Changes in classroom assessment practices during emergency remote teaching due to COVID-19

Ernesto Panadero [1 & 2]; Juan Fraile [3]; Leire Pinedo [1]; Carlos Rodríguez-Hernández [4];

Fernando Díez [1]

[1] Universidad de Deusto, Spain

[2] IKERBASQUE Basque Foundation for Science, Spain

[3] Universidad Francisco de Vitoria, Spain

[4] Institute for the Future of Education, Tecnológico de Monterrey, Mexico

Correspondence concerning this article should be addressed to Juan Fraile, Universidad Francisco de Vitoria. Carretera Pozuelo-Majadahonda km. 1,8. 28223 – Pozuelo de Alarcón, Madrid. Email: juan.fraile@ufv.es

Recommended citation:

The authors declare to not having any conflict of interest regarding this manuscript.

Abstract

This study explores the effects of the shift to emergency remote teaching on assessment practices due to COVID-19 lockdown. A total of 936 Spanish teachers from all educational levels ranging from early childhood to university participated in this nationwide survey. Four aspects were explored: (1) changes in the use of assessment instruments (i.e. exams); (2) changes in assessment criteria, standards and grading; (3) changes in the delivery of feedback and use of rubrics; and (4) changes in students' involvement in assessment (i.e. self- and peer assessment). In general, results are mixed, with some areas undergoing certain changes with the aim of adapting to the new situation (e.g. primary education teachers lowering their grading standards), whereas many other assessment practices have remained similar, especially among higher education teachers. Unfortunately, some of the assessment practices have worsened, such as students' involvement in assessment which has decreased.

Keywords: *Assessment practices, emergency remote teaching, COVID-19, feedback, rubrics, self-assessment, peer assessment, online teaching.*

Changes in classroom assessment practices during emergency remote teaching due to COVID-19

Classroom assessment is a crucial element in educational systems, as the way in which assessment is implemented influences instructional settings (Wiliam, 2011). Students deploy learning strategies in accordance with different assessment practices put in place by the teacher, ultimately shaping their learning (Balloo et al., 2018; Wiliam & Thompson, 2007). As an example, students that do not receive any performance feedback show less academic achievement than students who only receive grades as feedback which are outperformed by students receiving comments as feedback (Koenka et al., 2021). This shows how assessment changes such as grading or not, commenting or not, have a significant impact on students' learning.

The COVID-19 pandemic forced governments to instate lockdowns in most countries. These lockdowns interrupted normal classroom settings, necessitating the implementation of online teaching in a matter of days. This instructional situation has been termed 'emergency remote teaching' (Hodges et al., 2020) and has compelled changes to assessment practices among other aspects (Bozkurt et al., 2020; Rapanta et al., 2020; Tejedor, et al., 2021). Examples of assessment changes range from challenges in delivering feedback to students to a lack of control during high-stakes examinations (e.g. for university entrance) as well as regular classroom exams (Gamage et al., 2020; Rapanta et al., 2020). However, new opportunities may also have emerged, including automatic feedback or easier delivery of grades with the potential to make assessment more sustainable (Dawson et al., 2018). For these reasons, in this study we explored how emergency remote teaching has altered assessment practices.

Exploring the effects of emergency remote teaching on assessment practices

There is today a bulk of evidence regarding the importance of assessment practices on students' performance and learning. In their systematic review of meta-analyses, Schneider and Preckel (2017) concluded that assessment practices have medium to large effects on educational achievement among higher education students. Assessment practices such as feedback (e.g. Hattie & Timperley, 2007), self-assessment (e.g. Brown & Harris, 2013), peer assessment (e.g. Double et al., 2020), exams (e.g. Yan et al., 2021) or the use of rubrics (Brookhart & Chen, 2015) have been shown to have a positive impact on students' performance. Therefore, it is relevant to explore whether emergency remote teaching has affected how some of these practices are implemented. Next, we present the main variables explored in this study.

A first determinant of assessment is the type of instruments teachers implement in their classrooms. A vast array of assessment instruments is used in our classrooms, from popular ones such as exams or assignments to less common examples like one-sentence summaries, student-generated test questions or minute papers (Angelo & Cross, 1993; Brookhart & Nitko, 2015). Importantly, the use of different assessment instruments requires that students activate distinct learning strategies. In this study we explored whether teachers have changed the frequency of use of three popular assessment instruments: exams, individual assignments and group assignments.

For decades, there has been a tension between formative and summative uses of assessment information (Wiliam, 2011). In general, the tendency is to consider formative uses as better practices to increase students' learning (Black & Wiliam, 1998), although summative practices can also have positive effects (e.g. Koenka et al., 2021). Importantly, in most formal educational contexts around the world, grades are compulsory, so there is always a summative side present (Brookhart et al., 2016).

Therefore, it may be more a matter of combining the two than secluding summative uses. Formative assessment ensures that students and teachers improve their learning and instruction, while summative assessment ensures that grades are anchored in meaningful criteria and standards, as requested by educational systems. Regarding the summative side, in this study we explored if the emergency remote teaching has changed three crucial assessment variables: assessment criteria, standards and grading (Brookhart et al., 2016; Lipnevich et al., 2021; Popham, 1994).

