



Behind the scenes: Unpacking student discussion and critical reflection in lectures

Kristine Ludvigsen , Rune Johan Krumsvik  and Jens Breivik 

Kristine Ludvigsen is an assistant professor. Her research interests are how the usage of digital tools facilitate formative assessment and feedback practices. Rune Johan Krumsvik is Professor of Education, and Head of the DLCs at the University of Bergen. Krumsvik is also Honorary Research Fellow at the University of Bristol. His main research interests are formative assessment, digital learning and doctoral education. Jens Breivik is associate professor in education. His area of research is teaching and learning in higher education. His background is in scholarship of teaching and learning in HE and philosophy of education. He has a particular interest in analysis of students' critical thinking and argumentation skills. He is a member of the research group Pedagogical Philosophy at The Arctic University of Norway. Address for correspondence: Kristine Ludvigsen, Department of Education, University of Bergen, Cristiesgt. 13, Box 7807, Bergen 5020, Norway. Email: Kristine.Ludvigsen@uib.no

Abstract

This study investigated the characteristics of peer discussions used to support formative assessment in lectures, facilitated by a student response system, in an undergraduate qualitative methods course for psychology students. The intent was to examine the characteristics of peer discussions in which student response systems are used to facilitate the practice of formative assessment lectures. The research was guided by the following research questions: (1) What patterns of talk can be identified in the discussions? (2) How do the students use subject-specific vocabulary in the discussions? (3) How the students' understanding of the subject matter displayed in these discussions? To examine the characteristics of peer interactions, 87 student discussions were recorded and analysed. The concept of *exploratory talk* was used as a lens to examine the discussions. In 68 of the 87 discussions, the students exchanged ideas and elaborated on their peers' ideas and understanding of the concepts. In the remaining 25 discussions, the process of reasoning was less visible. The findings are relevant for teaching designs that aims to use digital tools to facilitate formative assessment.

Introduction

This study investigated the characteristics of peer discussions facilitated by a student response system (Turning Point) to support formative assessment in lectures, in a qualitative methods course for undergraduate psychology students. A key process in formative assessment is to make learning visible through “moments of contingency” (Black & Wiliam, 2009, p. 10), situations and activities in which students are encouraged to articulate their thinking and understanding, so that information generated from these activities can be used to shape ongoing teaching and learning activities. Furthermore, formative assessment should support self-regulated learning (Boud & Soler, 2016; Clark, 2012; Evans, 2013). A vital question when examining formative assessment activities is, to what extent do they provide students with opportunities to share their thinking, and what opportunities become available to students and teachers to draw inferences through these activities to shape teaching and learning (Furtak et al., 2016)—referred to as moments of contingency.

Practitioner Notes

What is already known about this topic

- Student response systems can support formative assessment and feedback in lectures.
- The most common approaches used in research on student response systems to support formative assessment is questionnaires or interviews.
- Few studies have provided detailed analyses of clicker-supported peer discussions.

What the paper adds

- Provides insights into the micro-processes in clicker-supported discussions.
- Critically discusses the role of discussions facilitated by student response systems to support formative feedback in the classroom.
- The paper contributes to scholarship in this field by drawing attention to the qualities of the activities created for formative assessment.
- Discusses the validity of the inferences drawn from the use of student response systems in the classroom.

Implications for policy and practice

- Analysis of discussions shows that there is not necessarily a correlation between aggregated answers and students' understanding. To ensure that valid inferences can be drawn from activities as a basis for feedback, the questions used and threats to the validity of possible inferences should be critically examined.

In lectures, formative assessment activities often include the use of different kind of student response system that elicit quantitative or qualitative answers to questions. The value of student response systems to support formative assessment lies in technological opportunities to make the reflections of the students visible through interaction and problem solving (Egelandstad, Ludvigsen, & Ness, 2019). Hattie and Timperley (2007) conceptualised feedback as “information provided by an agent (a teacher, a peer, a book, a parent, oneself, experience) regarding aspects of one’s performance or understanding” (p. 81). Asking questions and explaining their thoughts could create awareness and increase the students’ understanding of the topic under discussion. Therefore, participation in discussions provides opportunities for receiving feedback at both the task, often corrective of individual or group performance (p. 95) and self-regulation level, level refers to the students’ ability to monitor and to regulate their own learning processes, including opportunities to “create internal feedback and to self-assess” (p. 95) their progress towards achieving learning goals, which is essential for students ability to make decisions about their learning.

A growing body of empirical research suggests that the use of student response systems in lectures can enhance both the quality and the quantity of peer discussions (Chien, Chang, & Chang, 2016). Participation in discussions allows students to argue and to provide justifications, to question their own and others’ assumptions, to co-create knowledge, to explain and to clarify the subject matter. When students make their thinking explicit in peer discussions, students will be exposed to different ways of thinking, which can help them to become aware of their own understanding and make better-informed decisions about their learning process (Dawson *et al.*, 2018).

