What Went Well, What Went Badly? Teachers' and Students' Perspectives on Remote Mathematics Teaching During Pandemic School Closure

Paul Drijvers, Filip Moons, Marcel Klinger, Daniel Thurm, Ellen Vandervieren, Heleen van der Ree, and Bärbel Barzel

Contents

P. Drijvers (✉)
Freudenthal Institute, Utrecht University, Utrecht, The Netherlands
e-mail: p.drijvers@uu.nl

F. Moons · E. Vandervieren
University of Antwerp, Antwerp, Belgium
e-mail: Filip.Moons@uantwerpen.be; ellen.vandervieren@uantwerpen.be

M. Klinger · B. Barzel
University of Duisburg-Essen, Duisburg, Germany
e-mail: marcel.klinger@uni-duisburg-essen.de; baerbel.barzel@uni-due.de

D. Thurm
University of Siegen, Siegen, Germany
e-mail: daniel.thurm@uni-siegen.de

H. van der Ree
Utrecht University, Utrecht, The Netherlands
e-mail: h.m.vanderree@uu.nl

B. Pepin et al. (eds.), *Handbook of Digital Resources in Mathematics Education*,
Springer International Handbooks of Education,
https://doi.org/10.1007/978-3-030-95060-6_33-1

Abstract

When schools closed in March 2020 due to the COVID-19 pandemic, mathematics teachers and their students were confronted with emergency remote mathematics teaching (ERMT). To investigate how they experienced this in the first months, we set up an online questionnaire for mathematics teachers and their students in Flanders (the Dutch-speaking part of Belgium), Germany, and the Netherlands. The questionnaire contained two open items, in which respondents were asked to report on what went well and what went badly during ERMT. The responses of 1599 teachers and 2196 students were analyzed using a theory-based codebook with the code families: teaching practices, didactics, assessment, situational circumstances, and student behavior. Among the main results, we find that both teachers and students report positively on the ERMT teaching practices that student-teacher interaction is an important topic and that teachers struggle with assessment in ERMT. These findings seem in agreement to other research findings. For future research, we recommend monitoring whether these experiences persist after longer periods of ERMT.

Keywords

COVID-19 · Emergency remote teaching (ERT) · Distance teaching · Emergency remote mathematics teaching (ERMT)

Introduction

In early 2020, the COVID pandemic caused a worldwide disruption of educational practice. According to UNESCO, in March 2020, close to 1.4 billion students around the globe were forced to stay home, due to the closure of schools and universities (Statista 2020). Clearly, this sudden transition to emergency remote teaching had a huge impact on both teachers and students. Teachers were faced with the challenging question of how to set up teaching at distance. Their students found themselves at home and having to find ways to deal with the new teaching formats that put high demands on self-discipline and home conditions. In addition, students missed their peers from school and experienced drastic changes in their regular lives.

To describe the initial reactions by teachers and students to emergency remote teaching, Kamanetz (2020) introduced the term panic-gogy, a contamination of panic and pedagogy. Panic-gogy concerns both teachers and their students, as it includes "how teachers are going to move into this environment with their teaching approaches" and also "understanding students' practical resources and problems,

including availability of devices and the internet, and family responsibilities" (Engelbrecht et al. 2020, p. 836).

In addition to these general challenges encountered by teachers and students alike, the different school subjects may prompt specific issues for emergency remote teaching. For the case of mathematics, teachers needed ways to deal with the use of mathematical representations such as formulas and graphs in online teaching. Also teachers wondered how to interact with students and how to engage them in formative assessment and discourse practices, which are so crucial for mathematics learning. For students, the stacking and vertical nature of mathematics raised the danger that they might get completely lost – once you miss a step in the chain of procedures and concepts, it is hard to catch up. Also, the direct exchange between the student and teacher to help reveal where the problems might lie was no longer self-evident. Both teachers and students encountered mathematics-specific issues in emergency remote teaching practices. A special issue of *Educational Studies in Mathematics* reports further on these mathematics-specific aspects (see Chan et al. 2021).

Although initial experiences with emergency remote mathematics teaching (ERMT) have been reported from the teachers' perspective (Aldon et al. 2021; Drijvers et al. 2021) and, to a more limited extent, from the students' perspective (Thurm et al. 2023), little is known about the relationships between the students' and the teachers' perspectives. It is this gap that we aim to address in this paper. The overall goal, therefore, is to explore what both teachers and students experienced as going well or going badly in the initial phase of emergency remote mathematics teaching, a crucial and an overwhelming period.

To address this topic, in March 2020, we designed and sent out a theory-based online questionnaire to mathematics teachers and their students in three neighboring, Western-European countries or regions: Flanders (FL, the Dutch-speaking part of Belgium), Germany (GE), and the Netherlands (NL). Further to earlier reports on the teachers' perspectives (Drijvers et al. 2021), we here focus on the open reactions by teachers and students on what they considered good and bad in their early experiences.

Theoretical Background

Recent Literature

Worldwide, the educational research community responded immediately to the new drastic changes in educational practices. In a very short term, many research studies were conducted and published that concerned the emergency remote teaching caused by school closure. Without the illusion of being complete – the number of publications on this topic increases daily – we highlight some research findings relevant to the topic of this paper.

Several studies stress the unexpected character of emergency remote teaching. For example, Hodges and colleagues write: "In contrast to experiences that are

planned from the beginning and designed to be online, emergency remote teaching (ERT) is a temporary shift of instructional delivery to an alternate delivery mode due to crisis circumstances" (Hodges et al. 2020, p. 7). The main issues in ERT in literature include the delivery modes and the student emergency remote education readiness (Crompton et al. 2021). In response to the school closure, Harper et al. (2021) report that schools mostly moved to asynchronous delivery modes and claim that this puts high demands on parental engagement. In most of the cases, no ready-to-use concepts for such a shift in instructional paradigm existed (Eickelmann and Drossel 2020). Whereas distance education in general is not new (Pregowska et al. 2021), notions like ERT (e.g., Ferri et al. 2020) or Panic-gogy (Engelbrecht et al. 2020) underline the pragmatic and, in many cases, ad-hoc nature of this shift. Bergdahl and Nouri (2020) stress that teachers lacked the pedagogical strategies needed in the emerging learning landscape of distance education.