Additionally, we wanted to explored changes on two vastly implemented assessment practices: feedback and rubrics. Feedback is crucial for students' learning and its impact has been comprehensively examined (e.g. Hattie & Timperley, 2007; Lipnevich & Smith, 2018). Given that students greatly depend on teachers' feedback to know whether they have performed an activity correctly, changes in the delivery method and the quality of feedback may have a considerable impact on learning, thus the need to analyze the effects of emergency remote teaching. Rubrics are one instrument used to deliver feedback that have gained a major presence (Dawson, 2017), affecting academic achievement, self-regulated learning, self-efficacy and more (e.g. Brookhart & Chen, 2015). As rubrics have become commonplace across educational levels and countries, it is interesting to explore whether their implementation has changed because of online teaching.

Finally, a significant factor in the implementation of formative assessment is the involvement of students, which is achieved via self- and peer assessment (Black & Wiliam, 1998). Importantly, self- and peer assessment represent the ultimate goal of formative assessment: to potentiate students as the creators of their own feedback (Andrade, 2018; Panadero et al., 2019). Indeed, both have a proven influence on students' achievement (Brown & Harris, 2013; Double et al., 2020). We hypothesized

that students' involvement in assessment during emergency remote teaching might have followed two scenarios. First, emergency remote teaching implied a greater workload for teachers, as they had to become accustomed to the new context and experience screen fatigue (Bozkurt et al., 2020; Jung et al., 2021). Furthermore, as the instructional setting moved further away from the classroom, teachers had no control over what was happening in their students' learning environments (Bozkurt et al., 2020; Hebebci et al., 2020). Therefore, they might have given their students greater responsibility in their assessment. Second, as self- and peer assessment can increase teachers' workload (Panadero & Brown, 2017; Panadero et al., 2014), given the above circumstances, they might have decided against continuing with students' involvement in assessment.

Aim and research questions

The aim of this study was to analyse how emergency remote teaching altered assessment practices in the classroom. We explored four research questions (RQ):

RQ1. Have teachers changed their assessment instruments?

RQ2. Have teachers changed their assessment criteria, standards and grading?

RQ3. Have teachers changed their use of formative feedback and rubrics?

RQ4. Has students' involvement in assessment (i.e. self- and peer assessment) changed?

Method

Participants

A total of 936 Spanish teachers participated in this study, representing early childhood education -3 to 6 years in Spain- (n=64; 6.8%), primary education -6 to 12 years- (n=207; 22.1%), secondary education -12 to 18 years- (n=337; 36%), vocational education -from 15 years- (n=85; 9.1%), higher education -from 18 years- (n=192; 20.5%) and other educational contexts (n=51; 5.4%). Regarding the gender distribution, 641 (68.5%) were female participants. A total of 798 (85.3%) participant teachers

worked in public institutions, 90 (9.6%) in state-subsidised institutions and 48 (5.1%) in private institutions. The participants' mean age was 44.8 years ($SD = 10.88$) and their mean teaching experience was 15.8 years ($SD = 10.66$). In terms of qualifications, all of the participants had a university degree, 160 (17.1%) teachers had a master's degree and 190 (20.3%) had a PhD. In terms of geographical distribution, the 17 autonomous communities of Spain were all represented.

We used a convenience sample based both on the authors' direct distribution of the survey and the participants' own active dissemination to their colleagues. Participation was voluntary, and those participants that gave us their email received the survey results and the publications based on their data.

Instrument

The self-report survey included a total of 91 questions organized in blocks. In the first block, teachers were asked about their demographic and relevant personal information: gender, age, location, educational level, school type (public, state-subsidised or private), qualifications, years of teaching experience and whether their teaching duties had ceased due to COVID-19. In the second block, we asked about their teaching subject, the nature of the subject theoretical vs. practical, the access to technological equipment and their students' socioeconomic profile. In the third block, we asked about instructional aspect (e.g. what is your opinion about online teaching) and assessment changes (e.g. scoring weights, contents, criteria, assessment evidence/instruments and formative assessment practices). Finally, in the fourth block we asked about teachers' well-being, motivation and emotions with regards to the confinement and emergency remote teaching; this data is not presented in this study.

COVID-19 in Spain

Educational institutions at all levels were progressively closed between Monday 9 March 2020 (starting in Madrid and the Basque Country) and Monday 16 March ('COVID-19 pandemic in Spain', 2021). The Spanish government activated the state of alarm on 13 March 2020 and the following day an absolute lockdown was imposed on the population that was ordered to stay home. For the rest of the 2019/20 academic year, from March to June, education at all levels changed to online. There are three studies that offer insights about the implementation of the emergency remote teaching in the Spanish context. In higher education, where 85% of students are enrolled in face-to-face programmes, teaching was moved to fully online with some universities opting for synchronous lessons (Iglesias-Prada et al, 2021) while others universities interpreted the national emergency regulation with more flexibility allowing their teachers to opt for asynchronous or intervallic face-to-face (e.g. one week half of the group face-to-face while the other half online, changing the next week) (Area-Moreira et al., 2021). In Secondary and Primary education, from March to June of 2019/20 teaching was moved to fully online that, when resources allowed for it, was synchronous but a very large percentage had to be conducted asynchronously (Moliner et al., 2021).