Using student response systems provide lecturers with an awareness of their students’ understanding of the course material, which supporting a contingent teaching approach in lectures

(Chien *et al.*, 2016; Dawson *et al.*, 2018; Hunsu, Adesope, & Bayly, 2016; Liu *et al.*, 2017). Often it is assumed that a straightforward relation between the collection of answers and the process of providing feedback:

One of the key benefits of using an ARS is that instruction can be modified based on student feedback gathered throughout a class (...) If feedback from a majority of students indicates that confusion or misconceptions are evident, an experienced instructor can offer alternative explanations of the concepts in question. (Kay & LeSage, 2009, p. 822)

Inferences are based on what we can observe, and thus, they are characterised by uncertainty (Bennett, 2011). Findings derived from recordings of student discussions suggest that voting for the correct option did not necessarily demonstrate an understanding of the topic under discussion, and an incorrect answer did not necessarily indicate inadequate understanding of the concept (Knight, Wise, & Southard, 2013; Nielsen, 2012; Wood, Galloway, Hardy, & Sinclair, 2014). This can lead to misleading feedback, both for students and for lecturers (Nielsen, 2012), which might problematise the validity of the inferences drawn from multiple-choice clicker questions designed for formative assessment in lectures. Studies that have used recordings of peer discussions have found that instructing students to argue, instead of simply ask student to discuss can improve argumentation (McDonough & Foote, 2015). The evidence suggests that groups engage in more argumentation when they are required to justify a position rather than merely to discuss a topic (Knight *et al.*, 2013) and that an initial thinking period increases the number of arguments in a discussions Nielsen, Hansen, & Stav, 2016). In identifying the features of high- and low-quality discussions (based on the level of reasoning in each discussion), Knight *et al.* (2013), Knight *et al.* (2013) and James and Willoughby (2011) each found that the quality of discussions was not dependent on the cognitive level of the questions posed.

This current study

The research reported in this paper is a part of a longitude design-based research project (Barab & Squire, 2004) that explored the role of discussion-based activities in supporting a dialogue approach to formative feedback in the lecture. It also assessed the quality of dialogue among students and between students and lecturers. The point of departure for the project was to address a challenge identified in practice as well as by theory and prior research: the transformation of the lecture from a mode of transmission to a format that includes student active learning approaches. For this purpose, the project aimed to incorporate case-based activities in lectures to promote critical reflection, to connect the course material to the students' own language and experiences and to provide opportunities for feedback.

In three previous studies based on survey and interviews from the same course (Krumsvik & Ludvigsen, 2012; Ludvigsen, Krumsvik, & Furnes, 2015; Ludvigsen, Ness, & Timmis *et al.*, 2019), students claimed that the quality of the discussions had improved, compared to discussions without the support of a student response system. The reason for this was that the alternatives in the questions posed helped structure the discussion, and the fact that they had to submit a response made the activities felt more authentic than discussions that did not require a response (Krumsvik & Ludvigsen, 2012). Discussions among peers were highlighted as valuable for reflecting on one's own understanding. One student explained: "I notice that I cannot answer the questions until I discuss them out loud [...] You argue with someone about why [your ideas] are right, and then suddenly you find arguments for why it is right and why it is wrong" (Ane) (Ludvigsen *et al.*, 2015, p. 47). The act of explaining to a peer is a way of explaining their perspectives to themselves: "Even though you remember the words, then you should explain it to others, then they ask what it means, and then you realise that you did not know, then you notice" (S3) (Ludvigsen

et al., 2019, p. 11). Articulating their thinking through explanations to their peers helps students to reflect on their learning, as stated by another student: "I can sit and read or hear and believe that I understand these things. But, if you are to formulate yourself, with no help in front of you, then I realise if I understand" (S2) (Ludvigsen *et al.*, in press, p. 11). The quotations illustrate how students are drawn into reflection on their own thinking and understanding in which are vital in establishing a formative feedback practice in lectures. According to students, discussions provide opportunities for sharing perspectives, arguing, explaining and listening, which are all indicators of high-quality discussions (Hennessy *et al.*, 2016). The discussions were short, lasting only 1 or 2 minutes. This created the desire to explore them in greater detail. What would argumentation and explanation look like in these discussions?

In their meta-study of how use of student response systems support learning, Chien *et al.* (2016) thus argue that peer discussion should be examined from a social aspect: "Future studies are needed to investigate how students interact with peers within the context of clicker integrated instruction research on this line will also be helpful to understand how the use of student response systems mediate the process and outcome of peer discussion" (Chien *et al.*, 2016, p. 15). Such studies are also essential because there might be a gap between how activities play out and the lecturers' and students' perceptions of them (Nielsen, Hansen, & Stav, 2012). By exploring the activities as they unfold might give valuable insights into discussions supported by response systems that are not possible to identify using other methods. Recording students' discussions is important for understanding what is achieved, to arrive at a nuanced understanding and identify affordances as well as constraints to be assessed in relation to the purpose of using the technology. Recording discussions helps us pay attention to the qualities of the new spaces these tools might provide. Knowledge of the characteristics of peer discussions stimulated by multiple-choice questions is also essential to inform practice and to promote high-quality discussions (Barth-Cohen *et al.*, 2016; Knight *et al.*, 2013; Wood *et al.*, 2014), understanding how tools shape feedback practices is crucial for research into feedback in higher education in general (Evans, 2013).

The goal of this study thus, was to examine the characteristics of peer discussions facilitated by student response systems to support formative assessment in the lecture. The research was guided by the following research questions: (1) What patterns of talk can be identified in the discussions? (2) How do the students use subject-specific vocabulary in the discussions? (3) How the students' understanding of the subject matter displayed in these discussions?

By exploring these questions, the paper offers insights into the characteristics of peer discussions supported by student response systems and situating the research within the context of formative assessment.

Context, methods and analysis of the discussions

The context was a qualitative methods course for undergraduate psychology students. For students to judge the choices that must be made when approaching the qualitative research process, critical reflection is crucial (Cooper, Chenail, & Fleming, 2012; Cooper, Fleisher, & Cotton, 2012). To provide hands-on experience and to address such core concepts as sample, validity and triangulation, the students were engaged in a discussion of authentic case questions. Each lecture started with the introduction of a theme or concepts (a "mini-lecture"). A question for peer discussion followed; the students were given between 1 and 2 minutes to discuss each question. Each student was given a device ("clicker") to respond ("vote") to the multiple-choice questions

through a student response system (Turning Point). The answers were aggregated and projected onto the screen for clarification and whole-class discussion.

