity. Sanna, a 22-year-old female studying biochemistry, put it this way:

> Yes, I agree with Gøran. Most of the teachers have a PowerPoint and a lecture in which they go through the presentation. And most often this is related to what we are going to learn and receive in the exam. So, in a way, it is all fine. But, for my own part and the way I learn the best, I think it's very nice to be given some questions or tasks closely connected to what the lecturer takes us through. Then you can work on this later and don't have to worry about a feeling that says, "I should have worked more on this course, but I don't know what to do." (Focus-group interview 2nd May 2023)

To gain insight into whether the students agreed, or had complementary or contrasting views, the correlations and main findings in Table 1 were presented to them. The students consistently agreed on the importance of teachers being well prepared, and that this contributes to their overall satisfaction with a course. As one student said, "I fully understand why it is like this. If, on the other hand, the lectures were not well organized, it would be difficult for the student to follow. So, clearly, the students will be most satisfied if the lectures are well organized."

However, the students do not consider a direct relationship between preparation for leading learning among university teachers and students' overall satisfaction with a course in the same way, as apparent from the quantitative analysis. The reason drawn from the focus-group interviews is that there are many aspects that result in students being satisfied with a course, and satisfaction does not exclusively depend on how well prepared the subject teachers are for the different learning activities that are part of their

didactical practice. However, the students also orally claim a strong correlation when it comes to the connection between their satisfaction and the teachers' degree of preparedness.

Roger (23 years), who studies data science, elaborated, "I think they've been well-prepared for the content they're supposed to be lecturing. But, on some courses, they've been well-prepared for what they're supposed to teach, but they don't know where we are as students" (focus-group interview 24th April 2023). The ability to "see the students" and their individual needs is something that all the students value. Roman, a 30-year-old male studying electrical engineering, went on to say:

> I am very pleased and satisfied overall. But I still think that some things could have been better. The teachers could have been more flexible than they are now. They could have paid more attention to each student and what he/she tries to do, especially in relation to students who have a particular situation – like me, who had to travel a lot. I think that, if the teachers were a little more flexible, I would have had the chance to keep up with the pace of all the other students. (Focus-group interview 24th April 2023)

Throughout the interviews, the students appeared to find it reasonable that the teachers' preparation is one of the most important influencing factors for overall satisfaction. In addition, the way teachers approach the students socially and carry out the teaching is also seen as a key issue. Anna, a 23-year-old female studying electrical engineering, underscored the importance of how the university teachers interact with her. "Some teachers take their time, greet you in a nice manner and answer your questions, appreciating that we want to learn. Other teachers are more negative, and then you don't want to ask any questions" (focus-group interview 24th April 2023). She went on to describe this kind of behavior as something that really influenced her satisfaction negatively.

4 Discussion

Through the quantitative analysis, we have seen that Q1 (teaching and learning methods) and Q2 (conveying the teaching material) have the strongest correlation with the overall satisfaction with a course (Q8). This aligns well with the open-ended student feedback and thematic analysis, where 'Teaching and Learning' emerged as the most prominent theme, revealing students' preference for active, varied and engaging teaching methods.

Considering that the relationship between Q1 and Q8 and between Q2 and Q8 is stronger than that between Q5 and Q8, it appears reasonable to say that students do not interpret teachers' preparedness for teaching as covering didactical reflections and decisions taken up front to support students' learning. It seems reasonable to assume that the students, rather, interpret a well-prepared university teacher as someone who has spent time preparing to convey content effectively through particular teaching methods. A broader didactical perspective would imply that dimensions Q1 and Q2 were embedded within Q5, and lead to a stronger relationship between Q5 and Q8 than the results indicate. The emergence and prominence of the 'Course Structure' theme in our text-analysis, do emphasize that this is something the students are concerned about. Their occupation on the 'Assignments and Exams' theme, and 'Feedback and Evaluation'

as distinct themes in the free text, also indicate that students do consider a broader didactical perspective, not fully captured in the quantitative measures. These results tell us that a holistic didactic perspective (Bjørndal and Lieberg, 1978; Hiim and Hippe, 2009) is not something that is emphasized by the students scoring on the first 7 questions in UiS' questionnaire, and that quantitative analysis alone is insufficient.