Procedure

The survey was administered in April 2020. We decided to disseminate then because the national and autonomous educational governments had implemented the first measures and teachers had already experienced emergency remote teaching for some weeks. At that point, it was public knowledge that the rest of the academic year would take place exclusively online at the vast majority of educational institutions.

The survey was created on an online platform and distributed via email, text message and social media (e.g. Twitter). In addition, using a previous database, it was distributed by email to more than 8,000 teachers from all educational levels.

Data analysis

Several analyses were carried out to answer our RQs. First, a cross-tabulation analysis was conducted for RQs 1, 2, and 3 to identify whether teachers' assessment practices varied because of emergency remote teaching. By doing so, it was possible to summarize the relationship between teachers' educational level and changes in their: use of assessment instruments; use of criteria, standards, and grading; and implementation of formative assessment and rubrics.

Second, both descriptive analyses and paired t-tests were conducted to address RQ4 to determine whether students' involvement in self and peer assessment changed due to emergency remote teaching. Once again, we considered teachers' educational level in the analysis, to gain a better understanding of students' involvement and to identify the advantages and disadvantages of both types of assessments within each educational stage.

<div align="center">

Results

</div>

Before answering the RQs, we report three aspects from our sample in Figures 1a, 1b, and 1c. The majority of participants had not enrolled in courses on educational assessment (Figure 1a), nor courses regarding formative assessment (Figure 1b), withe largest frequencies of no-enrolment in primary and secondary education teachers. Interestingly, and as can be seen in Figure 1c, the vast majority of participants deemed their assessment rather formative, being this result very similar across educational levels.

Figure 1a

Enrolment in courses on Educational Assessment (n=936)
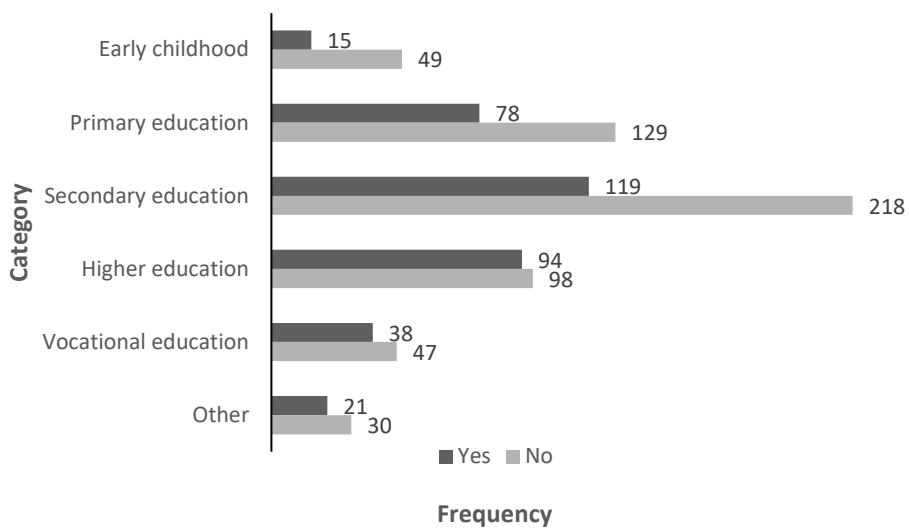
Figure 1b

Enrolment in courses on Formative Assessment (n=936)
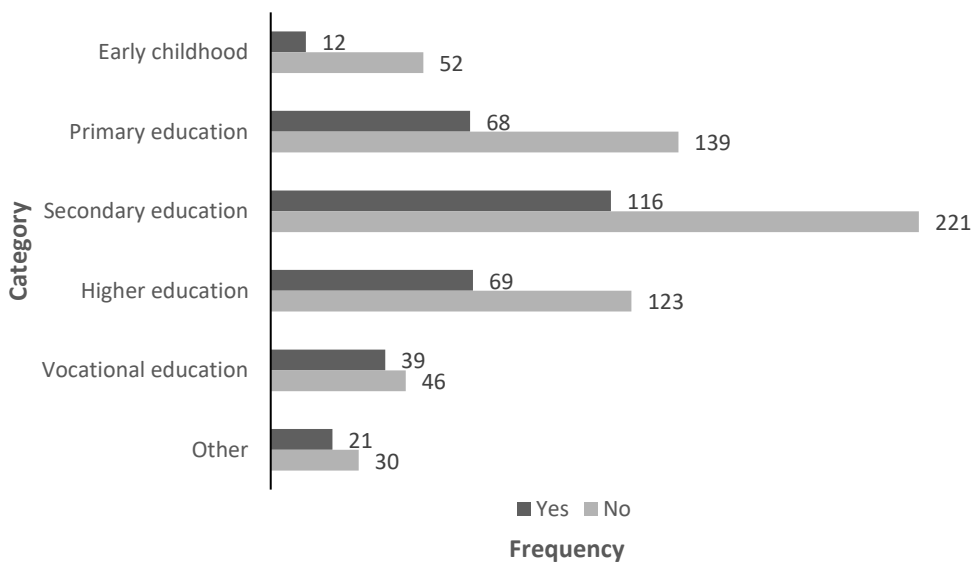
Figure 1c

