Teaching and Teacher Education 110 (2022) 103593

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Research paper

Teacher feedback on procedural skills, conceptual understanding, and mathematical practices: A video study in lower secondary mathematics classrooms

Roar Bakken Stovner*, Kirsti Klette

Department of Teacher Education and School Research, University of Oslo, Norway

HIGHLIGHTS

- We analyzed 47 mathematics teachers' feedback in 172 lessons in Norwegian lower secondary school.
- Both the feedback quality and its focus (procedures, concepts, or mathematical practices) was charted.
- Teachers provided clear and specific feedback that largely focused on procedures.
- Conceptual feedback was also common, but feedback on mathematical practices was rare.

ARTICLE INFO

Article history:
Received 3 June 2021
Received in revised form
20 September 2021
Accepted 18 November 2021
Available online 4 December 2021

Keywords:
Mathematics education
Feedback
Teaching practices
Classroom observation
Lower secondary school

ABSTRACT

Feedback is a prevalent teaching practice in mathematics classrooms, but few studies have documented how mathematics teachers enact feedback in classrooms. We investigated how 47 teachers provided feedback in 172 mathematics lessons in Norwegian lower secondary schools. We analyzed the quality of feedback, the quantity of feedback, and whether the feedback addressed students' procedural skills, conceptual understanding, or engagement in mathematical practices. Teachers spent large amounts of time providing concrete and specific feedback, most of it addressing procedural skills while conceptual feedback was less common. The study highlights details of feedback relevant for both pre- and inservice mathematics teacher training.

Hundreds of studies have investigated feedback and its effects on student learning, and meta-analyses and reviews indicate that teachers' use of feedback has large impacts on student learning (Hattie & Timperley, 2007; Kluger & Denisi, 1996; Shute, 2008). Feedback is a prevalent teaching practice, as eighty percent of teachers state that they frequently or always "observe students when working on tasks and provide immediate feedback" (OECD, 2019, p. 61). Despite the abundance of studies on feedback and its importance in regular classroom teaching, few studies have investigated feedback provided in classrooms (Ruiz-Primo & Li, 2013; Shavelson, 2003; Wiliam, 2018). This is particularly the case for the mathematics education: research on feedback in mathematics has mostly studied evaluative written feedback (correct/incorrect) on assessments or homework (M. Li et al., 2011) and there are almost no studies of how feedback is provided in regular mathematics lessons (Ruiz-Primo & Li, 2013). The field needs studies of the feedback from in situ and authentic mathematics classrooms because this might give important information on the kinds of mathematical learning opportunities teacher feedback grants students. This can aid mathematics teachers and teacher educators by showing the possible feedback practices available, including best-practice examples, and provide guidance about where to focus teacher development efforts.

The few observational studies of classroom feedback in mathematics lack a mathematical specification, either because they study generic aspects of feedback such as frequency and praise (N. Li et al., 2016; Voerman et al., 2012) or because they conflate mathematics classrooms with other subjects (Gamlem & Munthe,

* Corresponding author. Department of Teacher Education and School Research, University of Oslo, Norway.
E-mail addresses: robast@oslomet.no (R.B. Stovner), kirsti.klette@ils.uio.no (K. Klette).

https://doi.org/10.1016/j.tate.2021.103593

2014; Voerman et al., 2012). Such studies provide information on possibly very important aspects of feedback, but no information on how teachers' feedback practices support students' development of the mathematical competencies that current mathematics curricula specify. Mathematics educators and curricula stress that mathematical competency includes procedural skills, conceptual understanding, and the ability to engage in mathematical practices such as problem-solving, proving, and mathematical modeling (Ball, 2003; Burkhardt & Schoenfeld, 2019; Hiebert & Grouws, 2007; Ministry of Education and Research, 2015 [UDIR]; Moschkovich, 2013; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010 [NGACBP]). An important aspect of teacher feedback in mathematics is therefore how it may support the development of these competencies.

In the present study, we investigated teachers' oral feedback to students in 172 video-recorded lessons in 47 mathematics classrooms in lower secondary schools in Norway. We aimed at charting generic and subject-specific indicators of feedback, how much time teachers spend providing feedback, and in what activity formats (individual work/group/whole-class), as well as the mathematical competencies addressed in the feedback. First, to chart generic quality indicators of feedback, we used the standardized observation instrument Protocol for Language Arts Teaching Observation (PLATO; Grossman et al., 2013). Second, to describe how teachers focus on the different mathematical competencies in feedback, we conducted a qualitative analysis of feedback instances in the video-recorded lessons. This also afforded us to chart the amount of time teachers spend in feedback conversations with students and in which activity formats. The key questions investigated were:

(1) What is the quality of feedback as measured by the subject-generic dimensions?
(2) How much time do teachers spend providing oral feedback and in what activity formats?
(3) To what extent and how do teachers in mathematics classrooms provide feedback focusing on procedural skills, conceptual understanding, and mathematical practices?

1. Norwegian context

Norway has a comprehensive and non-tracked model of primary and lower secondary education for the first 10 years of schooling. Mathematics teaching in lower secondary classrooms has traditionally been characterized by teacher-led whole-class instruction and individual work (Bergem, 2014; Olsen, 2013), but more recent studies have seen larger instructional repertoires including group work (Klette et al., 2017). Feedback has been strongly emphasized in national policy: the national guidelines for teacher education state that prospective teachers must learn to give productive feedback (Gamlem, 2015; Munthe & Melting, 2016); between 2010 and 2018, a nation-wide professional development program was conducted emphasizing effective feedback in the Assessment for Learning (AfL) framework (UDIR, 2018); and students have a legal right to receive individual feedback (Education Act, 2006, § 3–10). Since 2006, the national curriculum has stressed the teacher's role in familiarizing students with mathematical practices, such as problem solving, mathematical discourse, justification and generalization (UDIR, 2006). The new national curriculum (UDIR, 2015) increases this emphasis by including many of these mathematical practices as “core elements.”

2. Conceptual framework

2.1. The concept of feedback

Feedback is information provided to a student about the student's performance or understanding (Black & Wiliam, 1998). A key feature of feedback is that it is evaluative, either through decomposing qualities of student work or suggesting improvements and ways forward (Sadler, 1989). This understanding is closely related to AfL (see Black & Wiliam, 1998), a set of teaching practices that have strongly impacted educational policy internationally (Birenbaum et al., 2015) and in Norway (UDIR, 2018). Following Black and Wiliam (1998), teachers teach according to AfL if they continually and informally assess student understanding and use this information to adapt their instruction or provide feedback.

In this study, we focus on the oral feedback provided by the teacher in mathematics lessons. Oral feedback can take the form of short praise, encouragement, and corrections, or it may be part of a longer conversation between teacher and student. Questioning plays a key role in providing feedback, because it is often through questions that the teacher learns about student understandings (Small & Lin, 2018). The distinction between teacher questioning and feedback may be blurry. To demarcate which utterances are considered feedback in the present study, we require that the utterances contain a) an evaluation or b) a suggestion for how to move forward. This will generally not include questions, although some questions may fulfill the requirements.

2.2. Qualities of feedback

Although aspects of quality feedback varies with the subject taught (Smith & Lipnevich, 2018), researchers have identified several aspects of high quality feedback that are generic (Hattie & Timperley, 2007; Shute, 2008). Feedback should focus on the student's work on the task at hand, and not on the student's skill level in comparison to others or to a grading scheme (Hattie & Timperley, 2007). Feedback is more effective when it not only provides information about answer correctness, but also elaborates on qualities of student work or how to improve (Shute, 2008). These elaborations should be specific and clear (Shute, 2008), but not so specific as to remove student agency from the learning process and making further work too easy (Hattie & Gan, 2011). Rather, they should address the overarching goal of the task and the underlying skills required to perform proficiently (Wiliam, 2018).

2.3. Feedback and activity formats

Feedback may be provided in all activity formats, like whole-class instruction, group work, or individual work. In whole-class instruction, the initiation, response, and evaluation/feedback (IRE/F) pattern has been used extensively (Inagaki et al., 1999; Mehan, 1979). The IRE/F pattern is an interaction pattern wherein the teacher initiates a question (often termed funneling questions (Wood, 1998)), a student replies, and the teacher gives evaluative feedback. Teacher replies are often short, even when students reply erroneously (Santagata, 2005). Other ways of providing feedback in whole-class instruction exist, for example by discussing students' solution methods (Stein et al., 2008), but they are less common. During group work, the teacher may provide feedback on the students' collaboration or on the quality of the group's work (Brodie, 2000). Individual feedback may have the biggest potential to be tailored to each student's individual needs (Krammer, 2009).