Two categories of questions were used. In the *textbook questions*, the students were asked to answer a multiple-choice question that addressed the definitions, basic concepts and characteristics of qualitative methods as presented in assigned readings. In the *case questions*, the students answered a multiple-choice question that addressed the procedures involved in qualitative methods and their application in authentic contexts. Both categories included questions that are allowed for multiple responses (“multiple responses”) and others that had only one correct answer (“closed”). Figure 1 illustrates a case question (Gibbs, 2013), and Figure 2 presents the response.

Each lecture included sequences of four to six questions. The lecturer (the second author of this paper) is a professor with extensive experience in the use of technology to support and enhance

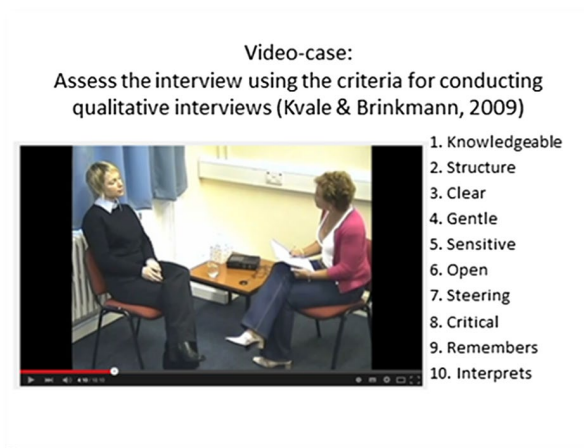


Figure 1: A question allowing multiple responses
[Colour figure can be viewed at wileyonlinelibrary.com]

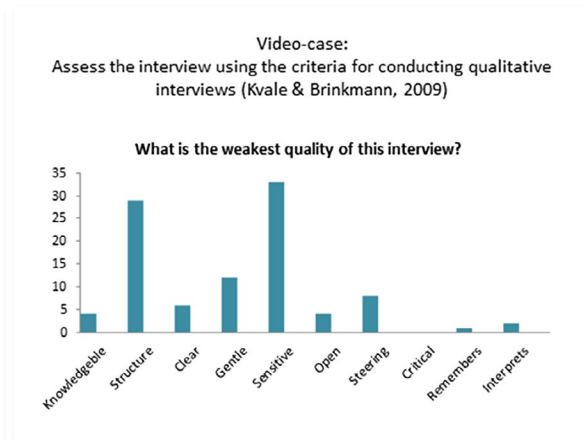


Figure 2: The corresponding responses
[Colour figure can be viewed at wileyonlinelibrary.com]

teaching and learning in higher education. Approximately 20 minutes of each lecture was devoted to discussion-based activities. Before the discussions, the students were asked to argue for their views; however, they were not provided guidelines on how to engage in exploratory talk. The activities were supposed to be dialogic in that they provided students with the space to engage with the material in the peer groups (Wegerif, 2013); they were also designed to gain insights into the students' thinking (to create moments of contingency). However, there is some tension between the infinite possibilities for multiple voices to be part of a dialogue and the closure needed to achieve structured learning objectives. This tension includes completion of the assignment within a teaching design developed to facilitate formative assessment through questions with only one correct answer.

Data collection

Discussions were gathered from six lectures spread over two semesters. In both classes, undergraduate students were given an introductory course in qualitative methods, with the same lecturer, and they discussed the same types of questions using the same course material and within the same context. Discussions were conducted over two semesters of classes so that more discussions could be included in the study than we were able to collect during only one semester. The sample was based on voluntary participation; thus, it was a convenience sample (Creswell, 2012). The students received written and verbal information about the study before the lecture began. Audio recorders were provided at the beginning of each lecture. Students who chose to participate received a short briefing on the recorders; they then decided whether to record a discussion. While 96 discussions were collected, not all of the recordings could be analysed because of audio quality issues. The analysis included 87 discussions distributed across 21 different questions asked in class. About 38 of the discussions addressed questions allowing multiple correct answers, and 50 addressed questions with one correct answer.

The study was performed in the authentic context of a lecture. Students were asked to discuss with students sitting beside them; for that reason, the discussion groups consisted of two to four students each. The number of students in each group depended on how they were seated in the auditorium, which reflects common practice in these lectures. An ethical question arose that the students might have found the questions stressful to discuss or that they might not speak as freely as they would have if the discussions had not been recorded. However, as the students were learning about qualitative research methods in this class, volunteering to participate in interviews would give them valuable experience in understanding the challenges and opportunities of using audio recordings of discussions as a research method.

Data analysis procedures

To analyse the discussions, we used five strategies. First was listening to the recordings several times to become familiar with the discussions. The material was then transcribed. Most of the discussions were short: between 40 seconds and 2 minutes. Second, the transcripts were read to examine their characteristics. Patterns of talk were coded using the framework of exploratory talk (Mercer, 2004). In the supplementary material, examples of this coding are provided. The transcripts showed that the students used subject-specific vocabulary in different ways. Three codes emerged through the reading of the discussion transcripts: no use of subject-specific vocabulary, references to subject-specific vocabulary and application or definition of subject-specific vocabulary. A majority of the discussions revealed student uncertainty. Questions and statements before their "vote" reflected the emergence of the students' questions and insecurities. Each of these occurrences was coded as "uncertainty." In the discussions, in which the students knew the answers, indicated their subjective certainty, or stated that the assignment was easy, the transcripts were coded as "confidence." Finally, the distribution of (1) uncertainty and (2) subject-specific

vocabulary across the patterns of talk were examined. The steps in this analysis are illustrated in Figure 3. In Appendix B, examples of this coding are provided.