Teachers' self-reported assessment types: Scale 1 for summative, 9 for formative

(n=936)
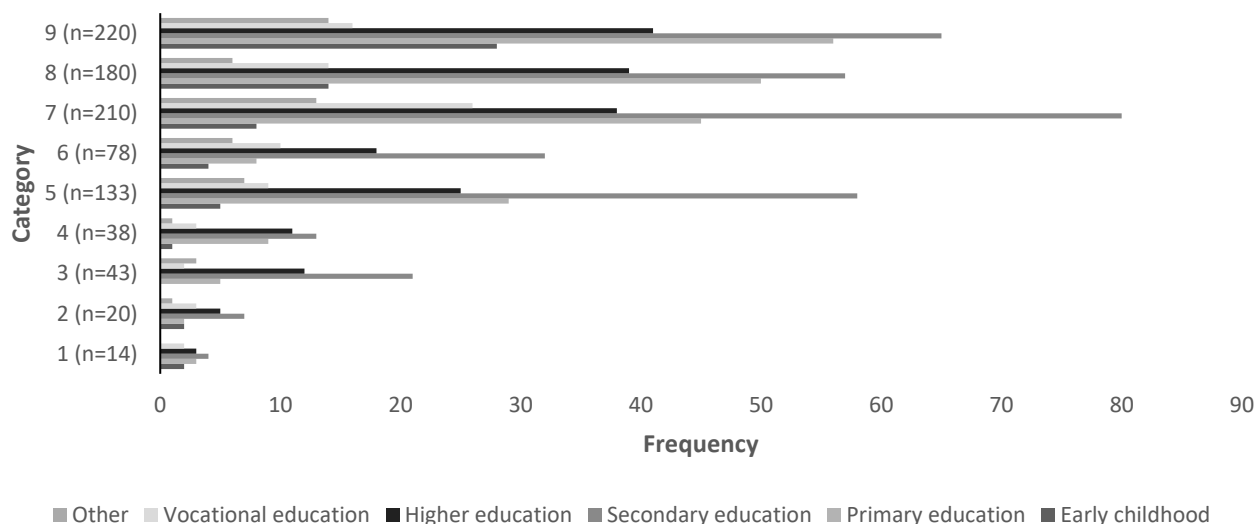
RQ1. Have teachers changed their assessment instruments?

Individual and group assignments followed different patterns (Table 1) in terms of (a) participants who reported an increase (18.3% for individual and 4.2% for group), (b) participants who reported a decrease (20.1% for individual and 34.1% for group), (c) participants who declared the same use (39.1% for individual and 23.3% for group) and participants who reported not using any type of assignment at all (22.5% for individual and 38.5% for group). With respect to educational levels, the largest decrease in the use of individual and group assignments was reported by primary and secondary education teachers (see Table 1). By contrast, higher and vocational education teachers were the most likely to declare using the same types of assignments as before the pandemic.

Considering all educational levels, a similar pattern of changes in the use of exams occurred among the three types: midterms, finals and practical exams (Table 1). The vast majority of participants who had been using exams before emergency remote teaching declared that they decreased their use (35.1% for midterms, 34% for final

exams, 33% for practical exams) or they used them in a similar manner (23.6% for midterms, 25.2% for final exams, 29% for practical exams), while a marginal percentage reported increased use. The other participants reported not using any type of exam (39.6% for midterms, 40.3% for final exams, 32.9% for practical exams). Regarding differences among educational levels, the largest decrease in the use of exams was reported by primary and secondary education teachers. By contrast, higher and vocational education teachers were the most likely to declare maintaining the same use of the three types of exams as before emergency remote teaching.

We then analysed whether the types of questions used in the exams had changed. We examined short-answer questions, essay-answer questions, problem-based questions and multiple-choice questions. As can be seen in Figure 2, except for multiple-choice questions, the other three types were either maintained in number or decreased. Interestingly, essay-answer questions showed the largest decrease. By contrast, there was an increase in the number of multiple-choice questions. Therefore, changes occurred not only in the use of exams during emergency remote teaching, but also in the exams' very design.

Table 1

Changes in the use of different types of exams and assignments, organised by educational level

| Type of exam	Early childhood (n=64)		Primary education (n=207)		Secondary education (n=337)		Higher education (n=192)		Vocational education (n=85)		Other (n=51)		Total (n=936)	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Midterm														
Increase	0	0	1	1	5	2	5	3	3	4	1	2	15	1.6
Decrease	16	25	91	44	158	47	24	13	28	33	12	24	329	35.1
Same	7	11	28	14	64	19	75	39	32	38	15	29	221	23.6
N/A	41	64	87	42	110	33	88	46	22	26	23	45	371	39.6
Final exam														
Increase	0	0	0	0	1	0	1	1	3	4	0	0	5	0.5
Decrease	16	25	78	38	141	42	40	21	31	37	12	24	318	34.0
Same	7	11	25	12	68	20	93	48	30	35	13	26	236	25.2
N/A	41	64	104	50	127	38	58	30	21	25	26	51	377	40.3
Practical exam														
Increase	1	2	13	6	18	5	4	2	8	9	4	8	48	5.1
Decrease	18	28	76	37	131	39	36	19	34	40	14	28	309	33.0
Same	9	14	53	26	94	28	70	37	22	26	23	45	271	29.0
N/A	36	56	65	31	94	28	82	43	21	25	10	20	308	32.9
Individual assignment														
Increase	5	8	28	14	87	29	24	13	23	27	4	8	171	18.3
Decrease	27	42	66	32	68	20	10	5	10	12	7	14	188	20.1
Same	9	14	72	35	130	39	91	47	40	47	24	47	366	39.1
N/A	23	36	41	20	52	15	67	35	12	14	16	31	211	22.5
Group assignment														
Increase	0	0	3	1	21	6	8	4	6	7	1	2	39	4.2
Decrease	24	38	97	47	146	43	17	9	21	25	14	28	319	34.1
Same	7	11	21	10	62	18	85	44	31	37	12	24	218	23.3
N/A	33	52	86	42	108	32	82	43	27	32	24	47	360	38.5