2.4. Quantity of feedback

There is no optimal quantity of feedback teachers should strive for. On the one hand, teachers must provide enough feedback that student work is productive. For example, teachers can provide corrective feedback to ensure that students are practicing procedures correctly. On the other hand, feedback may remove important learning opportunities, for example if students are working on challenging tasks and the feedback reveals too much of the solution (Stein et al., 1996). Feedback also comes with “opportunity costs”, namely that the time the teacher spends providing feedback might be better spent doing something else (Hays et al., 2010).

Studies from classrooms have found that teachers spend little time providing feedback (Krammer, 2009; N. Li et al., 2016; Voerman et al., 2012). Voerman et al. (2012) found that across 72 lessons (25 in mathematics), teachers provided an average of seven feedback interventions every 10 min of instruction. The authors “find these outcomes to be alarming, because feedback […] is one of the most important tools available to positively influence their students' learning” (Voerman et al., 2012, p. 1113). However, we could argue that seven feedback interventions in 10 min is quite a lot of feedback depending on the teaching situation. In the current study, we report on quantity of feedback teachers provide, but do not judge whether too much or too little feedback was provided. Rather, we view the quantity of feedback as one of several parameters that are relevant for understanding how teachers use feedback in the classroom.

2.5. The mathematical focus of the feedback

We stated earlier that feedback should elaborate on qualities of student work and what they need to do to improve. A key aspect of this study was to investigate what the mathematical focus of the feedback was. Especially, we were interested in feedback focusing on a) procedural skills, b) conceptual understanding, and c) mathematical practices such as problem-solving, proving, and mathematical modeling. All three areas are underscored as critical in current mathematical curriculum documents in the US and Norway (NGACBP, 2010; UDIR, 2013).

2.5.1. procedural feedback

Procedural feedback focuses on procedural skills by helping students arrive at the correct answer by applying the correct procedure or remembering facts and rules. This conceptualization stems from the classical definition of procedural knowledge, defined as knowledge of isolated facts or rules or rigid procedures (Hiebert & Lefevre, 1986). Examples of such feedback are: pointing out a computational error, showing the next step in a solution method (or even the entire procedure), reminding the student of a rule, and acknowledging whether a student's answer is correct without further elaboration. Verifying whether students' answers or methods are correct aids learning, especially for learners with low background knowledge (Fyfe et al., 2012). Procedural feedback presumably has the greatest effect on procedural outcomes, but some studies have shown improvement in conceptual outcomes as well (Fyfe & Brown, 2017). However, these findings are from experiments conducted individually outside of classrooms, and further research is needed to conclude that the findings are transferable to classroom settings (Booth et al., 2017).

2.5.2. Conceptual feedback

Conceptual feedback is feedback that addresses mathematical concepts, the relations between them, or students' understanding of them. An example is a teacher that draws a graphic representation of measurement division to a student that has made a division error in order to explain why it was an error. Here, only showing the rules of long division would be procedural feedback. Conceptual feedback is a way of explicitly addressing conceptual understanding, which is a critical aid for student learning (Hiebert & Grouws, 2007). As most studies of feedback concern feedback about answer correctness or procedural remediation of errors, not many studies of conceptual feedback exist, and their findings must be further confirmed and elaborated. Although procedural feedback has strongest effect on student learning on similar tasks, conceptual feedback might be more effective on far-transfer tasks where students must apply their knowledge in new contexts (Alibali, 1999). However, providing conceptual feedback is hard: Student teachers instructed in responding to students' concepts and ideas continue to provide feedback through praise or correcting the procedure (Son & Crespo, 2009) even when they acknowledge that the student lacks conceptual understanding (Bartell et al., 2013). Also, it is more difficult to identify student errors and respond to them under time constraints (Pankow et al., 2018). Because teachers must provide feedback “on the spot” during classroom teaching, we expect to find few examples of conceptual feedback in the present study.

2.5.3. Feedback on mathematical practices

Feedback on mathematical practices focuses on how students engage in mathematical practices (Moschkovich, 2015). The term *mathematical practices* is inspired by the practices mathematicians engage in while they work, such as problem-solving, justifying mathematical claims, or engaging in mathematical discourses (Moschkovich, 2013). A common explanation is that “[there are] two main (and deeply intertwined) aspects when doing mathematics: mathematical content and mathematical practices” (Schoenfeld, 2014, p. 500). Mathematics curricula have increasingly focused on mathematical practices as an important goal for student proficiency (UDIR, 2015; NGACBP, 2010). A problem is that mathematical practices may remain hidden to students if they are not addressed explicitly (Selling, 2016), and that they often devolve into mere prescriptions if a teacher explicitly prompts the use of them (Bieda, 2010). A proposed solution is to link students' work with a mathematical practice while they are engaged in them or immediately after (Selling, 2016), for example through feedback. An example of feedback on mathematical practices is a teacher that gives feedback on a student's use of problem-solving heuristics, for example by evaluating qualities of how the student problem-solved.

3. Methods

3.1. Data and design

This study was part of a larger classroom video study conducted to analyze teaching practices in Norwegian language arts and mathematics (Klette et al., 2017). During the 2014–2015 school year, the research team video-recorded 172 lessons in grade 8 mathematics classrooms in Norway. The students were 13–14 years old and in the first year of lower secondary school. Forty-seven formally trained mathematics teachers, each teaching in a different school, were video recorded. Three to four consecutive lessons were recorded in each classroom. The schools were carefully sampled to reflect variations in important school variables in Norway, such as rural/urban location, socioeconomic status, immigrant student population, and achievement level on national tests in reading and numeracy. The videotaped lessons followed the classes' normal schedules, and the teachers were asked to teach "as usual". Two discreet cameras and two microphones captured the whole-class discussions, teachers' talk, and teacher–student interactions (see Klette et al., 2017 for further descriptions of the camera setup). In the 47 recorded classrooms, there were 991 students: 456 boys, 500 girls, and 35 of undisclosed gender. The class sizes varied between 18 and 28 students and averaged 21. The students scored close to the national average on the national numeracy tests when they entered grade 8 and made slightly higher gains during the year.

3.2. Analyses

We performed the analyses in two steps. First, we used a standardized observation instrument to answer the research question about the feedback quality in the observed lessons. Second, we conducted a fine-grained qualitative analysis of feedback instances to answer the research questions on (a) the time spent providing feedback, (b) the activity formats in which it was provided, and (c) how teachers provided feedback with different mathematical foci (i.e., procedures, concepts, and mathematical practices).

3.3. First step of the analysis: the PLATO observation instrument

We used a standardized observation instrument (Bell et al., 2018) to obtain an overview of the generic quality of feedback across all lessons. Such instruments are suitable for transparently and consistently analyzing large observational data sets (Klette & Blikstad-Balas, 2017). We used the Protocol for Language Arts Teaching Observation (PLATO) instrument (Grossman et al., 2013) for two reasons. First, its conceptualization of feedback aligns with AfL, which is how teachers in Norway are expected to provide feedback. Second, it captures the generic quality indicators of feedback described previously; namely that high-quality feedback elaborates on qualities of student work or provides substantive suggestions for how to continue. PLATO was initially designed to capture elements of English language arts instruction in middle school but has since been used at the lower secondary level to rate mathematics instruction (Cohen, 2018; Luoto et al., in review; Tengberg et al., 2021). We did not use a mathematics-specific observation instrument, as these do not currently have dimensions that measure feedback (Praetorius & Charalambous, 2018).

The teaching practices that are covered by PLATO correspond with notions of quality mathematics teaching in Norway (Klette et al., 2017) and in the US (Cohen, 2018). Specifically, PLATO values teaching that stresses "rigorous content and intellectually challenging tasks, the centrality of classroom discourse in developing sophisticated understanding of content and disciplinary skills, and the critical role of teachers in providing instructional scaffolding for students to help them succeed" (Grossman et al., 2014, p. 295). Feedback within this view of teaching should not just help students by correcting their work or providing steps for how to continue but develop sophisticated understanding of the mathematics and discipline-specific ways of working.

When rating instruction with PLATO, lessons are subdivided into 15-min segments. The segments are rated on 12 elements of quality teaching, like feedback, time management, and quality of explanations. Each element is rated on an ordinal scale from 1 to 4. Only the feedback element was used in the present analysis.