Analytical framework for examining exploratory talk

Several frameworks are available for analysing the quality of student discussion (for a review, see Hennessy *et al.*, 2016). These frameworks can be distinguished based on their emphasis on changes in individual thinking, emergent understanding within a group or single utterances or episodes of talk (Mercer, Littleton, & Wegerif, 2004). The goal in this study was to examine the students' sharing of ideas and their processes of collaborative knowledge building. The unit of analysis was the discussion, and the patterns of talk were examined through sociocultural discourse analysis (Littleton & Mercer, 2013). This methodology is suited to analysing patterns of talk among participants engaged in problem solving (Mercer, 2004). It uses a quantitative approach to compare discussions under different conditions and a qualitative approach to examine student engagement in idea sharing and knowledge co-construction in a specific context (Mercer, 2004). As an indicator of the quality of education talk, Mercer (2004) refers to three modes as prototypes: *disputational*, *cumulative* and *exploratory*. Exploratory talk is characterised by the discussants' critical engagement with one another's ideas. The arguments or the reasoning is explicit or accountable in the discussions. The students offer alternative views or hypotheses, and they participate with the purpose of "joint consideration" (Mercer, 2004, p. 46). Cumulative talk is characterised by "repetitions, confirmations and elaborations" that build on one another uncritically (Mercer, 2004, p. 46). By contrast, disputational talk is characterised by disagreements, interruptions and individual decision making. Students do not ask follow-up questions or make additional contributions.

Their utterances are short, lack justification and are often confrontational (Mercer, 2004). The idea of "exploratory talk" is associated with the idea of a "dialogic space" (Wegerif, 2013). In a dialogical space, cumulative talk provides a widening of the space, and exploratory talk provides a deepening of the space (Wegerif, 2013).

The framework of exploratory talk is mostly used when examining the quality of talk in schools (Littleton & Mercer, 2013), however, the same pattern of talk has been identified in studies examining the quality of educational dialogue in the context of higher education (Havnes, Christiansen, Bjørk, & Hessevaagbakke, 2016) and in work place settings (Littleton & Mercer, 2013). Since this framework are mostly used within a school setting, our study is a contribution to bring this literature into analysis of student peer discussions in higher education.

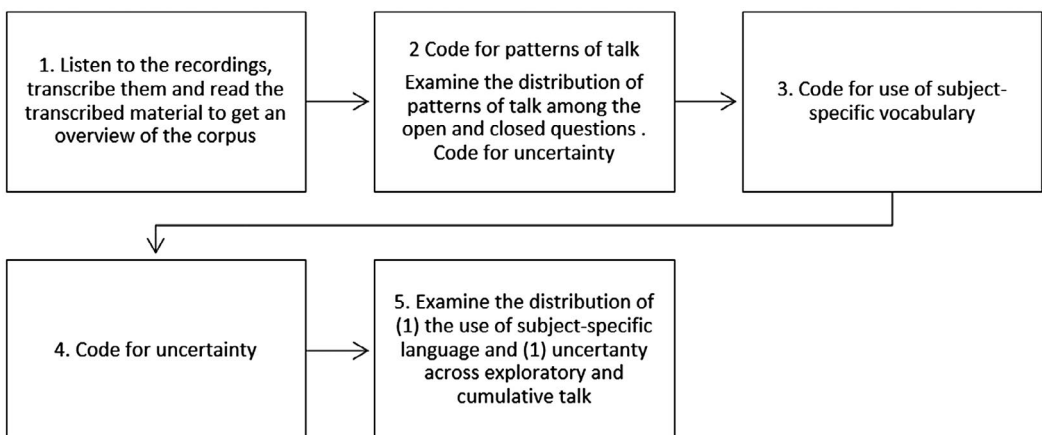


Figure 3: Data analysis procedure

The concepts of exploratory talk were operationalised in this study through the coding scheme (Appendix B). This coding scheme was inspired by the “Cam–UNAM Scheme for Educational Dialogue Analysis (SEDA: ©2015), developed by Sara Hennessy and Sylvia Rojas-Drummond” (Hennessy *et al.*, 2016, p. 42). The purpose of the scheme developed by Hennessy *et al.* (2016) was to “distil out the essence of dialogic interactions and operationalise them in the form of a new scheme of systematic indicators for these productive forms of educational dialogue.” (p. 42). The SEDA-framework thus describes indicators of qualities of talk (The complete coding scheme can be found at <http://tinyurl.com/BADialogue>). Four of these (invite elaboration or reasoning; build on other's ideas; make reasoning explicit; positioning and coordination) are from our point of view congruent with qualities described as exploratory talk in the literature (Mercer, 2004). Furthermore, it was suited for the material and we found that description of each of the categories were useful for examining qualities in our own context (peer-discussions). We used the condensed scheme as presented on the web page, however, with some adjustments, for example, that qualities described in “positioning and coordination” are put under the heading “make reasoning explicit.”