Similar to the overall educational research trend, the mathematics education research community also quickly engaged in research on emergency remote mathematics teaching practices. In contrast to the above general picture, Drijvers et al. (2021) found that teachers in Flanders and the Netherlands, and to a lesser extent in Germany, did use synchronous delivery modes such as videoconferencing and were getting increasingly confident with that. Aldon et al. (2021) showed how mathematics teachers in France, Israel, Italy, and Germany experienced challenges in supporting students' learning and developing assessment within ERT. Supporting students who face difficulties and exploiting potentialities for fostering typical mathematical processes turned out to be hard, according to the teachers. In line with this, Rodríguez-Muñiz et al. (2021) described that mathematics teachers reported adequate digital competence but also recognized a need for professional development. Hodgen et al. (2020) also claimed that mathematics teachers in the UK struggled with the limited opportunities that ERT offers to students for engaging in mathematical discourse or for receiving formative feedback. These findings, however, were not yet available at the time we set out the study reported in this paper.

Initial Theoretical Framework

In search for a framework that would help us capture students' and teachers' positive and negative experiences in the initial phase of emergency remote mathematics teaching in 2020, we distilled five key aspects from literature and previous research (Drijvers et al. 2021). These were labeled: teaching practices, didactical approaches, assessment, situational circumstances, and student behavior.

The first aspect of the *teaching practices* that are put into play is expected to be an important factor in teachers' and students' appreciation of education, and this is confirmed by recent studies (Aldon et al. 2021). The model of instrumental orchestration offers a lens to describe these practices and the ways in which they are used (Drijvers et al. 2010). Initial findings suggest that big differences in teaching practices exist both at the level of individual teachers and schools and at the level of regions and countries (Drijvers et al. 2021). These practices also put demands on

the available technological infrastructure (Eickelmann and Drossel 2020) and may impact on teacher self-management, workload, and personal well-being (Klapproth et al. 2020; Sokal et al. 2020).

The second aspect of *didactical approaches* refers to the choices that teachers make while designing and delivering emergency remote teaching. For example, how to deal with the wish to interact with students? How to find a balance between teaching basic skills and higher-order learning goals or between teaching new topics and rehearsing older content (Calder et al. 2021)? How to flexibly put orchestrations into action, and how to tailor the teaching to students' individual needs? For example, Hodgen et al. (2020) claim that lockdown has provided only limited opportunities for students to engage in mathematical talk or metacognitive activities. We wonder which didactical approaches can be recognized in teachers' and students' responses to the ERT.

As a third aspect, both *formative and summative assessments* are important in mathematics education and are a concern to teachers. In an ERT setting, the regular means of assessment used in face-to-face teaching cannot be applied, and the traditional paper-and-pen tests with students sitting in the classroom need to be replaced by assessment "through technology" formats (Stacey and Wiliam 2013; Drijvers et al. 2016). How do students and teachers experience the delivery of appropriate formative feedback and the use of online assessment tools?

Fourth, literature describes concerns about the impact of students' *situational circumstances* on their experiences of remote teaching. There are differences between students with respect to home access to technology (Ferri et al. 2020; Huber et al. 2020). Parental involvement and support may vary (Ferri et al. 2020). Some studies report challenges with respect to student equity and well-being (Engzell et al. 2021; Misirli and Ergulec 2021). We wonder whether these situational factors can be recognized in teachers' and students' reports on their experiences.

Fifth and final, it can be expected that *student behavior* will change in a remote setting, compared to the traditional classroom setting. We expect that higher demands are put on students' self-management, motivation, and concentration in ERT, compared to the regular school setting (Tan 2020; Altuntaş and Tekeci 2020). Also, teachers may feel they have less means to influence their students' behavior, for example, simply due to technological limitations or issues. How do students and teachers report on student behavior?

These five aspects – teaching practices, didactical approaches, assessment, situational circumstances, and student behavior – guided the phrasing of the research questions, the design of the online questionnaire, and the design of the codebook for analyzing the open questionnaire items central to this paper.

Research Questions

In line with the aims of the study, the available literature, and the initial framework, the study's research questions are as follows: What do mathematics teachers and their students in Flanders, Germany, and the Netherlands report as going well or

going badly in the early phase of emergency remote mathematics teaching in 2020 in relation to the following:

(RQ1) Teaching practices
(RQ2) Didactics
(RQ3) Assessment
(RQ4) Situational circumstances
(RQ5) Student behavior

Methods

To address the above research questions, we designed and distributed an online questionnaire for mathematics teachers and their students in Flanders, Germany, and the Netherlands. First, mathematics teachers were approached to fill in the teacher questionnaire. Next, they were encouraged to identify students from one of their classes where they felt emergency remote mathematics teaching worked best. Students from such classes were then invited to complete the student questionnaire. The teacher and the student questionnaires contained two common open questions: "What went well during the distance math teaching?" and "What went badly during the distance math teaching?". These questions, aimed to give the respondents the opportunity to freely express the ups and downs of the early phases of ERMT. The responses were analyzed using qualitative research methods, which we describe in the section "Data Analysis." The study design was approved by Utrecht University's Science-Geo Ethics Review Board (for German and Dutch data) and the University of Antwerp's Ethics Committee for the Social Sciences and Humanities (Flemish data). Guidelines with respect to privacy and data management were respected. The results of the teacher responses to the closed questionnaire items are described in Drijvers et al. (2021); the results of the student responses to the closed questionnaire items are presented in Thurm et al. (2023).

Research Context

Flanders (FL, the Dutch-speaking part of Belgium), Germany (GE), and the Netherlands (NL) are three adjacent countries in Western Europe, which share an education system of primary and secondary schools. Secondary schools include students of 12–18 years old in Flanders and the Netherlands and 10–18 years old in Germany. The three countries had similar COVID-19 pandemic restrictions in that the political decision was taken to close secondary schools from March 15, 2020, until early June (see Fig. 1). Whereas Flanders and the Netherlands have a nationwide education system, Germany's federal structure includes 16 states (the so-called "Bundesländer"), each with its own education system. In at least some of them, the local education ministries suggested that the teachers should focus their lessons on rehearsing and practicing existing curriculum content during the school closure.