A summary of PLATO's feedback element is provided in Table 1. At the high end, for a segment rating of 3 or 4, the feedback must (i) be clear, (ii) be specific to the student's work or ideas, and (iii) address "substantive elements" in the task or the student's contribution. "Substantive elements" refers to elements that go beyond a specific task. It refers to feedback that addresses students' understanding, underlying skills, or ways of working and is described in contrast to feedback providing prescriptive procedures or rules required for completing a task. For a rating of 3, although the segment may include many vague or procedural feedback instances, it must include at least two substantive feedback instances. For a rating of 4, the segment must include several substantive feedback instances and few procedural ones. Additionally, the rater must be able to reasonably infer that the feedback helps students. There is thus both a qualitative and quantitative difference between ratings 3 and 4. At the low end, for a rating of 1, the teacher does not provide feedback, and for a rating of 2, the feedback is either vague or perfunctory (e.g., "Good job"; "That's correct") or pertains only to correctly performing procedures and completing the task at hand.

Raters must be trained and certified in using PLATO, which requires rating a master-rated collection of 15-min segments with 80% exact agreement. A team of five certified PLATO raters conducted the ratings. All lessons were subdivided in 15-min segments and rated by one rater. The inter-rater reliabilities were estimated by double-coding 16 lessons sampled to include all PLATO rating levels and be taught by different teachers. This constituted 15% of the material since the lessons were slightly longer than average. The feedback element had a 72% agreement between the raters (Klette et al., 2017). Upon disagreement, the ratings from the first coding pass were used. We had some concerns about our level of reliability. Reported guidelines for interpreting percentages of exact agreement, propose that 80% is sufficient (Hartmann et al., 2004), but that for more complex and high-inference instruments (such as PLATO), an exact agreement of 70% between raters may suffice (Hill et al., 2012; Ho & Kane, 2013). The reliability level is also briefly discussed in the Results section.

3.3.1. Second step of the analysis: analyzing feedback instances

In the second step, we conducted a fine-grained qualitative analysis using *feedback instances* as the unit of analysis. We define a feedback instance as a sequence of utterances between the teacher and one or more students that was about their work or ideas and contained a teacher utterance that included an evaluation or a suggestion for improvement. A feedback instance started when a teacher and student began talking about the student's work or ideas and ended when the topic changed, which could be the start of a new feedback instance, or the discussion ended. Feedback instances included short one-line teacher utterances ("Good work.") as well as conversations where the teacher and the students talked and asked questions about the student's work at length. Commonly, the teacher added additional instruction—sometimes quite lengthy—to help the student. This was included in the feedback instance only if the teacher tied the instruction very closely to the student's work. Feedback instances could occur during individual or group work or whole-class instruction.

We identified feedback instances using the video-analysis software Interact (Mangold, 2019), which facilitated recording each instance's start and end times and appending labels and comments. Each instance was labeled with the appropriate classroom and 15-min segment, that segment's PLATO feedback rating (1, 2, 3, or 4), the activity format the feedback was provided in (whole-class/group/individual), and its mathematical foci. We also wrote short memos of each feedback instance, which contained a summary of what was said. An example of the information recorded is provided in Table 2. In analyzing the time spent, we calculated the total time a teacher spent in feedback utterances during each 15-min segment.

We then conducted a qualitative analysis of the feedback instances to gain in-depth information about the features and characteristics of how feedback with different mathematical foci was

Table 1
PLATO feedback dimension.

1 – No feedback	2 – Perfunctory or entirely procedural feedback	3 – Specific feedback with at least some substantive elements	4 – Consistently specific and substantive feedback
Teacher does not provide feedback. Erroneous or confusing feedback	Feedback is perfunctory (“well done,” “nice work,” “correct,” “No, wrong.”) Prescriptive suggestions for how to continue (“These should be on a common denominator.”) Showing correct procedure to complete student's task.	Feedback specific to students' work or contribution. At least two feedback instances contain substantive elements, such as explaining why the student's procedure is wrong, explaining a concept the student has misunderstood, or providing flexible suggestions for how to continue.	Feedback is consistently both specific to student's work or contribution and has substantive elements. It is reasonable to infer that the feedback helps students understand.

Table 2
Example information recorded for each feedback instance.

Start–End time (min:sec)	13:07–14:20
Classroom	37
Segment Feedback Rating	3
Activity Format	Individual work
Mathematical focus	Procedural and conceptual
Memo	Teacher tells the student he needs not put fractions on common denominator when multiplying and demonstrates how. Student says he was unsure if he needed common denominator. Teacher reminds student about the area representation for multiplying fractions

provided across the different PLATO rating levels. We coded features of how the feedback was provided to obtain themes (Braun & Clarke, 2006) of how feedback with the different mathematical foci was provided. Themes were gradually reworked by merging or refining existing themes to better characterize how feedback of different types was provided. Finally, we examined the coding by looking for inconsistencies, checking that the themes cohered across the material, and verifying our interpretations by double-checking the coding with the videos. We then translated and transcribed the examples that best typified the feedback categories.

In the second step of the analysis, we analyzed a subset of the 15-min segments ($n = 54$). This subset's size had to be big enough to represent typical patterns in the data set, but small enough to enable a careful in-depth analysis. First, we chose not to analyze segments rated 1 because these contained no feedback. Second, we randomly sampled segments rated 2, 3, and 4 until saturation (Saunders et al., 2018). Saturation was achieved when analyzing an additional 3–4 segments only yielded feedback instances that fit into the existing themes without developing new themes or changing the existing themes' characteristics. The final subset included 54 segments, in which we identified 603 feedback instances. We double-coded 79 feedback instances (13%) that covered all teachers and coding levels. Agreement was high for all codes, including activity format (96%), procedural feedback (90%), conceptual feedback (90%) and feedback on mathematical practices (99%).

3.3.2. Limitations

The study has several limitations in providing valid and reliable observational data on how feedback is provided in mathematics classrooms. We knew that the students in our sample had slightly higher gains on national numeracy tests than the national sample. Our sample might therefore have slightly more high-quality feedback practices than a random sample. Also, the camera's presence may have affected teachers and students, such that the lessons were not typical of normal teaching practices. Empirical research on camera reactivity has found, however, that it fades after the first few minutes and the teaching quality is unaffected (Praetorius et al., 2017), suggesting that camera reactivity was a minor problem. Additionally, the study focused on the teachers' feedback. However, it has been argued that feedback must be seen in relation to other factors, such as the learning goals, the tasks (Small & Lin, 2018), and the individual students (Van der Kleij, 2019). We did not investigate such factors, which would have required a different study design.

Although PLATO was developed to code language arts instruction, it has also been used to code mathematics instruction (Cohen, 2018; Klette et al., 2021; Tengberg et al., 2021) and other subjects such as foreign language and science (Brevik, 2019; Mahan et al., 2021). Current debates on using generic versus subject specific observation instruments to analyze instructional quality provide arguments for both approaches (Berlin & Cohen, 2020; Hill & Grossman, 2013; Schoenfeld, 2018). Praetorius and Charalambous (2018) reviewed 11 generic and subject specific observation instruments for use on mathematics instruction, and concluded that there are benefits and pitfalls of using both approaches. We decided to use PLATO as a generic instrument in the present study and supplement them with a mathematics-specific analysis in step 2.

4. Results

4.1. Distribution of subject-generic quality of feedback

Across the 172 lessons, 520 15-min segments were rated with the PLATO instrument. Of these, 72 (13.9%) were rated 1 indicating no feedback, erroneous feedback, or confusing feedback. These segments showed teachers performing administrative tasks, such as discussing an upcoming school-trip, lecturing, or conducting short tests. In our data, no segment was rated 1 because the feedback was erroneous. There were 349 (67.1%) 15-min segments rated 2, indicating vague, perfunctory feedback or feedback containing prescriptive rules and procedures. In line with the PLATO instrument, the feedback at this level could include one instance with substantive elements, but if two instances occurred in the same segment, it was rated 3. Feedback rated 3 or 4 is considered “high end” according to the PLATO instrument, meaning that they provide specific and substantive suggestions for how to improve students' work or understanding. In total, 91 (17.5%) segments were rated 3; that is, they contained at least two substantive feedback instances that were specific to students' work or ideas. Only eight segments (1.5%) were rated 4, indicating teaching that consistently provided feedback that was both substantive and specific to the