To limit possible biases in researching our own practice (Sikes, 2006) one researcher outside of our own institutions has been involved with the analysis of the discussion. There were two coders. To ensure their understanding of the coding scheme, the coders coded 10 discussions together. The coders then did a close reading of all the discussions and coded each transcript individually regarding “pattern of talk.” NVivo 11 was used to organise the coding of the discussions. Before analysis, each transcript was coded by assignment type (questions with one correct answer or multiple correct answers and textbook or case questions) in accordance with the coding scheme. In some cases, a discussion was characterised by more than one pattern of talk. They were coded as exploratory if indicators of exploratory talk were present in the transcript. Last, inter-rater reliability for patterns of talk was evaluated using Cohen's kappa coefficient, which was chosen because of the use of categorical rather than continuous coding. The inter-rater reliability, calculated using IBM SPSS Statistics, was a Cohen's kappa coefficient of 0.69, which could be considered a substantial fit (De Wever, Schellens, Valcke, & Van Keer, 2006). Disagreements were resolved after discussion. This was performed by reading the transcript aloud; then, each of the coders provided arguments for their interpretations. In this way, the transcript was discussed carefully, and through this process we also compared these discussions with other discussions we agreed upon. This made it easy to come to an agreement and in this way, we resolved each of the disagreements. To secure the validity of our conclusions and to be open for different possible ways of interpreting the discussion transcript, we have also presented excerpts and analysis of the peer discussions in several workshops involving experts in the field of higher education and in interaction.

Findings

The following section is in five parts. First, we provide a general characteristic of the material. Second, we provide examples of patterns of talk identified in the material. The third section discusses the students' application of subject-specific vocabulary. Fourth, we present how uncertainty was expressed in the discussions. The fifth section presents the distribution of subject-specific vocabulary and occurrences of uncertainty across the different pattern of talk.

General characteristics

Students used the answer options to structure the discussions, and they argued about the merits of each. In general, the discussions were of high quality. Another characteristic was the students' completion of one another's sentences, a phenomenon that was interesting in this context because it revealed the students' collaborative thinking processes as presented in the examples below:

Example 1a

<i>Student</i>	<i>Utterance</i>
S1:	The interview guide. It is the questions
S2:	It is the interview question
S3:	In a way, it is the order

Example 1b

<i>Student</i>	<i>Utterance</i>
S1:	You can measure stress or
S2:	Pulse or
S3:	Perspiration

Characteristics of cumulative and exploratory discussions

Of the 87 discussions, 25 were coded as cumulative talk. In some of the cases, the students knew the answers; therefore, discussion was not required. In others, they focused more on finding the right answer than making their reasoning visible to their peers: they did not justify or explain their claims. Only a few students addressed clarifications or justifications of the arguments, and in some cases follow-up questions were not asked. Subject-specific terms were used only at a superficial level, and the reasoning was visible only to a limited degree, as indicated in Example 1.

Example 1 Cumulative talk

<i>Student</i>	<i>Utterance</i>
S1:	It is not Two. It has to be One or Two? One?
S2:	Yes.
S1:	It cannot be Two.
S2:	No, because they are going to find out...
S1:	One, or Three? Hmm.
S3:	I think it is Number One.
S2:	Should we choose One then?
S1:	I think so. I'll go for that.
S2:	It is not Number Two. And it is not Number Three.
S3:	Then it has to be Number One.
S1:	I choose Number One.

In Example 1, the discussion was characterised by *cumulative talk*. The students did not use subject-specific terms. They suggest answer alternatives by referring to their number, rather than explaining and presenting reasons for their suggested alternative. The dramaturgy was simple: (1) opening, (2) suggestions for voting and (3) votes.

Exploratory talk

In 62 of the 87 discussions, the students exchanged ideas and elaborated on one another's ideas and understanding of course concepts. In making their ideas visible, the students enabled others to connect to these ideas, to build on them, to criticise them and to argue for or against them, thereby allowing for the development of multiple perspectives. The most striking characteristic of these discussions was students building on one another's arguments to come to a consensus. This was evident when they were arguing for one another's claims and completing one another's sentences. Second, the students' reasoning was visible to everyone in the group. This feature became evident in their justifications of their own or their peers' claims. Third, they asked for clarifications, or they addressed concepts that were unclear. However, there were only a few examples of the students being critical of one another's arguments. In the next discussion (Example 2), the group discussed the question after watching a news report about the off-task use of information and communication technology (ICT) in upper-secondary schools. The question invited the students to identify the features of quantitative and qualitative research questions.

Example 2 Exploratory talk*Student Utterance*

- S1: What do you think?
 S2: I think Number Four?
 S1: Why?
 S2: Because, the relationship is quantitative, and to *what extent* it is also quantitative, and the first is also totally yes or no questions.
 S3: But, how does the teacher experience the student's off-task ICT use in class, is qualitative, and then they do not problematise, and then it is not normative. Definitely, so? What do you think?
 S2: I do not know, and I have not started the process yet. I spent so much time reading this thing. Yes, I think Four is the best option: the point that it is particular. Experience?
 S3: What is?
 S2: It is qualitative.
 S3: Phenomenological.

VIDEO CASE: 'What is adequate for this teacher's experiences as well as in line with qualitative research questions?'



1. Is Facebook in class a disturbing factor for teachers?
2. To what extent do teachers find Facebook to be a disturbing factor in their lessons?
3. Is there a relation between student use of Facebook and the teachers' class management?
4. How does the teacher at 'Mountain school' experience student' use of Facebook in class?

Example 2 was characterised by *exploratory talk*. S2 asked for a justification for choosing Number Four. The sequences of the talk that brought the discussions into the explorative mode were characterised by the opening of the discussion space when S1 asked for a justification for S1's claim, "I think Number Four." S2 and S3 followed up by introducing new arguments in support their shared conclusion. The discussions exhibited a typical dramaturgy. First, one of the students invited the others into the discussion, and the students presented their immediate thoughts without explanations or justifications. This mode changed if or when problems were encountered or questions arose; to progress in these situations, the students needed to make arguments or explain their thoughts. This was the most exploratory part of the discussion. This confirms previous findings that exploratory talk was associated with productive ways of addressing the object of learning (Havnes *et al.*, 2016; Littleton & Mercer, 2013).