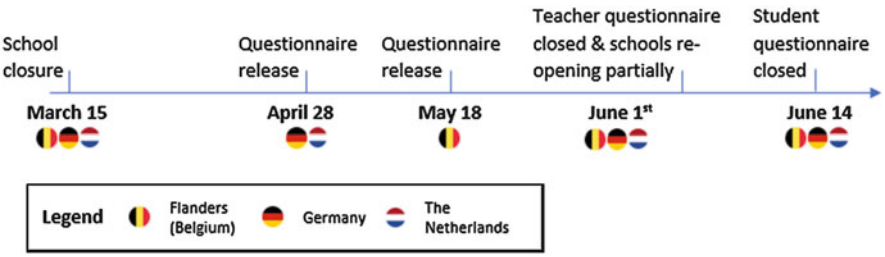

Fig. 1 Timeline for the 2020 school closures, re-openings, and questionnaire distribution in the three countries

Furthermore, in some German states, students' performance during school closure could not be used for grading purposes. In Flanders, teachers were obliged to rehearse topics already taught until the Easter holidays (April 19). After that, schools were encouraged to teach new topics. As Flemish education is organized in a decentralized way, schools were free to decide whether to re-open toward the end of May and if and how. They would organize school-based final exams. In the Netherlands, the Ministry of Education decided to cancel the national central final examinations (CE). Students received their secondary school diploma based on previously administered school-based assessments.

Data Collection

The online questionnaires, designed and implemented in Qualtrics, opened on April 28, 2020, in Germany and the Netherlands, and on May 18, 2020, in Flanders. The online teacher questionnaire closed on June 1, 2020. The student questionnaire closed 2 weeks later, on June 14, 2020. The invitation to participate was targeted to mathematics teachers through professional online newsletters, direct mails to associations of mathematics teachers, dedicated social media groups, teacher association websites, and messages to school principals. As a minor incentive to participate, all teachers were sent a report of the survey results by email when the research was completed.

Participants

In total, 1719 teachers completed the teacher questionnaire: 1599 of them (92.8%) answered at least one of the two open questions; 47 of them (2.9%) only answered the "What went good" question, and 64 teachers (4%) only reported what went badly. Of the 1599 teachers answering at least one open question, 358 (22.4%) came from Flanders, 1040 (65.0%) from Germany, and 201 (12.6%) from the Netherlands. The average age of the teachers answering at least one of the two open questions was 43.7 years, with a standard deviation of 11.1. The gender distribution of the teachers in the sample, and for each country, can be found in Table 1.

Table 1 Gender distribution of teachers and students per country

	Teachers				Students			
				Total				Total
Female	270 (75.4%)	583 (56.1%	113 (56.2%)	966 (60.4%)	623 (58.0%)	443 (55.3%)	190 (59.4%)	1256 (57.2%)
Male	88 (24.6%)	451 (43.4%)	87 (43.3%)	626 (39.1%)	448 (41.7%)	343 (42.8%)	129 (40.3%)	920 (41.9%)
Other	0 (0.0%)	2 (0.2%)	1 (0.5%)	3 (0.2%)	4 (0.4%)	10 (1.2%)	1 (3.1%)	15 (0.7%)
Missing	0 (0.0%)	4 (0.4%)	0 (0.0%)	4 (0.3%)	0 (0.0%)	5 (0.6%)	0 (0.0%)	5 (0.2%)
Total	358 (100%)	1040 (100%)	201 (100%)	1599 (100%)	1075 (100%)	801 (100%)	320 (100%)	2196 (100%)

Table 2 International class grades distribution of the sample

				Total
5th grade (10–11-year-olds) *Only in Germany*		60 (7.5%)		60 (2.7%)
6th grade (11–12-year-olds) *Only in Germany*		60 (7.5%)		60 (2.7%)
7th grade (12–13-year-olds)	32 (3.0%)	65 (8.1%)	17 (5.3%)	114 (5.2%)
8th grade (13–14-year-olds)	114 (10.6%)	67 (8.4%)	28 (8.8%)	209 (9.5%)
9th grade (14–15-year-olds)	131 (12.2%)	77 (9.6%)	30 (9.4%)	238 (10.8%)
10th grade (15–16-year-olds)	184 (17.1%)	149 (18.6%)	81 (25.3%)	414 (18.9%)
11th grade (16–17-year-olds)	279 (25.6%)	185 (23.1%)	155 (48.4%)	619 (28.2%)
12th grade (17–18-year-olds)	334 (31.1%)	113 (14.1%)	0 (0.0%)	447 (20.4%)
13th grade (18–19-year-olds)	1 (0.0%)	11 (1.4%)	0 (0.0%)	12 (0.5%)
Missing	0 (0.0%)	14 (1.7%)	9 (2.8%)	23 (1.0%)
Total	1075 (100%)	801 (100%)	320 (100%)	2196 (100%)

In total, 2552 student responded to the questionnaire, of which 2196 students (86%) answered at least one of the two open questions. Noteworthy in this respect is that 424 students, nearly 1 out of 5, only answered the "What went well?" question, leaving the "What went badly" question empty. Conversely, far fewer students (65, 3%) responded to the "What went badly" question without reporting anything good. Of the 2196 students answering at least one of the two open questions, 1075 (49.0%) were from Flanders, 801 (36.5%) from Germany, and 320 (14.5%) from the Netherlands. The average age of these students was 15.6 years, with a standard deviation of 1.94. The international class grades distribution of the student sample can be found in Table 2. It was obtained by converting each local class level to the international US class grades system. The gender distribution of the students is shown in Table 1. Relative to the country's populations (6.6 M for Flanders, 83.0 M for Germany, and 17.3 M for the Netherlands), Flanders had the highest response rate for both teachers and students.