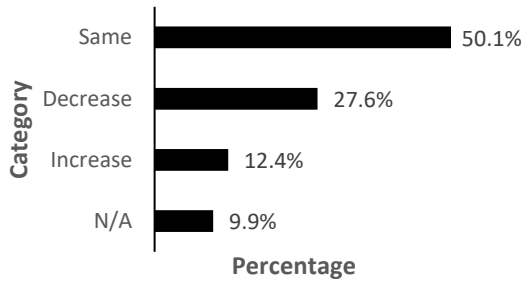**Figure 2**

Changes in evaluation design:

Short-answer questions (n=612)
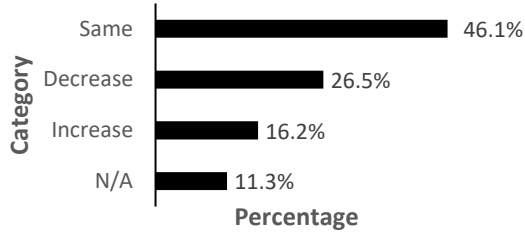
Essay-answer questions (n=590)
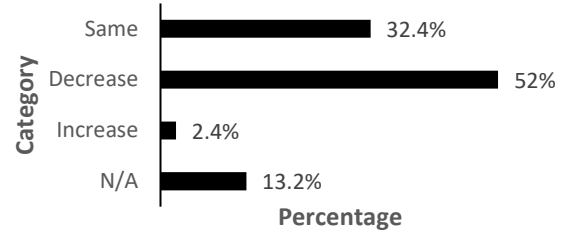
Problem-based questions (n=595)

Multiple-choice questions (n=599)
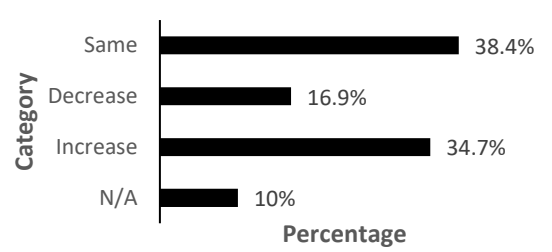
RQ2. Have teachers changed their assessment criteria, standards and grading?

First, we identified changes in the use of assessment criteria and standards (Table 2). A decrease in evaluation criteria and standards was reported by 40.7% of the participants, 31% reported maintaining similar ones, only 0.3% of the participants declared using tougher evaluation criteria and standards for assessing their students and 20.9% reported uncertainty as to whether they had changed or not their evaluation criteria and standards. Notice that 7.1% of the participants provided an answer that was classified as 'Other' (see Table 2). In terms of educational levels, the largest decreases in evaluation criteria and standards were reported by early childhood (42%), primary (59%) and secondary teachers (45%). By contrast, the majority of higher education teachers (72%) declared keeping the same evaluation criteria and standards as before the pandemic.

Second, as regards grading, 54.7% of the participants declared having reduced their grading standards (e.g. a piece of work scored as 6 out of 10 before, would turn into a 7 during ERT), 19.2% reported maintaining the same scoring standards, 19.2% declared being unsure as to whether they had changed their grading and only 0.4% of the participants reported being tougher when assigning grades. Notice that 6.4% of the participants provided an answer that was classified as 'Other'. Primary (68.6%) and secondary (63.8%) teachers declared the most flexibility when assigning grades. By contrast, higher education teachers (59.9%) were the most likely to declare an intention to maintain the same grading process.

Table 2

Changes in the implementation of assessment criteria/standards and grading, organised by educational level

	Early childhood (n=64)		Primary education (n=207)		Secondary education (n=337)		Higher education (n=192)		Vocational education (n=85)		Other (n=51)		Total (n=936)	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Use of evaluation criteria/standards														
Unsure	21	33	44	21	86	26	3	2	25	29	17	33	196	20.9
Tougher criteria	0	0	0	0	1	0	2	1	0	0	0	0	3	0.3
Decreased use	27	42	122	59	151	45	42	22	25	29	14	27	381	40.7
Other	6	9	11	5	30	9	7	4	8	9	4	8	66	7.1
Same	10	16	30	14	69	20	138	72	27	32	16	31	290	31.0
Grading														
Unsure	26	40.6	40	19.3	71	21.1	7	3.6	20	23.5	16	31.4	180	19.2
More flexible	30	46.9	142	68.6	215	63.8	57	29.7	49	57.6	19	37.3	512	54.7
More demanding	0	0,0	1	0,5	2	0.6	1	0.5	0	0.0	0	0.0	4	0.4
Same	5	7.8	12	5.8	25	7.4	115	59.9	12	14.1	11	21.6	180	19.2
Other	3	4.7	12	5.8	24	7.1	12	6.3	4	4.7	5	9.8	60	6.4

RQ3. Have teachers changed their use of formative feedback and rubrics?