Table 1: Patterns of talk by question allowing one correct answer and multiple correct answers

	Cumulative (25)	Exploratory (62)
Questions with one correct answer (50)	21	29
Questions allowing multiple correct answers (38)	4	35

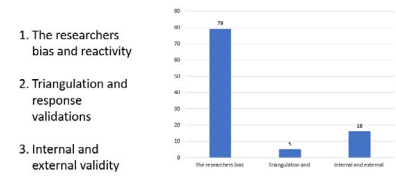
Table 1 displays the distribution of the patterns of talk engendered by the closed-ended and open questions. The discussions featuring exploratory talk were most likely to be generated by the questions allowing multiple correct answers,

Investigating the differences among the groups was beyond the scope of this study. However, it was noteworthy that while some groups engaged in exploratory discussions regardless of the questions, others addressed all of the questions superficially.

Use of subject-specific terms

There were discussions in which the students used no subject-specific vocabulary (3), referred to a course-related concept without any elaboration (33) or applied or defined the subject-related concepts they mentioned (51). In the discussion provided as Example 3, the assignment was to discuss the following question: What are the two most common threats to validity in qualitative research? The possible responses were the following: (1) researcher bias and reactivity, (2) respondent validation and triangulation and (3) internal and external validation.

Example 3 The students applied and defined subject-related concepts

Student	Utterance		
S1:	It has to be a ... Excuse me. What do you think?	<p>What are the most common validity threats in qualitative methods?</p>  <p>1. The researchers bias and reactivity 2. Triangulation and response validations 3. Internal and external validity</p>	
S2:	The researcher's trustworthiness? The influence of the researcher?		
S1:	Yes. Then, it is the two. Yes. It is a common threat with the method in general, is not it?		
S2:	However, particularly, in the qualitative, I guess? I think One.		
S1:	For triangulation.		
S2:	What is it?		
S1:	It is something you should use to get good validity, to have different entrance angles, different methods, for example. The way she is here now and making the recording, she could have chosen to ask us how it was to discuss the lecture.		S1 explains the concept of triangulation.
S2:	Moreover, got a response and used it as an answer. But, when she also records		S1 provides an example
	What is the truth, that strengthens the validity in a way, so it is the opposite.		
S1:	It increases the quality.		
S2:	Yes. It is not a threat.		

The two students discussed the reasons why triangulation was not a threat to validity in qualitative research. This demonstrated the process of proposing both a definition and an example. The students were inviting others into the discussion space and making their thinking visible by both posting questions (S2) and defining the concept and providing examples (S1).

An examination of the distribution of subject-specific terms across the patterns of talk indicated that the discussions coded as exploratory talk were more likely to feature application, definitions or examples than the discussions coded as cumulative (Table 2).

Uncertainty

Most of the recorded discussions revealed instances of the students' uncertainty. Below is an example of a discussion on the role of the Helsinki declaration.

Example 4 How uncertainty was expressed

<i>Student</i>	<i>Utterance</i>
S1:	Nuremberg. Is it something to do with World War Two?
S2:	I do not know what it is.
S1:	I think it is One.
S3:	Vulnerable group? What is that?
S1:	What does that mean?
S4:	I do not know, and I just think Nuremberg. I feel it was something with a trial with Nazism.
S1:	I do not know. I remember from history ... because of all the things that happened in the concentration camps.
S2:	That it is vulnerable groups do you think?
S1:	Possibly.
S2:	Anyway, what are vulnerable groups?
S1:	Yes. That also. Is it children?
S4:	Yes. This needs to be specified.
S2:	Is it minorities, Sami?

In this example, we see that each perspective that was added to the discussions engendered a new question.

In the corpus, there were as many as 20 occurrences of students' explicit statements of not understanding the topic at the moment before giving their votes. The following are examples: "Honestly, I don't have the slightest idea." "We're gambling." "I say One, and then, what if it's wrong?" "Frankly, I don't know. My heart says Four." "Ah, I don't want him to ask me to explain this." "We say Three, and then we have to sink into a hole in the ground if it's wrong." "I guess...." "Should we each take One?" These episodes in the discussions are interesting and important. They indicate that even though the students might have chosen an answer, they had expressed uncertainty about whether it was right or wrong. Table 3 presents the occurrences of articulations of uncertainty in the questions with multiple correct answers and questions with one correct answer.

Summary of the findings

The analysis showed the following:

- Almost all of the discussions focused on the assignment.
- The students expressed uncertainty in a majority of the discussions (68 of 87), and insecurity was evident with all types of questions.
- The students structured their discussions around the answer options.

Table 2: Use of subject-specific vocabulary in the discussions

	No use	Reference	Application/Definition/Examples
Cumulative (25)	3	20	2
Exploratory (62)	0	13	49

Table 3: Expressions of uncertainty in the discussions

	Confidence	Lack of confidence
Cumulative (25)	5	20
Exploratory (62)	14	48

- The alternatives were used both to open discussions (to clarify and explain concepts) and to shut down discussions (by referring to numbers only).
- Characteristics of exploratory talk were identified in 62 discussions.
- Most of the exploratory discussions were generated by questions allowing for more than one correct answer.
- The use of subject-specific terms beyond mere references to a concept was most likely to be found in discussions with characteristics of exploratory talk.

The next section focuses on the role of discussions in formative feedback practice. Guidelines for practice and further research are suggested.