Data Analysis

Thematic analysis was conducted on the two open questions using NVivo 12, software for qualitative data analysis. We started with the teachers' responses. First, an initial codebook was created based on the five perspectives introduced in the theoretical background (section "Data Collection"). The next step was an exploratory "bottom-up" data analysis approach which led to the identification of sub-codes within the five main codes. The units of analyses were:

- A single phrase
- Several phrases if they were clearly on the same topic and connected, e.g., if the second one explained the first one
- A part of a phrase about one topic

Hence, one response might be assigned multiple codes. For example, "Students find the quizzes that I do with Mentimeter very motivating" was double-coded to both *as motivation* and *as teaching activity.*

Next, three researchers (one from each country) used the codebook to code the same selection of 15 randomly chosen items from each country. After coding, the overall Cohen's kappa was calculated for each sub-code. Codes with a low Cohen's kappa value (<0.5) and a low agreement (<80%) were discussed and the code descriptions modified to achieve a more uniform interpretation within the research team. NVivo calculates the value of Cohen's kappa at the character level (Sang-Yeon et al. 2016), which can lead to an underestimation of the inter-rater agreement. Therefore, we also considered the percentage of agreed items. After four iterations, a final codebook was agreed that generated an acceptable Cohen's kappa value or high agreement for all codes. In the final iteration, the three researchers coded one specific random selection of 100 teachers' responses. The inter-rater reliability results can be found in Table 3 in the columns κ (Cohen's kappa) and % A (percentual agreement) for all codes. Finally, the three researchers coded all teachers' answers of their country using the final codebook for the teachers' data.

The analysis of the students' responses was done after the coding of the teachers' data. As a result, the students' codebook was initially based on the finalized teachers' codebook. Initially, one researcher oversaw the coding of all the students' answers. Based on the findings in the data, some additional codes were added to the students' codebook. Next, a second researcher coded a random selection of 20 students' answers to see if the adapted codebook made sense. After discussing the adapted students' codebook, this pair of researchers coded 10% of the total students' answers together: 100 students' answers from Flanders, 100 students' answers from Germany, and 50 from the Netherlands. Finally, they discussed the coding and amended the codes with a Cohen's kappa below 0.6. Finally, the first researcher coded all student answers using the finalized student codebook. The codebook and the inter-rater reliability results for the students can be found in Table 3.

Results

Table 3 provides a general overview of all codes, their inter-rater reliability estimates, and the frequencies of occurrence disaggregated by country and the two prompts "What went well"/"What went badly" for both teachers and students from all countries. Subitems preceded by a – symbol are included in both codebooks, items with a # symbol only in the teachers' codebook, and items with a * symbol only in the students' codebook. The colors of the cells indicate how often the code appeared proportionally to the frequencies within the same column. Light colors mean that the code appeared only a few times within a country or in total; dark colors indicate a frequently occurring code. In the following paragraphs, we draw on data from the table to discuss the research findings in relation to the five aspects of the research question.

Table 3 The teachers' and students' codebook, along with the number of occurrences of every code for each country and the countries in total, the Cohen's kappa κ, and the percentual agreement $\%A$

Teachers

	κ	$\%A$	What went well?				What went badly?			
						Total				Total
(RQ1) Teaching practices										
- Teaching Activity	0.60	87.4	187	385	96	668	14	91	13	118
- Technology	0.59	95.3	40	276	21	337	44	137	8	189
- Workload	0.70	97.7	4	16	0	20	42	76	1	119
- Creating new digital materials	0.84	98.6	62	92	10	164	2	8	0	10
# Colleagues	0.67	99.0	15	10	2	27	1	6	4	11
# Self-management	0.32	99.3	5	9	0	14	3	18	0	21
(RQ2) Didactics										
- Pedagogy	0.57	94.1	32	104	9	145	78	90	19	187
- Flexibility	0.80	99.7	59	17	4	80	8	20	1	29
- Interaction	0.59	90.5	37	276	12	325	107	307	42	456
- Content-specific comments	0.90	99.6	3	17	0	20	15	32	1	48
(RQ3) Assessment										
- Formative assessment	0.55	92.7	32	225	13	270	117	279	36	432
- Summative assessment	0.73	98.8	10	9	3	22	46	79	11	136
(RQ4) Situational circumstances										
- Student equity	0.53	95.7	1	26	0	27	16	230	2	248
- Technological infrastructure	0.60	96.2	5	9	0	14	46	82	2	130
- Well-being	0.40	99.7	4	2	0	6	2	12	3	17
# Parental work	0.68	97.9	1	19	1	21	5	30	0	35
(RQ5) Student behavior										
- Student motivation	0.60	96.7	21	54	6	81	30	68	8	106
- Student presence	0.65	96.7	10	25	3	38	44	79	22	145
- Student self-management	0.63	95.7	52	84	2	138	33	66	10	109

Students

	κ	$\%A$	What went well?				What went badly?			
						Total				Total
(RQ1) Teaching practices										
- Teaching activity	0.80	96.8	253	136	231	620	81	16	75	172
- Technology	0.79	98.5	166	33	76	275	33	12	20	65
- Workload	0.87	99.6	9	2	28	39	62	8	50	120
- Materials	0.66	99.4	52	5	65	122	10	1	18	29
* Teacher's (digital) skills	0.73	99.7	22	5	18	45	4	0	4	8
* Comparison F2F - online	0.81	99.4	19	18	6	43	13	11	9	33
(RQ2) Didactics										
- Pedagogy	0.89	99.6	47	7	33	87	38	17	23	78
- Flexibility	0.72	99.7	20	8	39	67	2	1	1	4
- Interaction student-teacher	0.86	98.4	110	36	100	246	137	48	49	234
- (no) difficulty with content	0.86	99.3	74	4	44	122	107	11	45	163
* Interaction student-student	0.64	99.3	3	2	15	20	17	9	10	36
* Structure and instruction	0.91	99.4	116	30	74	220	28	0	15	43
* Teacher's aid	0.99	99.9	4	2	12	18	38	3	34	75
(RQ3) Assessment										
- Formative assessment	0.97	99.9	26	7	22	55	9	13	9	31
- Summative assessment	0.99	100	6	2	4	12	24	6	1	31
(RQ4) Situational circumstances										
- Student equity	0.92	99.8	3	2	1	6	43	3	28	74
- Technological infrastructure	0.82	99.4	10	2	6	18	84	34	78	196
- Well-being	0.87	99.9	7	2	6	15	10	0	7	17
(RQ5) Student behavior										
- Motivation and engagement	0.91	99.7	6	7	8	21	32	9	18	59
- Presence	1	100	0	2	9	11	3	3	15	21
- Self-management	0.84	99.0	212	31	115	358	34	14	28	76
* Disturbance	0.99	100	16	17	19	52	9	7	12	28
* Concentration	1	100	26	8	4	38	48	22	12	82

Color index

What went well	What went badly
Max.	Max.
Average	Average
Min.	Min.