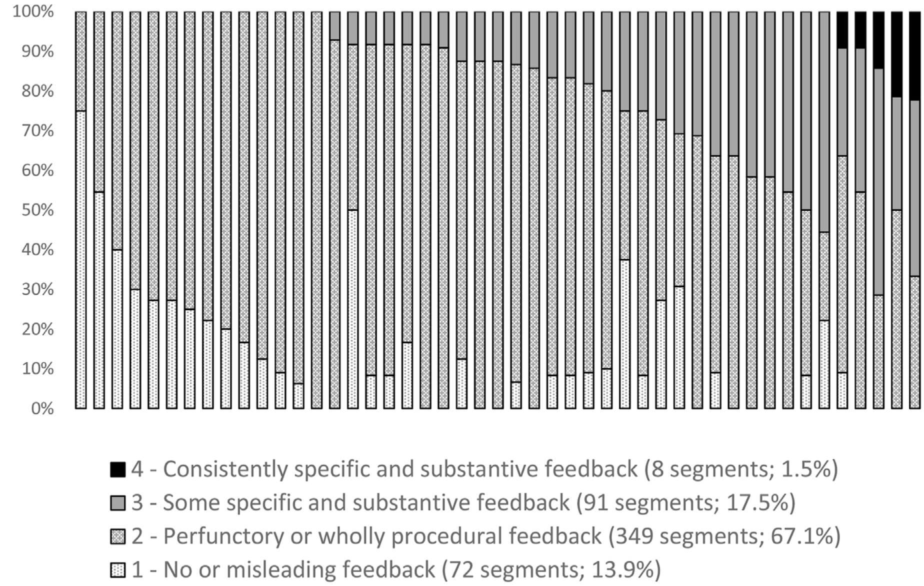

Fig. 1. PLATO feedback ratings for the 47 teachers. Each vertical bar represents one teacher's distribution of segment feedback ratings.

students' work, and it was reasonable to infer that the feedback helped the students.

These ratings were unequally distributed among the 47 teachers. As shown in Fig. 1, the teachers varied widely in the number of high-rated segments they taught. Fourteen teachers (30%) never received a score of 3 or 4; that is, they rarely or never provided substantive feedback. The remaining 33 teachers (70%) received at least one rating of 3 or 4 because they provided substantive feedback during the segments. Five teachers (11%) received at least one rating of 4, meaning that feedback was consistently substantive.

No teachers received high ratings for all their segments, which was neither expected nor deemed beneficial. We regard the PLATO results as a descriptive overview of the teachers' feedback practices, and higher ratings are not always better because many learning situations may benefit from not including feedback (Hays et al., 2010) or because procedural feedback may be sufficient for the learning at stake. Next, we describe how much time teachers spent in feedback instances, in what activity formats feedback was provided, and how teachers provided feedback on procedures, concepts, and mathematical practices.

4.2. Looking across feedback instances: activity format and time spent

We conducted the second step on a subsample based on the PLATO ratings as described in the Methods section. The final subsample in step 2 of the analysis contained 54 segments and 29 teachers: 29 segments rated 2, 17 segments rated 3, and all 8 segments rated 4. We identified a total of 603 feedback instances. A feedback instance was a sequence of utterances that was about student work or ideas that included an evaluation or a suggestion for improvement. An overview of how they were distributed according to PLATO feedback rating and activity format (i.e., individual work, group work, or whole-class instruction) is shown in Table 3.

Fewer feedback instances were observed during group work than in individual work and whole-class instruction. We should clarify that during individual work, the students were often seated in pairs or in groups of two, three or four, as is common in Norwegian classrooms. Students may talk about mathematics and ask one another for help, but they are supposed to solve the tasks individually. We labeled the feedback instance as occurring during group work only when it was clear that the assigned tasks were meant to be solved in groups.

Table 3
Number of feedback instances across activity formats and PLATO scores.

PLATO score	Group work	Individual work	Whole-class	Total
2	33	130	140	303
3	12	126	81	219
4	21	32	28	81
Total	**66**	**288**	**249**	**603**

The teachers spent a large amount of time providing feedback to students, see Table 4. Around half of the time in the 15-min segments analyzed in step 2 was spent in feedback instances. The average time spent providing feedback did not differ greatly in the segments rated 2, 3, and 4, which indicates that segments have different ratings because of the quality—not quantity—of feedback.

The time spent providing feedback occurred mainly during individual work or group work. Here, the teachers provided feedback almost constantly. The only time during individual and group work that teachers were not in feedback instances was when moving between desks or when students asked non-mathematical questions, such as about homework deadlines. A significant amount of time was spent providing feedback during whole-class instruction as well, as teacher lectures without student participation were rare. Rather, teachers asked questions that students answered. Students' answers were usually short, whereas the teachers' feedback was sometimes long. Thus, a considerable amount of time was spent in feedback instances in all activity formats. We now move to the

Table 4
Average time teachers spent providing feedback in each 15-minute segment.

Segment Feedback score	Time
2	06m56s
3	08m02s
4	08m15s

mathematical content in the feedback instances.

4.3. Characteristics of procedural feedback, conceptual feedback, and feedback on mathematical practices

Fig. 2 shows the proportion of feedback instances in analysis step 2 that addressed procedures, concepts, and mathematical practices for each PLATO rating level. It shows that higher rated segments have higher proportion of conceptual feedback and feedback on mathematical practices. This is an indication that segments with mostly procedural feedback was rated 2 and segments with substantive feedback was rated higher, despite our concerns about the inter-rater reliability of the PLATO coding. The proportion of procedural feedback is approximately the same at all rating levels. The reason is that when teachers gave conceptual feedback, they almost always addressed procedures as well. Qualitative descriptions of how this was done is given in the next sections.

Only feedback instances with a mathematical focus are shown in Fig. 2, resulting in the columns for rating levels 2 and 3 adding up to less than 100%. The remaining feedback instances contain mostly vague praise ("Good work") and some feedback on effort ("It was hard for me to get you to start working today, class. This will need to improve tomorrow."). This means that most feedback was focused on mathematics, in contrast to earlier research showing that feedback mostly functioned as emotional support (Gamlem & Munthe, 2014).

In the following, we present characteristics of feedback instances that focused on procedures, concepts, and mathematical practices. The presentation is structured according to PLATO rating such that the results from step 2 can elucidate the distribution of PLATO scores in Fig. 1 above.

4.3.1. Feedback rated 2: procedural feedback

In the 29 segments rated 2, we identified 303 feedback instances (see Table 3). A PLATO feedback rating of 2 specifies that the feedback is vague, perfunctory, or only attends to procedural aspects of the task. In the segments rated 2, we distinguished between two categories of feedback instances: *showing students correct solution methods* and *brief evaluation of student answers.*

Procedural Feedback: Showing Students Correct Solution Methods. A prominent feature of the feedback in the segments rated 2 was how the teachers helped students by giving clear and unambiguous guidance in completing the task on which the students were working. We call this *showing students correct solution methods*. It was present in all segments in the subsample analyzed in step 2 and was especially prominent during individual work and group work. Some feedback instances were brief, such as when the teachers only stated what students should do without showing them the complete procedure (e.g., "remember to put on a common denominator here"). However, many feedback instances were longer, wherein the teachers carefully demonstrated how to conduct the procedure. In the following example, the students were simplifying fractions during individual work.

Example 1:

1 Student: Teacher? [signals the teacher to come over] I did the other ones, but I do not see how I can do 15 twenty-firsts?
2 Teacher: Factorize 15.
3 Student: Is that three?
4 Teacher: Just factorize 15, the same approach as the others. What do you get if you factorize 15?
5 Student: Five times … No, three times five?
6 Teacher: Yes. Write that, three times five. Okay, then you factorize 21.
7 Student: Three times seven?
8 Teacher: Yes. Then you have two equal factors to cross out. So, 15 twenty-firsts equals five-sevenths. Good!

In this example, the teacher provided all the steps but activated the student by prompting her to perform the simple calculations. Procedural feedback generally included such student prompting or funneling questions (Wood, 1998) about the next step. Note that the teacher did not include substantive elements, such as relating the explanation to general rules or concepts that could help the student understand the procedure.

We warn against inferring that the instruction in segments rated 2 was "traditional" or "teacher-led" because of the above descriptions. Indeed, the students were active in most segments, and peer-to-peer interaction was the norm regardless of whether the students solved routine textbook exercises or more demanding problems.