The results from the quantitative findings further indicate that a narrow understanding of what teaching is all about (Hanssen et al., 2017) still exists both on an institutional macro-level (Felten, 2019) expressed in the quality system at UiS, and on a meso level at The Faculty of Science and Technology (TN-Faculty). Academic leaders and teachers, paying attention to the satisfaction scoring of the first 7 questions, risk paying little attention to the whole diversity of the work a university teacher actually does when leading learning in higher education. At least, this is "true" in relation to our quantitative analysis of results stemming from the standardized institutional student evaluation of educational quality. As we have described, the results in the quantitative analysis show a clear relationship between how students answer Q1 and Q2 in the survey. The two questions are oriented towards the teacher's choice of teaching methods (Q1) and her ability to convey the scientific content in an understandable way (Q2). Built into the phrasing of the two questions, the focus is on what the teacher does, and her ability to convey and do her part of the job. Less attention is put on the collaborative and shared obligation students and teachers have together in a didactical meeting (Tiller, 1995; Østrem and Hanssen, 2013). This kind of consideration keeps us aware of potential biases in using student evaluations as data. But the data comes from the questions raised, and therefore we need to first look at how we phrase and construct the questionnaires. It is on this systemic level that student evaluation data are generated. There are grounds for being critical concerning the way student evaluations are constructed, carried out, and handled today. The use of particular questions and phrases will carry a distinctive pedagogical meaning, and ultimately they can generate data and lay ground for findings that can be counterproductive (Braga et al., 2014; Boring et al., 2016; Uttl et al., 2017; Feistauer and Richter, 2018; Stroebe, 2020).

To illustrate this point, it is possible to look even deeper into the two first questions of UiS questionnaire, showing the highest correlation (Q1 and Q2). The first question, "To what extent do you think that the course's teaching and learning methods contribute to your learning?" underscores that teaching is about choosing concrete methods instead of taking didactical choices as part of all the dilemmas and considerations present at the same time in planning, conducting and evaluating didactical practices. Focusing too much on this question undermines that students also learn through the assessments and assignments used, their individual and group work, work in labs, problem-solving during different projects, and through a broad range of mediating tools shared by the teachers but also found outside what is being offered at the university. In the UiS evaluation, there is no possibility to express how these different mediating tools and various didactical approaches matter. The free-text responses give students a chance to do so, but these are still open and free, and not a guided opportunity through a particular question. Interestingly though, the free-text responses analyzed here show traces that students themselves are reflective about the evaluation process. In the 'Others' category, we found examples of students saying that they have "interest in more comprehensive and transparent course evaluation processes." This suggests that at

least some students are aware of the limitations of current evaluation methods and desire more nuanced ways of giving feedback on the whole didactical practice. Secondly, a narrow understanding of what teaching is all about (Hanssen et al., 2017) is also present in the second question: "To what extent do you think the teaching in this course conveys the subject content in an understandable way?" Here, the question states implicitly that teaching is first and foremost about the conveying of a message and some knowledge by someone who already has this knowledge, to somebody who lacks it. Here a behavioristic understanding of teaching is embedded, a perspective that for a long time has been challenged by theory, namely one that puts more weight on the social and collaborative aspects of learning through sharing psychological and physical mediating tools (Säljö, 2001, 2006; Wittek, 2012). This is also something that we find traces of in the free texts. In the theme 'Communication and Interaction,' the students placed value in the communication going on between students and teachers. They emphasized their preference for on-site teaching and varied ways of engaging in face-to-face classroom situations. Here they also show appreciation of interactive learning environments, which align more with sociocultural theories of learning.