Teaching Practices

Concerning teaching practices (RQ1), Table 3 shows that teachers in the three countries frequently reported that *teaching activities* went well, to which students seem to largely agree. These positive comments by both teachers and students clearly compensate for the negative comments. German teachers and Flemish students reported more positive outcomes concerning the technology than their colleagues from the other countries. If teachers reported on the creation of new digital materials, they were almost unanimously positive, with few bad experiences were reported. The students' answers also show a similar pattern in *materials* and *teachers' (digital) skills*.

In their frequent positive comments on teaching activity, many teachers mentioned specific teaching approach that they report had worked well: "I put the exercises to be done on a Padlet, and the students send their solutions by email," a 46-year-old German teacher explained. A notable observation here is that most references to teaching activities having gone well involved a systematic structure in the planning of activities and technologies used. For example, this 27-year-old Dutch teacher reported:

> Instead of using a whiteboard, I used a Wacom tablet and OneNote to support my explanations with text and figures. This went really well. After the lessons, I uploaded these annotations to Moodle for the students, which they really liked. Students could ask questions on a forum outside of the fixed live lessons. Also, this worked smoothly.

When students reported positively about teaching activities, many of them confined themselves to somewhat superficial comments such as "the explanation was clear." Nevertheless, many students also recognized that having a systematic structure for the ERMT helped them a lot:

> After the teacher explained the theory and demonstrated some example exercises, we got some time to make exercises independently or as group work. Everything the teacher writes down during the live lessons is sent out afterwards as a pdf, just in case you missed something. (18-year-old Flemish student)

Whereas "clear explanations" was often reported by students as a teaching practice that had gone well, surprisingly, most of the students who grumbled about teaching activities mention "unclear explanations," "explanations are too short," or "explanations go too fast." Only a few teachers noted that conveying new materials was problematic online, by mainly addressing the limitations of the online environment. Some teachers also observed that students often had difficulties understanding mathematical course materials and exercises independently, without the teacher's explanation.

Turning to technologies, most teachers stated some examples that worked well for them such as Zoom, bettermarks, the ANTON app, and Moodle. In addition, individual telephone calls and emails were noted positively with respect to the technology. The bad experiences concerned mostly practical struggles with technology for both teachers and students (e.g., unreliable connections, login issues). Many

teachers reported difficulties with the communication of written mathematical language to students. Graphics tablets and creative combinations of technologies, such as an extra webcam or the teachers' smartphone, were used to capture what was written on paper during live sessions and were often reported as technologies that had worked well. However, writing with the use of a keyboard or mouse was often reported negatively. The lack of a smooth way to write mathematics was noticed by a lot of students as well. "Sometimes you have to ask more questions to the teacher before you understand something because the teacher can't write as much," explained an 18-year-old Dutch student. In addition, the necessity to write Latex equations in some digital tools was often a hurdle for teachers who had no prior Latex experience. In addition, many teachers complained that most audience response and assessment tools often lacked the functionality to process mathematical responses. "Online, it's difficult to support students in their mathematical thinking as online tools often don't require intermediate steps. Moreover, computer language is sometimes 'too strict' for assessment purposes," said a 33-year-old Flemish teacher. The experiences with creating new digital materials mainly referred to self-made videos. Most teachers reported some advantages: students can watch them at their own pace and rewatch parts of these videos if necessary.

Regarding the sub-themes workload, colleagues, and self-management, many teachers report that organizing ERMT was more labor-intensive than regular face-to-face lessons. However, a few indicated that their work/life balance was better during the lockdown. Mixed feelings about the opportunities to collaborate and agree arrangements with colleagues were reported, which was noticed by students. Some complained about having too many tasks to handle for different subjects, ambiguous deadlines, and vague expectations that had all resulted from a lack of whole-school coordination. However, many students reported positively the many efforts made by their mathematics teachers to make the best learning environment possible in the bizarre period.

Didactics

Concerning didactics, addressed in RQ2, teachers' bad experiences with classroom interactions are striking. It is the most frequent code for the "What went badly" item for both teachers and students. Teachers also shared many good experiences on the matter (and, certainly, German teachers) but to a lesser extent. Student results show even more nuances: many negative comments on *student-teacher interaction* are reported, but there were almost as many positive remarks on the topic. Note that students hardly complained about their interactions with each other (*student-student interaction*). Furthermore, teachers tended to report slightly more negatively in relation to *pedagogy* (e.g., classroom differentiation, speed of offering new content). This contrasts with the students, who were more positive in relation to pedagogy. Overall, students seemed to appreciate the structure and instruction provided during ERMT. Teachers only rarely reported content-specific issues, which contrast with the students' responses, which frequently mentioned in positive and negative comments that are related to the *difficulty* of the provided *content*.

Almost all teachers' negative comments relating to interaction referred to some students who had completely disappeared from view during the lockdown. Teachers reported receiving very few questions from students, and the impression that they were giving monologues as most of the communication was one-way. In addition, teachers reported the lack of spontaneous communication possibilities, with most comments linking implicitly to difficulties with formative assessment:

> The personal and direct contact with students is completely missing. I have to rely on what students ask me. . . but I have never had to answer so few questions as I do now. (31-year-old German teacher)

Many students reported they needed the opportunity to ask questions but mentioned many hurdles to do so. Firstly, many students expressed some sense of shame to ask questions: "I found it difficult to ask questions, yes, we had a forum for questions, but then all your classmates see your question. I found that a bit strange," said a 16-year-old Flemish student explained. Secondly, some students noted that messaging the teacher whenever they had a question was cumbersome:

> In class, some questions are answered in less than half a minute. Now, it is difficult to communicate. I'm not going to text or call my teacher for such small questions! (14-year-old Dutch student)

Thirdly, some students stated they had difficulties making their questions clear at distance:

> If you have questions, it's sometimes tricky to pinpoint the part your question is about. In real life, you just point at it, but that is impossible for now. (16-year-old Dutch student)

Comments that reported positive interactions almost always described some aspect of practice that had gone well. For example, some practice going well, such as putting question marks in the chat so a teacher could see who was struggling or organizing "question hours" at fixed times during the school week. When students were positive about the interaction between them and the teacher, it related to their math teacher being very reachable for their questions.