Procedural Feedback: Brief Evaluations of Student Answers. In segments rated 2, most feedback was very short, and most feedback instances lasted less than 15 s. In these instances, the teachers mainly acknowledged whether the students' answers were correct, perfunctorily re-voiced the students' answers, and often included brief praise, such as "That's correct, great!" We termed this kind of feedback *brief evaluations of student answers*. During whole-class discussions, brief evaluations of students' answers accounted for almost all feedback instances. For example, one teacher asked in the beginning of the lesson whether the students knew examples wherein statistics was used. Many students gave examples, and the teacher replied, "Great example," or "Well, I am not so sure about that one."

A feature of these *brief evaluations of student answers* was that the teachers did not dwell on the students' errors but proceeded quickly to the correct answer. The following example shows two feedback instances (lines 2–3 and 4–5) during a discussion in which the teacher explained how to solve the equation $\frac{x+1}{10} = 2$.

Example 2:

1 Teacher Any ideas?
2 Student 1 You can divide x by 2.
3 Teacher Well, that is not quite true. We must first find what x is.
4 Student 2 Multiply by ten and subtract one.
5 Teacher Yes! [writes on the board] What do I need to do now?

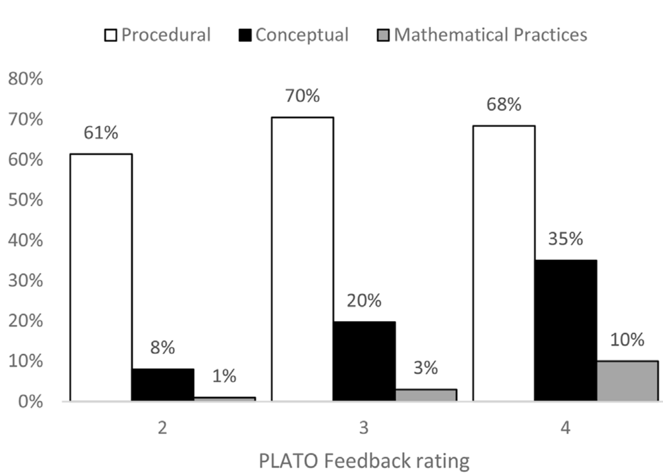

Fig. 2. Feedback instances in segments rated 2, 3, and 4. The percentages of the columns may add up to more than 100% because each feedback instance can be of several types.

The student's response in line 2 was apparently incorrect, which the teacher clearly stated in line 3. The teacher commented briefly on why it was incorrect ("We must first find what x is"), but it is unclear how this comment related to the student's answer, and the teacher quickly asked a new student to answer. This is typical of how the teachers responded to erroneous responses in procedural feedback instances: it only gives information on whether if a correct step has been executed.

4.3.2. Feedback rated 3 and 4: specific and substantive feedback

In the 17 segments rated 3 and the eight segments rated 4, we identified 219 and 81 feedback instances, respectively. In PLATO, to gain a rating of 4, the feedback must be consistently specific to student work and ideas and attend to substantive elements of the task. A rating of 3 is similar, except that the feedback need not be consistently substantive (but it must have at least two substantive feedback instances). That is, a rating of 3 indicates a mix of procedural and substantive feedback. The below categories thus apply strongly to segments rated 4, whereas segments rated 3 also include elements similar to segments rated 2. We distinguish between three feedback instance categories in segments rated 3 and 4, *showing the correct procedure with explicit attention to concepts*, *substantive evaluation of student answers,* and *feedback on mathematical practices* (e.g., problem-solving, proving.

Conceptual Feedback: Showing the Correct Procedure with Explicit Attention to Concepts. Showing students procedures, especially during individual and group work, was a characteristic of feedback instances labeled "conceptual". That is, they were labeled both "conceptual" and "procedural". The reason was that in addition to being procedural, the feedback addressed the procedure's conceptual underpinnings or linked the procedure to the students' preconceptions. That is, conceptual feedback was most often intertwined with procedural feedback. Also, to a much higher degree than in the segments rated 2, the teachers asked the students follow-up questions while providing feedback and modified their feedback depending on the answers. The next example clearly shows these features.

Example 3:

A student is working individually on a task that requires the insertion of $x = -2$, $y = 1.5$, and $z = -0.5$ into the expression $\frac{3x-5y}{z}$. The student has already done some work has written this in her notebook:

$$\frac{-6-7.5}{-0.5} = \frac{13.5}{-0.5}$$

The student puts her hand up to ask the teacher a question.

1 Student: Is positive and negative?
2 Teacher: Yes. [Talks to the student about what she has written.] We have to consider this, that you have gotten a positive answer. [Pointing at the numerators.]
3 Student: But negative and negative becomes positive?
4 Teacher: Not here. I think we need to look at the number line.
5 [Teacher draws a number line and shows the student how adding and subtracting positive numbers can be visualized as counting or jumping on the number line.]
6 Teacher: … [So] when you subtract a positive number you always go left on the number line. So, if you are at − 6 and go to the left, then you come to …
7 Student: Ah, I see! And when I get the answer, I remove the minus?
8 Teacher: Why should you remove the minus?
9 Student: Because minus and minus is plus?
10 Teacher: That only occurs when the signs are adjacent to each other.
11 Student: Ah, so this is minus and plus?
12 Teacher: No, no, no, try to think of what they are really asking. You have a number that is − 6. That may be, for instance, the temperature.
13 [Teacher proceeds by explaining that −6 − 7.5 models a 7.5°C drop in temperature from one day with −6°C to the next.]
14 Teacher: … and in such situations I have to go left on the number line without messing with the signs.
15 Student: A-ha!
16 Teacher: So, you see, the number line is the best way of understanding calculations with negative numbers. The rules that we have here are only mnemonics to help us after we have understood the values of the numbers. [The teacher further questions the student to ensure that she thinks correctly and answers correctly.]

In this example, the teacher chose to show the student how to complete the task while making conceptual connections to the number line and a temperature drop. Similar to the procedural feedback in segments rated 2, the teacher guided the student through a procedure. The utterances in lines 6 and 14 tell the student exactly what she should do to complete the task: go left on the number line from − 6, and divide by −0.5 to obtain a positive answer. The feature that distinguishes this feedback instance from the one in Example 1 is the explicit attention to concepts. The first concept is seen in lines 4 to 6 where the teacher linked the question to a mathematical representation, the number line. The second concept is after line 11 when the student indicated that she was still confused, which prompted the teacher to point to a real-life representation, a drop in temperature. The teacher thus provided two representations that a student may use when thinking about the subtraction of negative numbers.

Not all the teacher's utterances in this example make explicit references to concepts. In line 10, the teacher resorts to a rule to explain the subtraction, multiplication, and division of negative numbers. This illustrates that the feedback instances could contain several of the same features of segments rated 2.

Note that the task was a procedural exercise and that the student asked for a confirmation of a procedural rule. Even though both student and the task focused on rules and procedures, the teacher chose to respond by focusing on conceptual understanding. Thus, the task's nature did not determine the teacher's feedback: The teachers sometimes gave procedural feedback on non-routine tasks and conceptual feedback on procedural tasks.

Conceptual Feedback: Substantive Evaluation of Student Answers. The second most prevalent feature of feedback instances labeled "conceptual" was evaluations which focused on conceptual understanding after the students had finished a task. The teachers' evaluations were focused on why the students' solution methods were correct or incorrect and sometimes on why the solution methods were efficient or elegant. The teachers provided students with evaluative feedback during individual work (Example 4) and whole-class discussions.

Example 4:

This example occurred during individual work while the students were solving exercises in the textbook. We understand from the context that the student question in line 1 concerns whether 24 is a common denominator of 8 and 12.

1 Student Is 24 correct?
2 Teacher: Yes, it is. Are you not certain about that?
3 Student: I was just thinking that 8 times 12 is 96, so that must be the common denominator.

4 Teacher: Ah, but why are you suggesting 24?
5 Student: Because I can make 24 from 8 to 12, like here it is … [inaudible]
6 Teacher Yes, good! You have accomplished the task of making the denominators equal. You do not need to multiply them all the time. Here, you have made them equal, and it is smaller, so that seems to me to be the better option. It will be easier to calculate with later.

In line 6, after listening to what the student was thinking, the teacher commented on why the student's solution was correct. In segments rated 2, the teachers stopped after acknowledging the correct answer (line 2: “Yes, it is.”). However, in segments rated 3 and 4, the teachers sometimes continued by asking a question (as in lines 2 and 4) or elaborating on the answer's qualities (line 6).

The teachers sometimes did these evaluations to the whole class. When teachers noticed that many students were struggling to complete a task, they brought the class's attention to a solution used by a student who had completed the task successfully. For the whole class, the teachers spent more time evaluating the student methods and carefully addressing the problems that the class was having.