In the focus-group interviews, we also found that, even though the students consistently expressed an opinion that their teachers had been well prepared for *lecturing*, they were talking about a broader conception of this term, given a chance to reflect upon the meaning of being prepared for teaching. Here, they embraced a comprehensive understanding of preparedness for teaching, encompassing a conscious and explicit argumentation about the alignment between teaching methods, content, aims and the ability to meta-communicate (Bateson, 1972; Husebø, 2012) the rationale behind the didactical practice of the teachers. In the free-text opportunity given in the institutional student evaluation, we also find expressions that provide us with a deeper and more nuanced understanding of what the students find important for their learning. Here, we find expressions pointing to them preferring on-site teaching and varied ways of engaging in face-to-face classroom situations. Here, they also talk about important structures and didactical frames that could strengthen their learning experience and outcome.

Through the focus-group interviews, other aspects also became apparent that were not as visible through the quantitative analysis. The students emphasized the importance of the social interaction and social learning environment among students, but also stressed the importance of how they were "greeted" and met by the university teachers. The importance placed on social aspects dealt with the creation of a conducive learning environment that encouraged sharing and collaboration, along with sociocultural ideas of learning, and the ideas of teaching as didactic meetings, where students and teachers meet to understand some scientific content better together. These findings highlight the importance of a mutually engaged and committed social learning environment, where many different mediating tools are shared and brought into use. Moreover, the written free-text feedback on technical and digital aspects underscores the growing importance of integrating technology into the learning process. Specifically, this feedback suggests that teachers' proficiency with embedding digital tools in a sufficient digital learning environment has become an increasingly crucial aspect of the university teachers' didactical practice.

5 Conclusion

In this article, we have presented and discussed findings from a mixed-method study concerning students' evaluations of higher education STEM courses in the TN-faculty at UiS. We raised the research questions: 1) What is the connection between students´ report of overall satisfaction in STEM courses and the university teachers' didactical practice? and 2) What can we learn from students´ evaluations and feedback that is relevant to how we perceive and strive for teaching and educational quality enhancement in higher education?

In the article, we show that there are strong correlations between students' overall satisfaction and their evaluations of the appropriate use of teaching and learning methods contributing to their learning and teachers' ability to convey the teaching material in an understandable way. We have shown that this might be the most prominent result correlating with students' satisfaction. On the other hand, this might reflect a potential bias built into the evaluation questions, one in which the students are guided to evaluate and find their university teachers well prepared for a narrowly understood lecturing practice.

Related to this, we show that the very same evaluations and instruments used for quality enhancement face the danger of perpetuating counterproductive ways of perceiving and carrying out teaching and discussing quality in higher education. Besides being occupied by students' scoring and satisfaction on particular questionnaires, we should pay more attention to what they actually say when given the chance to communicate more freely in both texts and orally. Through the article we have shown and demonstrated how to deepen the insights generated through quantitative analysis by utilising AI-assisted analysis of free-text responses to process large volumes of qualitative data. This is a previously very labor-intensive task and rarely undertaken in the institution's quality enhancement systems. Now this is an approach that could offer an effective and more comprehensive understanding of student satisfaction in higher education STEM courses and beyond.

Finally, we argue that there is a continuous need for working on teaching quality, and thereby also striving for educational quality. This effort will never come to its end. Our take is to listen more carefully to students' voices, and the feedback they provide us, and make use of this as data in critical research. The evaluations should take as their starting point established pedagogical and didactical theory and language. They should focus on how university teachers prepare and lead the learning process explicitly and in a holistic didactical manner. When the university teachers can articulate and involve the students in regard to *what* content they are meant to learn, *how* they are supposed to learn this and *why*, it might be clearer to both students and teachers how the complex didactical relationships could be taken into consideration at the same time. This could strengthen the evaluation procedures, and once again place the center of quality discussion and systems at the heart and "moment of truth" of the didactical meetings, where study program- and course-plans are made real and experienced by both the students and the teachers.

Acknowledgements

We would like to thank Håvard Ronold, Jon Bjelland and Marit Cecilie Farsund at the University of Stavanger for providing us with access to the data. Finally, we thank the eight students taking part in the focus group interviews, and the three anonymous reviewers for their valuable feedback, which has greatly contributed to the improvement of this work.