Discussion and conclusion

This paper contributes to research in its field by providing insights into the dynamics of classroom peer discussions facilitated by educational technology. The study demonstrates the potential of clicker-supported discussions to promote critical reflection on the concepts in qualitative methods. By connecting *moment of contingency* to *exploratory talk*, we draw attention to the quality of the activities employed to support formative assessment in lectures. The idea *moments of contingency* emphasise that such activities should be used to support and adjust learning and teaching activities. Analysing *moments of contingency* using the framework of “exploratory talk” enable us to critically examine the quality of the dialogues enabled by this teaching design. This helps us to focus how such activities can enhance contingent teaching and formative assessment.

Exploratory discussion must be valued and promoted. Thus, the alternatives provided are an important factor in stimulating discussion. Our study shows that the alternatives offered in multiple-choice questions have the power both to trigger and open discussions, and to limit discussions. In some of the analysed discussions, students opened up dialogues by using the alternatives as means for clarifying concepts or arguing for their view. In other of the analysed discussions, we found that the alternatives discourages the students from articulating their knowledge and sharing their thinking. For example, some students were simply guessing at an answer without elaborating, or they told each other what to vote for an alternative without offering an explanation, only referring to numbers: “sure, it has to be c.” This is similar to findings by Wood *et al.* (2014), Knight *et al.* (2013), James and Willoughby (2011) and McDonough and Foote (2015).

When students use the MC-alternatives this way, this leads them to superficial approach to their learning, rather than opening up spaces for reflection. To support formative assessment and deep learning, it is vital to stimulate dialogues where students articulate and share their understanding. This happens when the students use the alternatives as points of departure for clarification and argumentation. Open-ended questions enabled more exploratory talk and should thus be used to create moments of contingency.

A common claim in the research literature on the use of student response systems is that aggregated responses on MC-questions provide feedback to the lecturer about the students' knowledge and understanding (Chien *et al.*, 2016; Dawson *et al.*, 2018; Hunsu *et al.*, 2016; Liu *et al.*, 2017). Our data analysis revealed considerable uncertainty in a majority of the discussions, even though the students voted for the right answer (in questions with only one correct answer) or a reasonable answer (in questions allowing multiple correct answers). Regarding formative assessment, this is important for several reasons: (1) Exploring uncertainty allows students to be aware and reflect on their own understanding. (2) Our analysis demonstrates that conclusions based on clicker responses might be fragile and that the response students given by choosing an answer provides limited information about the students' understanding of subject concepts. This raises questions about the validity of the inferences that can be drawn from the aggregated responses to clicker questions. Wood *et al.* (2014), James and Willoughby (2011) and Knight *et al.* (2013) have expressed similar concerns. Dall'Alba & Bengtson (2019) argue that underneath what is visible or apparent "we might become aware [of] disconnected thoughts, broken arguments and doubt" (Dall'Alba & Bengtson, 2019, p. 1486). When this uncertainty is brought to the scene, there is a potential to open "moments of contingency" and allow the lecturer to enter into dialogue with the students' thinking. We argue that when the students' responses are received through a system without options for the lecturer to unpack the students' reasoning behind their aggregated response, there is a risk to neglect valuable opportunities for learning. In any activities created for the purpose of formative assessment, engaging in dialogue with students allows more sensitivity towards students' ideas, which helps in drawing inferences and strengthens the possibilities for the lecturers to follow up student responses in a formative way. To embrace activities that allow both aggregated and qualitative answers, or to include different ways of displaying knowledge would allow complexities to emerge, and this would provide richer insights into the students' understanding and thereby allow other inferences to be drawn.

Limitations

This study was conducted in an authentic setting, and thus, has several limitations regarding design, data collection and analysis. Video recordings of the lectures might have provided greater insights into the quality of the discussions. For example, body language would have been captured, which could provide valuable information. Also, we did not know what each individual student voted. If this information was connected to the discussions, we would be able to assess the quality of the discussions and how students responded. Furthermore, the students recorded their own discussions; therefore, it is possible that they might have chosen not to record some of their discussions if they were unsure of the answers, which could have influenced the results. Some groups decided to record their discussions, while other groups chose not to record their discussions. The use of audio recorders may also have affected the quality of the discussions (eg, students may have tried their best to engage in a productive discussion or were afraid to talk if they were unsure).

Implications for practice

To create, recognise and capitalise on those moments as an integrated part of learning activities helps teachers adjust their teaching to the needs of their students and help students to take

decisions on their learning process. When using student response system as a part of a formative feedback practice, the key questions to consider are to what extent and how the activities allow students to share thinking and what kind of thinking is being shared. To secure the quality of the inferences that can be drawn from the activities, it could be a strength, to include questions on confidence and/or design an environment, where uncertainty and questions might surface and to take time to clarify and elaborate on these questions.

To ensure that valid inferences can be drawn from activities as a basis for feedback, the questions used and threats to the validity of possible inferences should be critically examined.

Implications for research

Future studies should compare clicker-supported discussions and discussions without technology within the same overall teaching design. A following up study could design typologies of questions to investigate more rigorously the influence of question type on student discourse. Further studies should explore the role of the lecturer, and how lecturers use quantitative or qualitative information received through these systems to adjust their teaching to the needs of the students.

Acknowledgements

First, we would like to thank the students who participated in the recorded peer-discussions. We would also like to thank insightful comments from five anonymous reviewers. Your comments surly contributed to the quality of the paper.

Statements on open data, ethics, and conflict of interest

The study was conducted under the Norwegian Centre For Research Data framework for research ethics.

There is no conflict of interest in the work reported.

The recordings of peer-discussions in the study are not currently accessible due to agreement with the participants.

Funding

This research was funded by University of Bergen. It is located in the Research Group Digital Learning Communities.