Feedback on Mathematical Practices. Only one segment showed a teacher consistently giving feedback on mathematical practices:

Example 5:

The teacher posed a problem to the students: find a mathematical expression showing how many chairs are needed for the table arrangement in Fig. 3, when there are (a) 4 tables, (b) 25 tables, and (c) *n* tables. The teacher told the students to work in pairs and emphasized that the most important part of the task was to discuss their thoughts with their partner, that the thoughts behind the answers were as important as the answer itself, and that their thoughts must be written down.

While the students worked, the teacher talked to nine pairs of students in turn. First, she asked what they had discussed. After the students answered, she gave suggestions about how to continue, providing feedback according to each pair's specific challenges. Here are her suggestions to four of the pairs:

Teacher: I am wondering if the calculations you are telling me match the expressions you have written in your books. Have you talked to Hedda about this? Go over that again. Together. Check if both agree.

Teacher: You have proposed a formula here. One should check if one's proposals are correct. Continue now by trying to find out if it is correct. How would you argue? Discuss this.

Teacher: Now you have solved (a) and (b) and struggle with (c). It is often a good idea to think about whether the same approach will work or not. […] Discuss what they really mean by *n* tables.

Teacher: Good that you drew a figure. That's very clever because you can see patterns that are hard to see without [it].

In the four instances, the feedback gives feedback on the students' problem-solving, justification, and mathematical discourse, which are considered important mathematical practices that students should acquire. For example, the feedback in the first instance serves to make the students aware that their answers lack justification and proposes a way forward (i.e., “Discuss this”). The feedback in this segment addressed neither procedures nor concepts, but it evaluated and gave suggestions for the students’ engagement in mathematical practices, such as problem-solving, argumentation, and discourse.

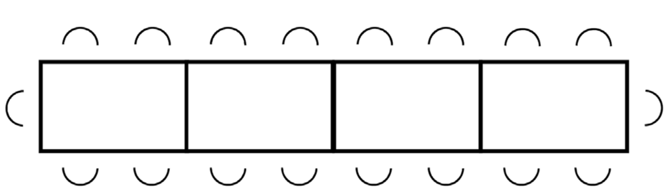

Fig. 3. Seating arrangement of the task in Example 8.

There were isolated examples of feedback on mathematical practices in other teachers’ lessons. This feedback was often provided to a student who was working on more difficult tasks than those being solved by the rest of the class. For example, in one class, all the students were working on textbook exercises except for one girl who was working on more cognitively demanding tasks that the teacher had assigned her. The teacher provided procedural feedback to all the students except this girl. The teacher provided her the feedback that one answer was not fully justified and, later in the segment, that an answer could be further generalized. Both feedback instances supported the students in the mathematical practices of justification and generalization.

In sum, we have analyzed teachers use of feedback using frequency and distribution of feedback drawing on the Feedback element as defined in PLATO observation manual (step 1), and characteristics of these feedback instances focusing for procedural feedback, conceptual feedback, and feedback on mathematical practices (step 2).

5. Discussion

The study's first aim was to determine whether the teachers' feedback aligned with generic indicators of quality feedback. In rating every 15 min of instruction ($n = 520$) of 172 lessons using a standardized observation instrument, we found that in about 67% of the segments, the teachers provided only procedural feedback or vague and perfunctory feedback. However, in 19% of the segments, the teachers' feedback included substantive elements. The feedback ratings were unequally distributed among the teachers; 14 of 47 teachers (30%) did not teach a higher-rated segment. Of the 34 teachers (70%) that did teach higher-rated segments, only five (11%) received the highest rating for a segment, meaning that they consistently provided substantive feedback that was likely to help the student.

We warn against overgeneralizing the finding that 30% of teachers never provided feedback other than procedural or vague and perfunctory feedback. Our results do not imply that these teachers did not attend to conceptual understanding or support students' engagement in mathematical practices, only that they did not do so in their feedback. The teachers may have supported these competencies through other instructional practices, such as providing tasks or explanations. Providing procedural feedback may also lessen student's cognitive load, making it easier for them to focus on understanding the concepts. Moreover, the prevalence of procedural feedback during individual work and whole-class instruction should not be assumed to imply that the teaching in these classrooms was “traditional,” “teacher centered,” or similar. The feedback was often procedural and prescriptive also when the teacher used extended group-work and the tasks were cognitively demanding. Conversely, Example 3 shows that feedback could attend to students' conceptual understanding when the classroom was organized for individual work and whole-class lecturing and the students worked on procedural tasks. To summarize, the teachers' instructional repertoires were varied, but based on the oral feedback analyzed in the present study, some of the classrooms had a narrow procedural focus.

The second aim was to investigate the amount of time teachers spent providing oral feedback in mathematics lessons. By identifying and time-stamping all the feedback instances in a subsample of the 15-min segments ($n = 54$), we found that—on average—the teachers spent approximately seven and a half out of the 15 min providing feedback. Orally providing feedback to students was thus a frequent teaching practice present across all the teachers and instructional formats. In particular, the teachers constantly provided feedback during individual and group work. When teachers spend so much lesson time providing feedback, it is crucial that students benefit from it. In this regard, it is positive that almost all feedback focused on mathematics and not solely on praise or encouragement, but we can question the usefulness of spending so much time providing feedback that shows students correct solution methods during individual work. Students often had the same problems with the same tasks, and the teachers gave lengthy procedural feedback to each student. The time might be better spent by highlighting and contrasting student solutions in a class discussion (Stein et al., 2008) or by re-teaching the whole class, because "if students lack necessary knowledge, further instruction is more powerful than feedback information" (Hattie & Timperley, 2007, p. 91).

The study's third aim was to identify how much and how teachers gave feedback that focused on procedures, concepts, and mathematical practices. The predominant way that the teachers provided feedback across rating levels was to explain procedures during individual work. We expected that the teachers would provide mostly procedural feedback, given that (a) mathematics teaching has a procedural focus in many countries (Hiebert et al., 2003), (b) it is difficult to provide conceptual feedback (Son & Sinclair, 2010), especially "on the spot" in classroom situations (Pankow et al., 2018), and (c) teachers may have been taught mathematics as a collection of procedures themselves so they may not have the required conceptual understanding. However, in 19% of the segments the teachers gave feedback on "substantive elements" that was mainly conceptual feedback. A central feature of how the teachers provided conceptual feedback was that they provided it while showing, explaining, or commenting on a procedure, as in Example 3. Why did the teachers address concepts in their feedback mainly by explaining procedures? Due to the challenges mentioned above, it seems reasonable to speculate that the concrete solution procedures may have served as a scaffold for the teachers in explaining concepts to the students, but further work must be conducted before concluding.

To the best of our knowledge, this study is the first to investigate feedback on mathematical practices across several classrooms (for case studies, see Moschkovich, 2015; Selling, 2016). Although several teachers sometimes provided feedback on mathematical practices to individual students, only one teacher consistently did so (Example 5). It is unsurprising that feedback on mathematical practices is rarely provided because of the lack of research, systematic training, and professional development that would enable mathematics teachers to provide it in the classroom. Now that mathematical practices are a core part of curricula in many countries, research on how to enable teachers to support them—for example with feedback—should follow suit (Schoenfeld, 2015).

5.1. Implications for teacher education and professional development

The study has important implications for teacher education and professional development. First, if teacher feedback shall support students in developing the competencies required in newer curricula, teachers must provide feedback on procedural skills, conceptual understanding, and mathematical practices. The results show that helping teachers provide conceptual feedback does not necessarily imply making major changes to their lessons, as procedural feedback and conceptual feedback were provided in similar situations. This does not imply that it will be easy to change teachers' feedback to have a conceptual focus. Teachers themselves often lack strong conceptual knowledge (Hannigan et al., 2013) and this is needed along with pedagogical knowledge to know when a student's struggle stems from conceptual problems and how to create a correct conceptual explanation (Son & Sinclair, 2010).

To further strengthen this practice, teachers and teacher candidates must be trained in how to provide conceptual feedback, not only as a general guideline, but as a specific practice that can be trained, rehearsed, and perfected. Videos and observation instruments are promising tools in this regard. Observation instruments have been used to support teachers in having collaborative conversations about their own practice that lead to lasting changes in teaching (Gore et al., 2017). Reports from using videos in teacher education and professional development that focus on teachers' use of scaffolding techniques such as feedback, suggest that a targeted focus on specific teaching practices strengthens prospective teachers' practical classroom repertoires (Borko et al., 2014; Stockero et al., 2017).