References

Aarstad, J. (2012). Studentevalueringer i høyere utdanning: Hva kan den internasjonale forskningslitteraturen lære oss? [Student evaluations in higher education: What can the international research literature teach us?] *Uniped*, *35*(1), 34–45.

Austen, L. (2020). The amplification of student voices via institutional research and evaluation. In T. Lowe and Y. E. Hakim (Eds.) *A Handbook for Student Engagement in Higher Education*. Routledge. https://doi.org/10.4324/9780429023033

Barber, M., Donnelly, K., and Rizvi, S. (2013). *An Avalanche is Coming.* London, UK: The Institute for Public Policy Research.

Bateson, G. (1972). *Steps to an Ecology of Mind: A Revolutionary Approach to Man's Understanding of Himself.* New York: Ballantine Books.

Bergfjord, O. J. (2014). Studentevaluering i høyere utdanning, en empirisk studie fra HIB [Student evaluation in higher education, an empirical study from HIB]. *Uniped,* 37(2), 33–43. https://doi.org/10.3402/uniped.v37.22433

Bishop, D. (2018). More than just listening: The role of student voice in higher education, an academic perspective. *IMPact: The University of Lincoln Journal of Higher Education*, 1(1), 1–15.

Bishop, D., Crawford, K., Jenner, N., Liddle, N., Russell, E., and Woollard, M. (2012). Engaging students in quality processes. *Enhancing Learning in the Social Sciences*, 4(3), 1–6.

Bjørndal, B., and Lieberg, S. (1978). *Nye veier i didaktikken* [The New Paths in Didactics]. Oslo: Aschehoug.

Boring, A., Ottoboni, K., and Stark, P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *Science Open Research* 0(0), 1–11. DOI: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1

Bovill, C. (2017). A framework to explore roles within student-staff partnerships in higher education: Which students are partners, when, and in what ways? *International Journal for Students as Partners*, 1(1).

Braga, M., Paccagnella, M., and Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, 41,71–88.

Davey, T., Meerman, A., Orazbayeva, B., Riedel, M., Galán-Muros, V., Plewa, C., and Eckert N. (eds.) (2018). *The Future of Universities Thought Book*. Amsterdam, Netherlands: University Industry Innovation Network.

Dunne, E., and Zandstra, R. (2011). *Students as Change Agents. New Ways of Engaging with Learning and Teaching in Higher Education*. Bristol, UK: Escalate.

Dubey, A., Jauhri, A., et al. (2024). The Llama 3 Herd of Models. arXiv preprint arXiv:2407.21783. https://doi.org/10.48550/arXiv.2407.21783

ENQA (2015). *Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG)*. Brussels, Belgium. Retrieved from https://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf

European Commission (2013*). Report to the European Commission on Improving the Quality of Teaching and Learning in Europe's Higher Education Institutions*. Luxembourg.

European Political Strategy Centre (2017). *10 Trends Transforming Education as We Know It. European Commission*. Retrieved from https://ec.europa.eu/epsc/publications/other-publications/10-trends-transformingeducation-we-know-it_en.

Feistauer, D., and Richter, T. (2018). Validity of students' evaluations of teaching: Biasing effects of likability and prior subject interest. *Studies in Educational Evaluation*, 59, 168–178. https://doi.org/10.1016/j.stueduc.2018.07.009

Felten, P. (2019). Student Engagement in the United States. In Tanaka, M. (ed.) *Student Engagement and Quality Assurance in Higher Education: International Collaborations for the Enhancement of Learning*. Retrieved from: https://www.researchgate.net/publication/326839119_Chapter_Five_Student_Engagment_in_the_United_States_-_Felten_2019.