When it comes to pedagogy and flexibility, a frequent remark of teachers concerned the enormous positive effect for the more proficient students to have the opportunity to work at their own pace and at a time of their choosing. These remarks often accompanied a comment that concerned student equity. However, both students and teachers often found distance learning to be rather monotonous, with few opportunities for different forms of lesson activities. Classroom differentiation and group work were only exceptionally reported as having gone well: "I divided my class into small groups during the live lessons to accommodate the different learning levels better," said a 40-year-old German teacher, while a 17-year-old German student confessed: "Individual support and challenges are no longer provided."

Assessment

In line with the many negative responses about the lack of *interaction*, teachers' responses in relation to assessment (RQ3) were also negative about *formative assessment* within ERMT. However, students' responses related to assessment practices were rare, and it appears that the picture might be more nuanced for students. When teachers reported positively on assessment, most remarks referred to the opportunities that online tools offered for immediate, automatic feedback. Teachers reported that getting an overview of the whole class in this way was useful.

Multiple teachers reported the practice of students sending solved exercises to them, mostly by taking pictures with their smartphones of handwritten exercises. Teachers corrected these exercises and sent them back. However, this approach to providing student feedback is also reported often as a negative experience:

> Every day I spent hours writing feedback, although it is unknown whether students read it at all. Moreover, tasks are possibly completed by their parents, so I lack the insight on whether students really understand the material. (29-year old German teacher)

This teacher was not the only one making this comment, as concerns about the time taken were very frequently reported in combination with giving feedback. Many colleagues also reported issues with plagiarism and students copying solutions from each other: "some students suddenly had suspiciously super good tasks," wrote a 34-year old Flemish teacher. Moving to the students, rarely made comments relating to assessment, most said that they missed the feedback and the pressure generated by tests and tasks. Only one student reported that tests were solved together with the whole class.

Another frequently reported phenomenon concerning formative assessment was the lack of teachers' observation of students, which hampered teachers' noticing behaviors (Gibson and Ross 2016):

> You don't get any visual feedback from students, this takes away so much information you naturally use in a face-to-face classroom. (41-year old German teacher)

The lack of teachers' noticing was also observed by several students: "The teacher can't see if we're following and understanding everything, so sometimes the live lessons go too quickly," said a 16-year-old Flemish student.

Teachers reported less on *summative assessment*, but again, if they did, it was almost unanimously a negative comment. These included complaints about the reliability of assessments and the associated time spent correcting. However, there were probably fewer comments relating to summative assessment because most teachers explained that it was forbidden to grade students during ERMT.

Situational Circumstances

For the situational circumstances mentioned in RQ4, German teachers reported many negative experiences concerning *student equity.* Students also almost always reported on student equity in a negative way but mention it less frequently than their teachers. Also, the *technological infrastructure* was a topic that almost only featured in the responses to the "What went badly" question for both students and teachers. However, proportionally many more students complained about *technological infrastructure* (one of the most frequent "What went badly" codes for students) than their teachers.

Teachers mentioned student equity issues mostly in a practical way, and the limited access to the internet and printers for some groups of students was the most frequently coded response. Some teachers and their schools coped with this issue by providing more extensive, self-explanatory, course materials, for students who could not attend the live distance lessons. Although most teachers only expressed negative views on student equity, the few who reported good experiences tended to concern the mathematically more proficient students: "Strong students are still getting better, can concentrate better and work diligently!", noticed a 36-year-old German teacher. Students mostly referred to classmates who struggled to attend due to the circumstances at home or technical issues. Only a few students acknowledged that it was not straightforward for them to follow the distance lessons: "There is a lot of arguing at home due to the distance lessons. I really want to go back to school as quickly as possible," said an 11-year-old Dutch student.

When students complained about the technological infrastructure, almost all complaints were about problems with certain platforms that crashed regularly, connectivity problems, slow Internet connections, Wi-Fi connectivity issues, and poor audio/video quality during video conferences, which made distance lessons hard to follow. Teachers shared the same complaints and also reported limitations to writing mathematical symbols during live lessons:

> I take the webcam of my computer to film the paper I'm writing on. However, the video quality of it is very poor because my webcam does not focus well. I should buy a writing pad, but I haven't done so yet. (45-year-old Flemish teacher)

Also, the lack of support was often noticed by teachers: "Nobody said what and how we could do it, and I'm really not an ICT person, nobody can help me," said a 47-year old German teacher.

Well-being was not frequently addressed by both students and teachers. Only a few students reported being quite happy with the distance lessons, as they felt math classes were more relaxed. Other students confessed that they had become stressed by the distance learning and the more blurrier expectations that came with it. Most teachers reported the lack of social contact and community building.

When it comes to parents, teachers welcomed the many positive comments they had received from their pupils' parents. However, they also reported that some parents and their children were unreachable.

Student Behavior

Finally, for student behavior addressed in RQ5, a very nuanced view emerges for teachers: they seem to be equally positive and negative about the *motivation* and *self-management* of the students; only the *presence* (e.g., in live sessions) was frequently reported as more negative. On the other hand, students were very positive about math distance learning giving rise to good *self-management*. Only *concentration* and *motivation and engagement* were minor negative experiences.

With respect to student motivation, teachers reported that students who already struggled with mathematics before the lockdown were even more lost during the lockdown:

> Catching up with students who have dropped out completely is almost impossible. Most of them already lost their motivation long before the school closure, but they were at least present in your classroom back then. (39-year-old Flemish teacher)

Also, the lack of possibilities for spontaneous interaction was viewed by teachers as a threat to students' motivation. Students, on the other hand, frequently linked their lack of motivation to the cancellation of tests and exams, which provided some extrinsic pressure in normal times. Surprisingly, none of the teachers made this connection. "In general, I'm much less motivated to learn and understand math really well. The positive pressure stemming from the learning for tests and the good understanding of the content is a big loss. This must change!", reported a 16-year-old Flemish student. In some places (see section "Research Context"), for some time teachers were only allowed to teach already seen materials before the lockdown. Some students felt that had negatively affected their motivation: "I feel a little unmotivated because similar tasks are given over and over again," a 17-year-old German student explained. Positive comments on students' motivation were often linked to students being actively engaged during the live sessions, some feelings of "being in the same boat" by the whole class, and high levels of students' presence.