Through the AfL policy initiative, training in feedback has spread to countries around the world (Birenbaum et al., 2015), an example being the nation-wide professional development effort in Norway. Most of these initiatives have not been subject-specific (Bennett, 2011; Birenbaum et al., 2015), and our study suggests that they could benefit from that: both the procedural feedback and the conceptual feedback seen in our study fulfill many of the common prescriptions for high-quality feedback described in seminal reviews like Hattie and Timperley (2007) and Shute (2008). That is, they do not direct attention to the self, they are specific to the students' work, and they contain information about how to improve. However, students may learn very different aspects of mathematics when teachers provide feedback only on the correct execution of procedures and when they also include conceptual links or a focus on mathematical practices. An implication is that teacher trainers can make teachers aware of how classroom feedback tends to focus students' attention on a procedural view of mathematics and provide examples of alternative foci for feedback.

6. Conclusion

Because of the large number of lessons analyzed, and the focus on the three central aspects of mathematical competency, our findings contribute substantially to the knowledge of classroom feedback practices in mathematics. The study provides knowledge about the prevalence of feedback addressing procedural skills, conceptual understanding, and engagement in mathematical practices and detailed examples to display teachers' feedback repertoires. Such knowledge is useful to ground future research, teacher education and professional development in actual classroom practice.

Funding

This work was funded by the Research Council of Norway. The authors declare that there were no conflicts of interest.

References

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*(1), 127–145. https://doi.org/10.1037//0012-1649.35.1.127
Ball, D. L. (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education*. Rand Corporation.
Bartell, T. G., Webel, C., Bowen, B., & Dyson, N. (2013). Prospective teacher learning: Recognizing evidence of conceptual understanding. *Journal of Mathematics Teacher Education, 16*(1), 57–79. https://doi.org/10.1007/s10857-012-9205-4
Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2018). Qualities of classroom

observation systems. *School Effectiveness and School Improvement*, 1–27. https://doi.org/10.1080/09243453.2018.1539014
Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678
Bergem, O. K. (Ed.). (2014). *STEM education (REALFAG), white paper from national expert group on STEM education in Norway*. Ministry of Education and Research. https://www.regjeringen.no/globalassets/upload/kd/vedlegg/rapporter/rapport_fra_ekspertgruppa_for_realfagene.pdf.
Berlin, R., & Cohen, J. (2020). The convergence of emotionally supportive learning environments and college and career ready mathematical engagement in upper elementary classrooms. *AERA Open, 6*(3). https://doi.org/10.1177/2332858420957612, 2332858420957612.
Bieda, K. N. (2010). Enacting proof-related tasks in middle school mathematics: Challenges and opportunities. *Journal for Research in Mathematics Education, 41*(4), 351–382.
Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., & Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education, 13*(1), 117–140. https://doi.org/10.1177/1478210314566733
Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102
Booth, J. L., McGinn, K. M., Barbieri, C., Begolli, K. N., Chang, B., Miller-Cotto, D., Young, L. K., & Davenport, J. L. (2017). Evidence for cognitive science principles that impact learning in mathematics. In *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 297–325). Academic Press. https://doi.org/10.1016/B978-0-12-805086-6.00013-8.
Borko, H., Jacobs, J., Seago, N., & Mangram, C. (2014). Facilitating video-based professional development: Planning and orchestrating productive discussions. In Y. Li, E. A. Silver, & S. Li (Eds.), *Transforming mathematics instruction: Multiple approaches and practices* (pp. 259–281). Springer International Publishing. https://doi.org/10.1007/978-3-319-04993-9_16.
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
Brevik, L. M. (2019). Explicit reading strategy instruction or daily use of strategies? Studying the teaching of reading comprehension through naturalistic classroom observation in English L2. *Reading and Writing, 32*(9), 2281–2310. https://doi.org/10.1007/s11145-019-09951-w
Brodie, K. (2000). Teacher intervention in small-group work. *For the Learning of Mathematics, 20*(1), 9–16. https://doi.org/10.2307/40248313
Burkhardt, H., & Schoenfeld, A. H. (2019). Formative assessment in mathematics. In *Handbook of formative assessment in the disciplines* (pp. 35–67). Routledge. https://doi.org/10.4324/9781315166933-3.
Cohen, J. (2018). Practices that cross disciplines?: Revisiting explicit instruction in elementary mathematics and English language arts. *Teaching and Teacher Education, 69*, 324–335. https://doi.org/10.1016/j.tate.2017.10.021
Education Act. (2006). *Act no. 724 of 23 June 2006 Regulations to the education Act*. https://lovdata.no/dokument/SF/forskrift/2006-06-23-724.
Fyfe, E. R., & Brown, S. A. (2017). Feedback influences children's reasoning about math equivalence: A meta-analytic review. *Thinking & Reasoning, 24*(2), 1–22. https://doi.org/10.1080/13546783.2017.1359208
Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology, 104*(4), 1094–1108. https://doi.org/10.1037/a0028389
Gamlem, S. M. (2015). Feedback to support learning: Changes in teachers' practice and beliefs. *Teacher Development, 19*(4), 461–482. https://doi.org/10.1080/13664530.2015.1060254
Gamlem, S. M., & Munthe, E. (2014). Mapping the quality of feedback to support students' learning in lower secondary classrooms. *Cambridge Journal of Education, 44*(1), 75–92. https://doi.org/10.1080/0305764X.2013.855171
Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H., & Lubans, D. (2017). Effects of professional development on the quality of teaching: Results from a randomised controlled trial of Quality Teaching Rounds. *Teaching and Teacher Education, 68*, 99–113. https://doi.org/10.1016/j.tate.2017.08.007
Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2014). The test matters: The relationship between classroom observation scores and teacher value added on multiple types of assessment. *Educational Researcher, 43*(6), 293–303. https://doi.org/10.3102/0013189X14544542
Grossman, P., Loeb, S., Cohen, J., & Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *American Journal of Education, 119*(3), 445–470. https://doi.org/10.1086/669901
Hannigan, A., Gill, O., & Leavy, A. M. (2013). An investigation of prospective secondary mathematics teachers' conceptual knowledge of and attitudes towards statistics. *Journal of Mathematics Teacher Education, 16*(6), 427–449. https://doi.org/10.1007/s10857-013-9246-3
Hartmann, D. P., Barrios, B. A., & Wood, D. D. (2004). Principles of behavioral observation. In , *Comprehensive handbook of psychological assessment: Vol. 3. 108–127*. John Wiley & Sons Inc. Behavioral assessment.
Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. E. Mayer, & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (p. 249). Routledge.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112, 10.3102003465430298487.
Hays, M. J., Kornell, N., & Bjork, R. A. (2010). The costs and benefits of providing feedback during learning. *Psychonomic Bulletin & Review, 17*(6), 797–801. https://doi.org/10.3758/PBR.17.6.797
Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chui, A. M.-Y., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterback, W., Manaster, C., Gonzales, P., & Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2003013.
Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 371–404). Information Age Publishing (f) http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.405.3591&rep=rep1&type=pd.
Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27).
Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56–64. https://doi.org/10.3102/0013189X12437203
Hill, H. C., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review, 83*(2), 371–384. https://doi.org/10.17763/haer.83.2.d11511403715u376
Ho, A. D., & Kane, T. J. (2013). The reliability of classroom observations by school personnel. In *MET project*. Bill & Melinda Gates Foundation. Research paper. https://eric.ed.gov/?id=ED540957
Inagaki, K., Morita, E., & Hatano, G. (1999). Teaching-learning of evaluative criteria for mathematical arguments through classroom discourse: A cross-national study. *Mathematical Thinking and Learning, 1*(2), 93–111. https://doi.org/10.1207/s15327833mtl0102_1
Klette, K., & Blikstad-Balas, M. (2017). Observation manuals as lenses to classroom teaching: Pitfalls and possibilities. *European Educational Research Journal*. https://doi.org/10.1177/1474904117703228, 1474904117703228.
Klette, K., Blikstad-Balas, M., & Roe, A. (2017). Linking instruction and student achievement. A research design for a new generation of classroom studies. *Acta Didactica Norge, 11*(3). https://doi.org/10.5617/adno.4729
Klette, K., Roe, A., & Blikstad-Balas, M. (2021). Do all student groups benefit from what is regarded as high-quality teaching? In K. Klette, M. Tengberg, & M. Blikstad-Balas (Eds.), *Ways of measuring teaching quality*. Oslo University Press.
Kluger, A., & Denisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2). http://proxies.apa.org/journals/bul/119/2/254.
Krammer, K. (2009). *Individuelle lernunterstützung in schülerarbeitsphasen: Eine videobasierte analyse des unterstützungsverhaltens von lehrpersonen im mathematikunterricht*. Waxmann.
Li, N., Cao, Y., & Mok, I. A. C. (2016). A framework for teacher verbal feedback: Lessons from Chinese mathematics classrooms. *Eurasia Journal of Mathematics, Science and Technology Education, 12*(10). https://doi.org/10.12973/eurasia.2016.1298a
Li, M., Yin, Y., Ruiz-Primo, M. A., & Morozov, A. (2011). *Identifying effective feedback practices on student mathematics learning: A literature synthesis*. New Orleans, LA: Annual Meeting of the American Educational Research Association.
Luoto, J. M., Klette, K., & Blikstad-Balas, M. (in review). Patterns of instructional quality in Finnish and Norwegian lower secondary mathematics classrooms.
Mahan, K. R., Brevik, L. M., & Ødegaard, M. (2021). Characterizing CLIL teaching: New insights from a lower secondary classroom. *International Journal of Bilingual Education and Bilingualism, 24*(3), 401–418. https://doi.org/10.1080/13670050.2018.1472206
Mangold. (2019). *Behavior research with Mangold Interact*. https://www.mangold-international.com/.
Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Harvard University Press.
Ministry of Education and Research. (2006). *Læreplanverket for kunnskapsløftet*. https://www.udir.no/laring-og-trivsel/lareplanverket/finn-lareplan/lareplan-hele/.
Ministry of Education and Research. (2015). The school of the future—renewal of subjects and competences. *Official Norwegian Reports NOU 2015, 8*, 118.
Moschkovich, J. N. (2013). Issues regarding the concept of mathematical practices. In Y. Li, & J. N. Moschkovich (Eds.), *Proficiency and beliefs in learning and teaching mathematics* (pp. 257–275). SensePublishers. https://doi.org/10.1007/978-94-6209-299-0_15.
Moschkovich, J. N. (2015). Scaffolding student participation in mathematical practices. *ZDM, 47*(7), 1067–1078. https://doi.org/10.1007/s11858-015-0730-3
Munthe, E., & Melting, J. (2016). *National guidelines for the primary and lower secondary teacher education programme for years 5-10*. The National Council for Teacher Education. https://www.uhr.no/_f/p1/iecd98eeb-d012-44ce-b364-c8787ca51a95/national_guidelines_for_the_primary_and_lower_secondary_teacher_education_programme_for_years_5_10.pdf.
National Governors Association Center for Best Practices, Council of Chief State