Felten, P. (2020). Critically reflecting on identities, particularities, and relationships in student engagement. In T. Lowe and Y. E. Hakim (eds.*) A Handbook for Student Engagement in Higher Education*, Routledge. https://doi.org/10.4324/9780429023033

Fletcher, A.F. (2017). *Student Voice Revolution: The Meaningful Student Involvement Handbook*. Common Action Publishing. https://doi.org/10.4324/9780429023033-12

Fletcher, A (2020). Going beyond student voice through meta-level education transformation. In T. Lowe and Y. E. Hakim (eds.) *A Handbook for Student Engagement in Higher Education*, Routledge. https://doi.org/10.4324/9780429023033

Fung, D. (2017). *The Connected Curriculum*. London: UCL Press.

Gaebel, M., and Zhang, T. (2018*). Trends 2018: Learning and Teaching in the European Higher Education Area*. Brussels: EUA.

Gallaro, K., Butt, S. and Ceballos, H. (2023). Improvement of Teaching Competencies Training in Higher Education Faculty Based on Student Evaluations of Teaching and AI Systems, *Perspectives and Trends in Education and Technology*, Vol.366, p.555-563, DOI: 10.1007/978-981-99-5414-8_51

Goodlad, J. I. (1979). *Curriculum Inquiry. The Study of Curriculum Practice*. McGraw-Hill.

Gundem, B. B. (2008). Didaktikk – fagdidaktikk, anstrengte eller fruktbare forhold? *Acta Didactica Norge, 2*(1). Art. 1.

Handal et al (1990). *Innstilling fra studiekvalitetsutvalget*. Oslo: Utdannings-, og forskningsdepartementet.

Hainline, L., Gaines, M., Long Feather, C., Padilla, E., and Terry, E. (2010). Changing students, faculty, and institutions in the twenty-first century. *Peer Review*, 12(3).

Hanssen, B., and Husebø, D. (2022). Tilpasset opplæring som didaktisk profesjonsbegrep i grunnskolelærerutdanningen. (Adaptive learning as didactical professional concept in teacher education). *Acta Didactica Norden*, 16(3), 22 pages. https://doi.org/10.5617/adno.9233

Hanssen, B., Husebø, D., Moen, V. (2012). "Å forelese er en ensom affære. Jeg er sikker på at de fleste går inn i klasserommet med en sommerfugl eller tusen i magen [...]) In T. Løkensgaard-Hoel, B. Hanssen og D. Husebø (eds.) *Utdanningskvalitet og undervisningskvalitet under press? – Spenninger i høgere utdanning*.Trondheim: Tapir akademiske forlag, s. 185-204

Hanssen, B., Husebø, D., & Moen, V. (2017). Universitetslærerarbeidet - utfordring av den naturlige innstilling? [The University teachers' work – challenging the natural attitude]. *Uniped*, 40(1), 7–17.

Hanssen, B., and Østrem, S. (2013). *Rutinemessig plikt eller produktiv læring. En studie av praksis og arbeidskrav i en veilederutdanning knyttet til veiledning for nyutdannede lærere*. Universitetet i Stavanger. https://uis.brage.unit.no/uis-xmlui/handle/11250/185525

Hénard, F., and Roseveare, D. (2012). *Fostering Quality Teaching in Higher Education: Policies and Practices*. Institutional Management in Higher Education. OECD.

Hiim, H., and Hippe, E. (2009). *Undervisningsplanlegging for yrkesfaglærere*. (Planning Teaching in Vocational Teacher Education). Gyldendal.

Husebo, D. (2012). Bridging theory and practice in Norwegian teacher education through action research. *Educational Action Research*, *20*(3), 455–471. https://doi.org/10.1080/09650792.2012.697665

Kahu, E. R. (2011). Framing student engagement in higher education. *Studies in Higher Education*, 38(5), 758–773.

Knight-McKenna, M., Felten, P., and Darby, A. (2018). Student engagement with community. *New Directions for Teaching and Learning*, 154, 65–74.