Concerning students' presence, different teachers reported frequent absences of students during live sessions. With most students, this was a sporadic phenomenon, but teachers did report a small number of students who had completely vanished for the whole lockdown period. Some teachers also noticed that although students log in, it was not possible to really tell they were present: "If the students don't have their webcams turned on, you can't see them and their reactions. . . Are they even there?", said a 33-year-old German teacher. Only a couple of students confessed they struggled with being present in live classes: "sometimes I oversleep and come in too late," commented a 14-year-old Flemish student. Students also reported that they were not always "mentally present" and often faced disturbances and experienced concentration issues: "I can't understand why they [the teachers] don't notice we're doing completely other stuff during live sessions," added a 16-year-old Dutch student. When students' presence was reported positively, these teachers noted that they had reached (almost) all their students.

In the section "Teaching Practices," we have reported that students' view on their self-management was mostly positive. Indeed, many students appreciated that they

could work at their own pace and were not being forced to follow the rhythm of the classroom: "It went very smoothly: in a period of 3 weeks we've seen more materials than on a full trimester! Not because of our teacher, but because everyone could work independently at their own pace," said a 16-year-old Flemish student. In addition, students also saw the practical benefits of being able to choose when to work on what. Teachers noticed almost the same things: "A lot of students, definitely the mathematically stronger ones, even benefit from distance education, because there is a lot of room for their own ways of thinking, to seek for their own problem-solving strategies,. . .", said a 45 year-old Flemish teacher. When students reported negative experiences with self-management, the voluntary nature of most tasks was (again) seen as a major cause. In addition to that, some students complained they had received too many tasks to complete.

Conclusion and Discussion

Through two open items in an online questionnaire, we investigated what mathematics teachers and their students in Flanders, Germany, and the Netherlands had reported as going well or going badly during the early phase of ERMT with respect to (RQ1) teaching practices, (RQ2) didactics, (RQ3) assessment, (RQ4) situational circumstances, and (RQ5) student behavior.

As a conclusion on teaching practices, the overall results are fairly positive. The majority of comments, both by teachers and students, are positive, in particular in relation to teaching activity. Nearly one out of five students only made positive comments, leaving the "What went badly?" question response empty. For didactics, the most striking result is that many comments were coded in relation to student-teacher interaction for both teachers and students, and these were both positive and negative. Clearly, interaction is an important topic to consider in ERMT. Assessment seemed to be more of a concern to teachers than to students. Teachers struggled with ways to assess, whereas students minded less, although seemed to need it as motivation. Related to the topic of interaction, teachers have mixed feelings on opportunities for formative assessment. They comment less frequently on summative assessment, maybe because this was discouraged by education policies at the time. If they do comment on it, it is mostly in a negative sense. Within situational circumstances, teachers are mostly concerned with issues on student equity, whereas students struggle with the technological infrastructure and comment negatively on poor Internet connections and limited home facilities. With respect to student behavior, teachers have both positive and negative comments on students' self-management and motivation. Students' presence in the ERMT classroom was a concern for teachers. Students report positively about their self-management but struggled with their personal motivation and concentration.

Of course, we should consider these findings in the light of the study's limitations. A first limitation concerns the three countries involved. Even though they differ in terms of educational systems and policies, they share being West-European countries with well-established educational and technological infrastructure. These findings cannot be automatically extrapolated to other geographical regions. A second

limitation is that teachers took part on a voluntary basis, so the sample is unlikely to be representative of the teaching population in the countries. In addition, the teachers were asked to invite students to participate in the survey from classes in which the teachers had considered that the ERMT had worked best, which also might induce bias. Third, the comments on what went well and what went badly are self-reports, so may be judged in a somewhat subjective way, which of course is a general limitation of questionnaire methods. Finally, the data are collected in an early phase of ERMT, and the results cannot be extrapolated to later experiences. In fact, the development of these positive and negative comments over time is the topic of further studies.

How do these findings stand out against the body of literature described in the section "Theoretical Background"? First, we want to highlight that the overall image from the data tends to be positive. Recent literature, e.g., Kamanetz' (2020) paper on panic-gogy, suggests a state of panic in which both teachers and students found themselves after school closure. Of course, educational practices were disrupted, but the move to asynchronous teaching reported by Harper (2021) or the teachers and students being unprepared for remote teaching (Bergdahl and Nouri 2020; Crompton et al. 2021) does not resonate with the findings from our questionnaire analyses. Rather, our responses do not seem to reflect panic or stress, but appear rather relaxed, with student-teacher interaction and assessment as the most critical aspects. This might be a result of the specific countries under investigation or the voluntary basis of teacher participation, but nevertheless, we consider this as a remarkable result.

As a second point to discuss, one may wonder how specific these findings are for the subject of mathematics. Most of the reports by teachers and students seem quite general and do not refer to mathematics in particular. Do our findings, despite our focus on mathematics education, concern ERT in general, rather than ERMT, and would they be generalizable to other school subjects? Could it be that the students' responses were more based on their overall ERT experience, and not limited to the mathematics classes? Interestingly, most of the literature we refer to in the section "Theoretical Background" seem to reveal a similar phenomenon. Even if the context of most of the studies reported concern mathematics education, many results seem general and as such might be generalizable to other subjects. We conjecture that in this initial phase of ERMT the general concerns were dominating the mathematics-specific issues, even if we see some traces of them in the critical comments on teacher-student interaction and assessment, as mathematics offers some specific challenges here. It would be interesting to know whether the results for other subjects would be similar and whether the mathematics specificity would show more clearly after a longer period of ERMT experience.

Finally, the study suggests some implications for both research and ERMT practices. For educational research, a relevant next step would be to investigate the evolution of ERMT and following a longer period of school closure. A second focus would be to investigate what is general for all ERT in the results and what is specific for mathematics. For educational practice, it seems highly relevant to develop technological infrastructure that supports interaction between teachers and students and that facilitates formative and summative assessment, as these seem to be the main challenges of ERMT at present.