School Officers. (2010). *Common core state standards for mathematics*. National Governors Association Center for Best Practices, Council of Chief State School Officers. http://www.corestandards.org/Math/Practice/.
OECD. (2019). *TALIS 2018 results (volume i): Teachers and school leaders as lifelong learners*. OECD. https://doi.org/10.1787/1d0bc92a-en
Olsen, R. V. (2013). Undervisning i matematikk. In M. Kjærnsli, & R. V. Olsen (Eds.), *Fortsatt en vei å gå. Norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012* (pp. 121–156). Universitetsforlaget.
Pankow, L., Kaiser, G., König, J., & Blömeke, S. (2018). Perception of student errors under time limitation: Are teachers faster than mathematicians or students? *ZDM, 50*, 631–642. https://doi.org/10.1007/s11858-018-0945-1
Praetorius, A.-K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: Looking back and looking forward. *ZDM, 50*(3), 535–553. https://doi.org/10.1007/s11858-018-0946-0
Praetorius, A.-K., McIntyre, N. A., & Klassen, R. M. (2017). Reactivity effects in video-based classroom research: An investigation using teacher and student questionnaires as well as teacher eye-tracking. *Zeitschrift für Erziehungswissenschaft, 20*(1), 49–74. https://doi.org/10.1007/s11618-017-0729-3
Ruiz-Primo, M. A., & Li, M. (2013). Examining formative feedback in the classroom context: New research perspectives. In *SAGE Handbook of research on classroom assessment* (pp. 215–232). SAGE Publications, Inc. https://doi.org/10.4135/9781452218649.
Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. https://doi.org/10.1007/BF00117714
Santagata, R. (2005). Practices and beliefs in mistake-handling activities: A video study of Italian and US mathematics lessons. *Teaching and Teacher Education, 21*(5), 491–508. https://doi.org/10.1016/j.tate.2005.03.004
Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality and Quantity, 52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8
Schoenfeld, A. H. (2014). Reflections on learning and cognition. *ZDM, 46*(3), 497–503. https://doi.org/10.1007/s11858-014-0589-8
Schoenfeld, A. H. (2015). Summative and formative assessments in mathematics supporting the goals of the common core standards. *Theory Into Practice, 54*(3), 183–194. https://doi.org/10.1080/00405841.2015.1044346
Schoenfeld, A. H. (2018). Video analyses for research and professional development: The teaching for robust understanding (TRU) framework. *ZDM, 50*(3), 491–506. https://doi.org/10.1007/s11858-017-0908-y
Selling, S. K. (2016). Making mathematical practices explicit in urban middle and high school mathematics classrooms. *Journal for Research in Mathematics Education, 47*(5), 505–551. https://doi.org/10.5951/jresematheduc.47.5.0505
Shavelson, R. J. (2003). *On the integration of formative assessment in teaching and learning with implications for teacher education*. Padova, Italy: Paper Presented at Biannual Meeting of the European Association for Research on Learning and Instruction, 2003.
Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795
Small, M., & Lin, A. (2018). Instructional feedback in mathematics. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 169–190). Cambridge University Press. https://doi.org/10.1017/9781316832134.
Smith, J. K., & Lipnevich, A. A. (2018). Instructional feedback. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 591–603). Cambridge University Press. https://doi.org/10.1017/9781316832134.
Son, J.-W., & Crespo, S. (2009). Prospective teachers' reasoning and response to a student's non-traditional strategy when dividing fractions. *Journal of Mathematics Teacher Education, 12*(4), 235–261. https://doi.org/10.1007/s10857-009-9112-5
Son, J.-W., & Sinclair, N. (2010). How preservice teachers interpret and respond to student geometric errors. *School Science & Mathematics, 110*(1), 31–46. https://doi.org/10.1111/j.1949-8594.2009.00005.x
Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313–340. https://doi.org/10.1080/10986060802229675
Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455–488. https://doi.org/10.2307/1163292
Stockero, S. L., Rupnow, R. L., & Pascoe, A. E. (2017). Learning to notice important student mathematical thinking in complex classroom interactions. *Teaching and Teacher Education, 63*, 384–395. https://doi.org/10.1016/j.tate.2017.01.006
Tengberg, M., Bommel, J. van, Nilsberth, M., Walkert, M., & Nissen, A. (2021). The quality of instruction in Swedish lower secondary language arts and mathematics. *Scandinavian Journal of Educational Research*, 1–18. https://doi.org/10.1080/00313831.2021.1910564
UDIR. (2013). *Curriculum for the common core subject of mathematics (MAT1-04)*. Kunnskapsdepartementet. http://data.udir.no/kl06/MAT1-04.pdf?lang=http://data.udir.no/kl06/eng.
UDIR. (2018). *Observations on the national assessment for learning programme (2010-2018): Skills development in networks*. Final report. https://www.udir.no/contentassets/596e196f7b9d459e96e64ecc4b463faa/the-norwegian-assessment-for-learning-programme_final-report-2018.pdf.
Van der Kleij, F. M. (2019). Comparison of teacher and student perceptions of formative assessment feedback practices and association with individual student characteristics. *Teaching and Teacher Education, 85*, 175–189. https://doi.org/10.1016/j.tate.2019.06.010
Voerman, L., Meijer, P. C., Korthagen, F. A. J., & Simons, R. J. (2012). Types and frequencies of feedback interventions in classroom interaction in secondary education. *Teaching and Teacher Education, 28*(8), 1107–1115. https://doi.org/10.1016/j.tate.2012.06.006
Wiliam, D. (2018). Feedback: At the heart of—but definitely not all of—formative assessment. In A. A. Lipnevich, & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback*. Cambridge University Press. https://doi.org/10.1017/9781316832134.
Wood, T. (1998). Alternative patterns of communication in mathematics classes: Funneling or focusing? In H. Steinbring, M. G. B. Bussi, & A. Sierpinska (Eds.), *Language and communication in the mathematics classroom*. National Council of Teachers of Mathematics.