Firstly, we asked the participants how frequently they gave feedback to their students before the emergency remote teaching (Figure 3). In regular, face-to-face classroom settings, 49% claimed to give feedback daily, 35% weekly, 8% more than once a month, 6% another type of frequency and only 1% quarterly/biannually. Regarding educational levels, teachers coming from primary and secondary education levels reported the most providing feedback to their students on daily and weekly basis.

Changes in the formative use of feedback and rubrics are summarised in Table 3. Regarding feedback frequency and content, 38.6% of the participants claimed to maintain the same as before, 32.2% reported providing less feedback, while 29.3% declared providing more. In terms of educational levels, the largest increase in feedback was reported by secondary teachers (35%), while the largest decrease was reported by early childhood teachers (62.5%), followed by primary teachers (46.9%). Vocational (42.4%) and higher education teachers (53.6%) were the most likely to maintain the same feedback frequency and content.

Regarding the use of rubrics, 15.4% of the participants claimed to stop using rubrics, 6.4% began to use them, and 45% maintained the same use. Remarkably, 33.2% of the participants declared that they had not used rubrics in the past and they were not going to use them now. The largest decrease in the use of rubrics was reported by primary teachers (28%), whereas secondary teachers were the most likely to maintain the same use as before the pandemic (56.4%). The main uses of rubrics declared by the participants are depicted in Figure 4 and include fairness/objectivity for evaluation and grading (25.4%), students' access to assessment processes and criteria (15.3%) and scaffolding students' learning and self-regulation (8%).

Table 3

Changes in feedback and rubrics use, organised by educational level

	Early childhood (n=64)		Primary education (n=207)		Secondary education (n=337)		Higher education (n=192)		Vocational education (n=85)		Other (n=51)		Total (n=936)	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Feedback frequency and content														
Increase	8	13	55	27	118	35	57	30	26	31	10	20	274	29.3
Decrease	40	63	97	47	97	29	32	17	23	27	12	24	301	32.2
Same	16	25	55	27	122	36	103	54	36	42	29	57	361	38.6
Rubrics use														
Used them before, but not now	15	23	58	28	55	16	6	3	7	8	3	6	144	15.4
Used them before and now	12	19	80	39	190	56	75	39	42	49	22	43	421	45.0
Did not use them before, but uses them now	2	3	10	5	14	4	22	12	10	12	2	4	60	6.4
Did not use them before nor now	35	55	59	29	78	23	89	46	26	31	24	47	311	33.2

Figure 3

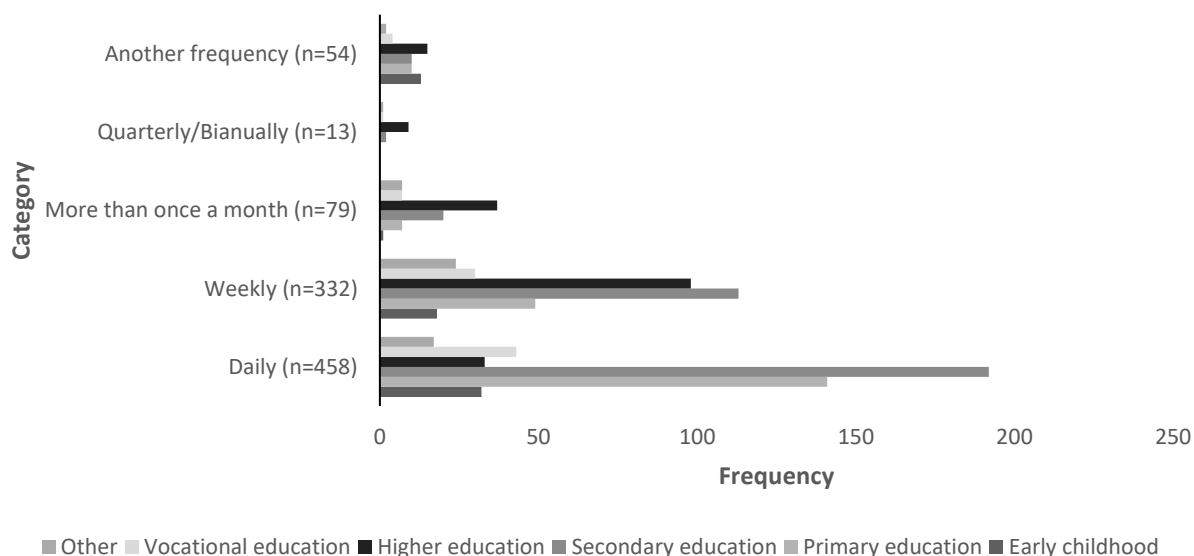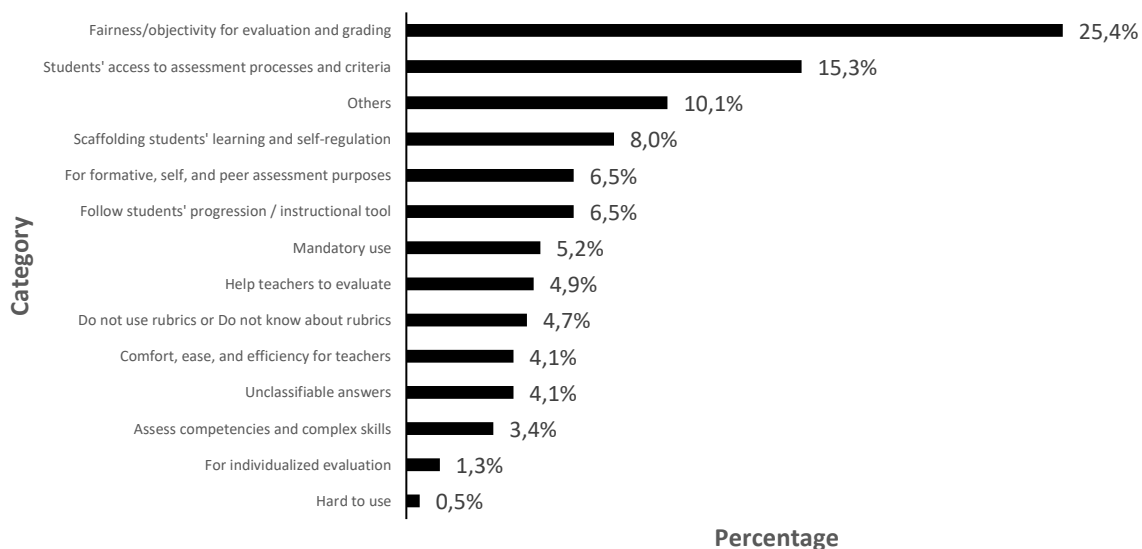*Teachers' feedback frequency (n=936)*