Kunnskapsdepartementet (2024). Forskrift for kvalitetssikring og kvalitetsutvikling i høyere utdanning og høyere fagskoleutdanning, https://lovdata.no/dokument/SF/forskrift/2010-02-01-96

Künzli, Rudolf 1997. Common Frame and Places of Didaktik. In: Bjørg B. Gundem & Hopmann, Stefan (eds.): *Didaktik and /or Curriculum – an International Dialogue* New York: Peter Lang Publishing, pp. 29–46

Lauvås, P., and Maurud, Ø. (2008). Evaluering av egen undervisning: En hjelp til selvhjelp for universitetene [Evaluation of own teaching: Self-help for universities]. *Uniped*, *31*(4), 7–32.

Loukkola, T., and Dakovic, G. (eds) (2017). *EUA's Learning and Teaching Initiative – Report from the thematic peer groups in 2017*. Brussels: European University Association.

Mausethagen, S., and Mølstad, C. E. (2014). Licence to teach? Læreplananalyse og profesjonsutvikling. In E. Elstad and K. Helstad (eds) *Profesjonsutvikling i skolen* (pp. 152– 169). Universitetsforlaget.

Meld. St. 16 (2016–2017). *Kultur for kvalitet i høyere utdanning* [Culture for Quality in Higher Education]. Oslo: Kunnskapsdepartementet.

McCowan, T. (2017). Higher education, unbundling, and the end of the university as we know it. *Oxford Review of Education*, 43(6), 733–748.

Neves, J. (2022). Using surveys to represent the student voice and demonstrate the quality of the experience. In M. Natzler *What is the Student Voice? Thirteen essays on how to listen to students and how to act on what they say*. Higher Education Policy Institute, Report 140 https://www.hepi.ac.uk/wp-content/uploads/2021/08/What-is-the-student-voice_HEPI-Report-140_FINAL.pdf

NOKUT (2016). Quality Areas for Study Programs. Report from The Norwegian Agency for Quality Enhancement and Assurance

NOKUT (2024). Studiebarometeret. Retrieved from https://www.nokut.no/studiebarometeret/

Rybinski, K., & Kopciuszewska, E. (2020). Will artificial intelligence revolutionise the student evaluation of teaching? A big data study of 1.6 million student reviews. *Assessment & Evaluation in Higher Education*, 46(7), 1127–1139. https://doi.org/10.1080/02602938.2020.1844866

Scott, C. L. (2015). *The Futures of Learning 3: What kind of pedagogies for the 21st century?* UNESCO Education Research and Foresight, Paris. ERF Working Papers Series, 15.

Speight, S., Moreira, G., and Husebø, D. (2020). Listening to students for tomorrow, today: engaging students to define the future of higher education. *Student Engagement in Higher Education Journal,* ISSN 2399-1836 3(1), 96–114.

Stokke, S., Solberg, I. L., Grøttum, P., Lundin, K. E. A., Heggen, K. M., and Breivik, J. (2019). Et system for studentevaluering av forelesninger ved medisinstudiet i Oslo [A system for student evaluation of lectures in the medical study program in Oslo]. *Tidsskriftet for den norske legeforeningen*, doi: 10.4045/tidsskr.19.0027

Stroebe, W. (2020). Student evaluations of teaching encourages poor teaching and contributes to grade inflation: A theoretical and empirical analysis. *Basic and Applied Social Psychology*, 42(4), 276–294. https://doi.org/10.1080/01973533.2020.1756817

Strømsø, H. (2016). Forførende entusiasme: 40 års forskning på Dr. Fox effekten [Seductive enthusiasm: 40 years of research on the Dr. Fox effectin Norwegian]. *Uniped tidsskriftet*, 29(2), 118–130. https://doi.org/10.18261/issn.1893-8981-2016-02-03

Säljö, R. (2001). *Læring i praksis. Et sosiokulturelt perspektiv* [Learning in Practice. A Sociocultural Perspective]. Cappelen AkademiskeForlag.

Säljö, R. (2006). *Læring og kulturelle redskaper* [Learning and Cultural Tools]. Cappelens Forlag AS.