References

Aldon G, Cusi A, Schacht F, Swidan O (2021) Teaching mathematics in a context of lockdown: a study focused on teachers' praxeologies. Educ Sci 11(2). https://doi.org/10.3390/educsci11020038

Altuntaş O, Tekeci Y (2020) Effect of COVID 19 on perceived stress, coping skills, self-control and self-management skills. Health Econ Outcomes Res. https://doi.org/10.21203/rs.3.rs-48393/v1

Bergdahl N, Nouri J (2020) Covid-19 and crisis-prompted distance education in Sweden. Technol Knowl Learn 26:443–459. https://doi.org/10.1007/s10758-020-09470-6

Calder N, Jafri M, Guo L (2021) Mathematics education students' experiences during lockdown: managing collaboration in eLearning. Educ Sci 11(4). https://doi.org/10.3390/educsci11040191

Chan MCE, Sabena C, Wagner D (2021) Mathematics education in a time of crisis – a viral pandemic. Educ Stud Math 108:1–13. https://doi.org/10.1007/s10649-021-10113-5

Crompton H, Burke D, Jordan K, Wilson SWG (2021) Learning with technology during emergencies: a systematic review of K-12 education. Br J Educ Technol 52(4):1554–1575. https://doi.org/10.1111/bjet.13114

Drijvers P, Doorman M, Boon P, Reed H, Gravemeijer K (2010) The teacher and the tool: instrumental orchestrations in the technology-rich mathematics classroom. Educ Stud Math 75(2):213–234. https://doi.org/10.1007/s10649-010-9254-5

Drijvers P, Ball L, Barzel B, Heid MK, Cao Y, Maschietto M (2016) Uses of technology in lower secondary mathematics education; a concise topical survey. Springer. http://www.springer.com/us/book/9783319336657

Drijvers P, Thurm D, Vandervieren E, Klinger M, Moons F, van der Ree H, Mol A, Barzel B, Doorman M (2021) Distance mathematics teaching in Flanders, Germany and the Netherlands during COVID-19 lockdown. Educ Stud Math 108:35–64. https://doi.org/10.1007/s10649-021-10094-5

Eickelmann B, Drossel K (2020) Schule auf Distanz: Perspektiven und Empfehlungen für den neuen Schulalltag Eine repräsentative Befragung von Lehrkräften in Deutschland. Vodafone Stiftung Deutschland, Düsseldorf

Engelbrecht J, Llinares S, Borba MC (2020) Transformation of the mathematics classroom with the internet. ZDM 52. https://doi.org/10.1007/s11858-020-01176-4

Engzell P, Frey A, Verhagen MD (2021) Learning loss due to school closures during the COVID-19 pandemic. PNAS 118(17). https://doi.org/10.1073/pnas.2022376118

Ferri F, Grifoni P, Guzzo T (2020) Online learning and emergency remote teaching: opportunities and challenges in emergency situations. Societies 10(4). https://doi.org/10.3390/soc10040086

Gibson SA, Ross P (2016) Teachers' professional noticing. Theory Pract 55(3):180–188. https://doi.org/10.1080/00405841.2016.1173996

Harper FK, Rosenberg JM, Comperry S, Howell K, Womble S (2021) #Mathathome during the COVID-19 pandemic: exploring and reimagining resources and social supports for parents. Educ Sci 11(2). https://doi.org/10.3390/educsci11020060

Hodgen J, Taylor B, Jacques L, Tereshchenko A, Kwok R, Cockerill M (2020) Remote mathematics teaching during COVID-19: intentions, practices and equity. UCL Institute of Education, London

Hodges C, Moore S, Lockee B, Trust T, Bond A (2020) The difference between emergency remote teaching and online learning. Educause Review. Retrieved from https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Huber SG, Günther PS, Schneider N, Helm C, Schwander M, Schneider JA, Pruitt J (2020) COVID-19 und aktuelle Herausforderungen in Schule und Bildung: Erste Befunde des Schul-Barometers in Deutschland, Österreich und der Schweiz. Waxmann, Münster

Kamanetz A (2020) 'Panic-gogy': teaching online classes during the coronavirus pandemic. NPR. Retrieved from https://www.npr.org/2020/03/19/817885991/panic-gogy-teaching-online-classes-during-the-coronavirus-pandemic

Klapproth F, Federkeil L, Heinschke F, Jungmann T (2020) Teachers' experiences of stress and their coping strategies during COVID-19 induced distance teaching. J Pedagog Res 4(4):444–452. https://doi.org/10.33902/JPR.2020062805

Misirli O, Ergulec F (2021) Emergency remote teaching during the COVID-19 pandemic: parents experiences and perspectives. Educ Inf Technol 26:6699–6718. https://doi.org/10.1007/s10639-021-10520-4

Pregowska A, Masztalerz K, Garlinska M, Osial M (2021) A worldwide journey through distance education: from the post office to virtual, augmented and mixed realities, and education during the COVID-19 pandemic. Educ Sci 11(3). https://doi.org/10.3390/educsci11030118

Rodríguez-Muñiz LJ, Burón D, Aguilar-González Á, Muñiz-Rodríguez L (2021) Secondary mathematics teachers' perception of their readiness for emergency remote teaching during the COVID-19 pandemic: a case study. Educ Sci 11(5). https://doi.org/10.3390/educsci11050228

Sang-Yeon K, Graham SS, Ahn S, Olson MK, Card DJ, Kessler MM, DeVasto DM, Roberts LR, Bubacy FA (2016) Correcting biased Cohen's Kappa in NVivo. Commun Methods Meas 10(4): 217–232. https://doi.org/10.1080/19312458.2016.1227772

Sokal L, Trudel LE, Babb J (2020) Canadian teachers' attitudes toward change, efficacy, and burnout during the COVID-19 pandemic. Int J Educ Res Open 1:100016. https://doi.org/10.1016/j.ijedro.2020.100016

Stacey K, Wiliam D (2013) Technology and assessment in mathematics. In: Clements MA, Bishop A, Keitel C, Kilpatrick J, Leung F (eds) Third international handbook of mathematics education. Springer, New York, pp 721–751

Statista (2020) COVID-19 has forced 1.4 billion students to stay home. Retrieved from https://www.statista.com/chart/21225/countries-with-country-wide-or-localized-school-closures/

Tan C (2020) The impact of COVID-19 on student motivation, community of inquiry and learning performance. Asian Educ Dev Stud 10(2):308–321. https://doi.org/10.1108/AEDS-05-2020-0084

Thurm D, Vandervieren E, Moons F, Drijvers P, Barzel B, Klinger M, van der Ree H, Doorman M (2023). Distance mathematics education in Flanders, Germany, and the Netherlands during COVID 19 lockdown – the student perspective. ZDM – Mathematics Education 55:79–93. https://doi.org/10.1007/s11858-022-01409-8