Figure 4

Main uses of rubrics declared by teachers (n=386)

RQ4. Has students' involvement in assessment (i.e. self- and peer assessment) changed?

Self-assessment

When we asked the teachers about their previous experiences with self-assessment in the face-to-face context, 68% reported positive experiences, 30% neutral and 2% negative (Appendix 1). In terms of educational levels, primary education teachers claimed to have the most positive experiences, whereas secondary education and early childhood teachers' experiences were the most negative. Sixty-two percent of the participants reported that the reliability of self-assessment is its main limitation, especially according to secondary teachers (but much less so according to early childhood teachers). By contrast, the main advantages reported were that self-assessment fosters students' awareness about their learning (46%), enables them to identify and correct mistakes (26%) and helps them to learn (22%). Importantly, saving teachers time was not regarded as an advantage of self-assessment, with only 4% mentioning it. Primary and secondary teachers reported the most advantages of self-assessment.

During emergency remote teaching, the participants declared using self-assessment as a pedagogical and learning strategy less ($M = 4.55$, $SD = 2.7$) than they had before ($M = 5.12$, $SD = 2.59$), t (935) = -7.774, $p < .001$. This decrease was found to be significant for all educational levels except vocational education. Considering that the majority of teachers reported positive previous experience with self-assessment, this finding suggests that the special situation had a negative effect. Several challenges of self-assessment that are specific to the remote context were reported by the participants: while 38% did not identify any challenges, concerns pertaining to reliability and objectivity (7.5%) and the difficulty of knowing how students self-assess (6.7%) were

the two main challenges identified. Primary teachers were the most likely to report concerns about reliability and objectivity, while secondary teachers were the most likely to report difficulties in knowing how students self-assess.

Appendix 1 displays several advantages of self-assessment that were identified as being specific to the remote context. While 53.4% of the participants declared that they did not identify any advantage, promoting students' autonomy, responsibility and sincerity/honesty (11.2%) and fostering students' learning (8.3%) were the two main advantages reported. The former advantage was mostly reported by secondary teachers, while the latter was most commonly reported by primary teachers.

Peer assessment.

When we asked the teachers about their previous experiences with peer-assessment in the face-to-face context, 72.4% reported positive experiences, 24.6% neutral and 2.9% negative (Appendix 1). These frequencies were slightly more positive than those for self-assessment. Early childhood were the ones with the lowest response rate (29%) while secondary ones were the highest (52%) which could be an indication of chances of applying this strategy as in early childhood might be more challenging.

Regarding the peer assessment modality used, 36.8% of the teachers allowed communication between the assessor and the assessee, while anonymous modes were less commonly used. The use of either group or individual feedback was most frequently reported by primary and secondary teachers and least by early childhood teachers.

The teachers reported the following as the main challenges of peer assessment: students do not trust their peers' grading (33%), reliability (27%) and creation of tension within the group (22%). These challenges were mostly reported by secondary teachers. The main advantages of peer assessment reported were that it fosters students'

awareness about their learning (32.3%), promotes teamwork (21.5%) and enables students to learn through its use (20.5%). Again, these advantages were primarily reported by secondary teachers.

The participants declared using peer assessment as a pedagogical and learning strategy less during the emergency remote teaching ($M = 1.83$, $SD = 1.82$) than before ($M = 3.56$, $SD = 2.69$), t (935) = -22.276, $p < .001$. It is noteworthy that this decrease was significant at all the educational levels. Among the participants who claimed to use peer assessment, 37.1% declared that they have not altered how they implement it, while 5.6% declared that they now organise groups of students. Higher education teachers were the most likely to report changes.