References

- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, 13(1), 1–14.
- Barth-Cohen, L. A., Smith, M. K., Capps, D. K., Lewin, J. D., Shemwell, J. T., & Stetzer, M. R. (2016). What are middle school students talking about during clicker questions? Characterizing small-group conversations mediated by classroom response systems. *Journal of Science Education and Technology*, 25(1), 50–61.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5–31.
- Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400–413.

- Chien, Y. T., Chang, Y. H., & Chang, C. Y. (2016). Do we click in the right way? A meta-analytic review of clicker-integrated instruction. *Educational Research Review*, 17, 1–18.
- Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, 24(2), 205–249.
- Cooper, R., Chenail, R. J., & Fleming, S. (2012). A grounded theory of inductive qualitative research education: Results of a meta-data-analysis. *The Qualitative Report*, 17(52), 1–26.
- Cooper, R., Fleischer, A., & Cotton, F. A. (2012). Building connections: An interpretative phenomenological analysis of qualitative research students' learning experiences. *The Qualitative Report*, 17(17), 1–16.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative: and qualitative research*. Prentice Hall Upper Saddle River, NJ: Pearson Education.
- Dall'Alba, G., & Bengtson, S. (2019). Re-imagining active learning: Delving into darkness. *Educational Philosophy and Theory*, 51(14), 1–13.
- Dawson, P., Henderson, M., Ryan, T., Mahoney, P., Boud, D., Phillips, M., & Molloy, E. (2018). Technology and feedback design. In M. Spector, B. Lockee, & M. Childress (Eds.), *Learning, design, and technology* (pp. 1–45). Cham, Switzerland: Springer.
- De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46(1), 6–28.
- Egelandsdal, K., Ludvigsen, K., & Ness, I. J. (2019). Clicker interventions in large lectures in higher education. In M. Spector, B. Lockee, & M. Childress (Eds.), *Learning, design, and technology*. Cham: Springer.
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70–120.
- Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: findings from a four-year intervention study. *Instructional Science*, 44(3), 267–291.
- Gibbs, G. (2013). *How to conduct a qualitative interview*. Retrieved from <https://www.youtube.com/watch?v=9t-hYjAKww>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81–112.
- Havnes, A., Christiansen, B., Bjørk, I. T., & Hessevaagbakke, E. (2016). Peer learning in higher education: Patterns of talk and interaction in skills centre simulation. *Learning, Culture and Social Interaction*, 8, 75–87.
- Hennessy, S., Rojas-Drummond, S., Higham, R., Torreblanca, O., Barrera, M. J., Marquez, A. M., & Ríos, R. M. (2016). Developing an analytic coding scheme for classroom dialogue across educational contexts. *Learning, Culture and Social Interaction*, 9, 16–44.
- Hunsu, N. J., Adesope, O., & Bayly, D. J. (2016). A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect. *Computers & Education*, 94, 102–119.
- James, M. C., & Willoughby, S. (2011). Listening to student conversations during clicker questions: What you have not heard might surprise you!. *American Journal of Physics*, 79(1), 123–132.
- Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53(3), 819–827.
- Knight, J. K., Wise, S. B., & Southard, K. M. (2013). Understanding clicker discussions: Student reasoning and the impact of instructional cues. *CBE—Life Sciences Education*, 12(4), 645–654.
- Krumsvik, R., & Ludvigsen, K. (2012). Formative e-assessment and learning outcome in higher education. *Nordic Journal of Digital Literacy*, 1(6), 36–54.
- Littleton, K., & Mercer, N. (2013). *Interthinking: Putting talk to work*. London, UK: Routledge.
- Liu, C., Chen, S., Chi, C., Chien, K. P., Liu, Y., & Chou, T. L. (2017). The effects of clickers with different teaching strategies. *Journal of Educational Computing Research*, 55(5), 603–628.
- Ludvigsen, K., Krumsvik, R., & Furnes, B. (2015). Creating formative feedback spaces in large lectures. *Computers and Education*, 88, 48–63.
- Ludvigsen, K., Ness, I. J., & Timmis, S. (2019). Writing on the wall: How the use of technology can open dialogical spaces in lectures. *Thinking Skills and Creativity*, 34. <https://doi.org/10.1016/j.tsc.2019.02.007>
- McDonough, K., & Foote, J. A. (2015). The impact of individual and shared clicker use on students' collaborative learning. *Computers and Education*, 86, 236–249.

- Mercer, N. (2004). Sociocultural discourse analysis. *Journal of Applied Linguistics*, 1(2), 137–168.
- Mercer, N., Littleton, K., & Wegerif, R. (2004). Methods for studying the processes of interaction and collaborative activity in computer-based educational activities. *Technology, Pedagogy and Education*, 13(2), 193–209.
- Nielsen, K. L. (2012). *Student response systems in science and engineering education* (doctoral dissertation). Norwegian University of Science and Technology.
- Nielsen, K. L., Hansen, G., & Stav, J. B. (2016). How the initial thinking period affects student argumentation during peer instruction: Students' experiences versus observations. *Studies in Higher Education*, 41(1), 124–138.
- Nielsen, K. L., Hansen-Nygård, G., & Stav, J. B. (2012). Investigating peer instruction: How the initial voting session affects students' experiences of group discussion. *ISRN Education*, 2012, 1–9.
- Sikes, P. (2006). On dodgy ground? Problematics and ethics in educational research. *International Journal of Research & Method in Education*, 29(1), 105–117.
- Wegerif, R. (2013). *Dialogic: Education for the Internet age*. Cornwall: Routledge.
- Wood, A. K., Galloway, R. K., Hardy, J., & Sinclair, C. M. (2014). Analyzing learning during peer instruction dialogues: A resource activation framework. *Physical Review Special Topics—Physics Education Research*, 10(2), 1–15.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.