Tiller, T. (1995). *Det didaktiske møtet – et møte mellom fag og hverdag: grunnlaget for en lærende skole*. Praxis forlag, ISBN 9788290830033

Uttl, B., White, C. A., and Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching rating and student learning are not related. *Studies in Educational Evaluation*, 54, 22–42.

Wittek, L. (2012). *Læring i og mellom mennesker* [Learning in and between Human Beings]. Cappelen Damm akademiske.

Worum, K. S., Hansen, I. M., and Olsen, R. (2022). Omvendt undervisning som talesjanger i veilederutdanningen – studenters perspektiv på forhold som har betydning for læringsprosesser [Flipped classroom as speech genre in mentor education – students' perspectives on factors influencing learning processes]. NORDVEI – *Nordisk tidsskrift i veiledningspedagogikk*, 7(1). https://doi.org/10.15845/ntvp.v7i1.3792

6 Appendix

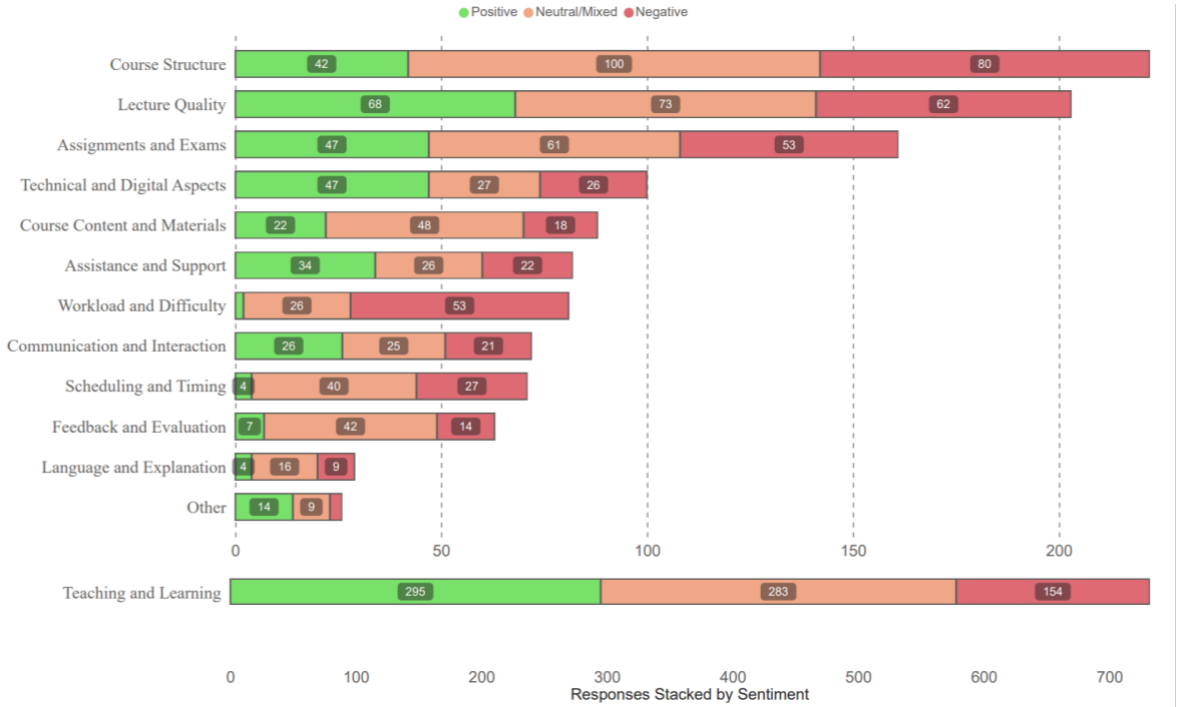

***Figure S1**: Absolute Counts of Sentiment Categories Within Identified Themes. Bars show the absolute number of responses categorized as positive, neutral/mixed, or negative for each theme. Themes are ordered by total response frequency. Teaching and Learning plotted **separately** to enhance visibility.*