The teachers identified several challenges of peer assessment that they deemed specific to the remote context. While 38% of the participants declared not to identify any problem, technical challenges (7.6%) and the difficulty of creating groups and managing from a distance (7%) were the two main challenges reported, mostly by secondary teachers.

Appendix 1 displays several advantages of peer assessment that the participants deemed specific to the remote context. While 67.4% of the participants declared not to identify any advantage, fostering students' learning (3.1%) and promoting students' autonomy, responsibility and honesty (1.6%) were the two main advantages reported, again most often by secondary teachers.

Discussion

We aimed to examine how emergency remote teaching changed assessment practices in the classroom. Even though our participants claimed to have limited experience in formative assessment, this was the most common strategy. We will next

discuss the results of our four RQs followed by a section in which we expand on the differences between educational levels.

Changes in Assessment Practices

Our RQ1 results showed that there were changes in how teachers used assessment instruments. Group assignments became less frequent, while exams contained more multiple-choice questions and fewer essay questions. These findings align with other recent studies on the effects of the pandemic on assessment practices (Almossa & Alzaharani, 2022; Bartolic et al., 2021; Senel & Senel, 2021). Because exams were required for accountability purposes, educational institutions (universities especially) conducted and invigilated them online using a variety of modalities (e.g., via platforms such as Zoom or proctoring). Consequently, and notwithstanding the ethical issues entailed, some studies have shown that there was no variability in grading compared with paper-based invigilated examinations (e.g., Linden & Gonzalez, 2021), while others found an increase in the use of contract cheating services (Hill et al., 2021; Montenegro-Rueda et al., 2021).

The results from RQ2 showed that teachers became more flexible in their evaluation strategies, lowering their assessment criteria and standards and amending grading procedures. Again, these findings accorded with other similar studies (Bartolic et al., 2021; Montenegro-Rueda et al., 2021). Reasons for these adjustments included reduced instructional contact; setting up online classes was difficult, and the participants suggested that they wanted to alleviate screen time for students and help them cope with complex living situations such as domiciliary lockdowns.

The results from RQ3 were mixed. The participants reported a high frequency of feedback delivery (daily or weekly). One third reported providing less feedback, while one third reported providing more. As with Jiang and Yu's (2021) study, the shift to

online teaching increased the participants' awareness of the importance of feedback, including the willingness to give it, and the need to learn how to use new technologies to do so. The delivery of feedback increased in secondary education and markedly decreased in primary education and early childhood, perhaps because the students were more dependent on teachers in the latter two settings (Gamage et al., 2020) and direct observation was needed to provide feedback. This was difficult to achieve online.

Two conclusions were drawn from the RQ3 results. First, the participants claimed that before emergency remote teaching they used rubrics mainly to make evaluations fairer or more objective -mostly a summative purpose- and to inform their students about assessment processes and criteria -a transparency purpose-. This confirmed previous research on the use of rubrics (Jonsson & Svingby, 2007). The use of rubrics for formative actions (e.g., scaffoldings for learning) was much less frequently reported than it was in the pre-pandemic period (Brookhart, 2018). Second, we learnt that the use of rubrics remained for the most part unaffected by emergency remote teaching. We hypothesised that this may have been because (a) the teachers believed in the educational value of rubrics even, or especially, in challenging situations; (b) rubrics were created in a digital form (e.g., Word and pdf files) so it was easy to distribute them amongst the students; (c) the students may have requested for the use of rubrics to be continued because they clarified learning goals, standards, and so on. Our results confirmed that rubrics were here to stay (Dawson, 2017) even in complex educational situations. Unfortunately, our data showed that the participants were still not using the full potential of rubrics for formative purposes despite the capacity of these tools to enhance teaching and learning (Andrade, 2005). We must therefore continue to train teachers on how to implement rubrics for formative and summative purposes (Brookhart, 2013).

In addressing RQ4, we explored students' involvement in self-assessment (Andrade, 2018) and peer assessment (Topping, 1998). We found that emergency remote teaching had a clear negative impact on students' involvement in both cases. This aligned with the findings of other studies (Senel & Senel, 2021). Despite their positive experience of face-to-face contact (70%), 62% of the teachers identified specific challenges regarding the use of self- and peer assessment in online teaching. In addition, 53% of the participants did not identify any advantages in online self-assessment; the figure was 83% for peer assessment. For the latter, increased difficulties saw its use fall by half when the teaching context changed. The problems most commonly reported by the participants were related to the online context rather than those indicated in the face-to-face context (Panadero et al., 2014; Panadero & Brown, 2017). Clearly, students' involvement in assessment had diminished, even though the participants were generally positive about its value.

All in all, our results showed the often negative impact of emergency remote teaching on assessment. It is important to consider that adapting to online teaching can take between 6 and 9 months (Hodges et al., 2020), and the sudden shift to emergency remote teaching left little time to adjust the materials and instructional setting, especially for teachers who had previously only delivered face-to-face education. Obviously, assessment practices were affected by this. Because assessment represents the materialisation of learning goals and standards (Brookhart & Nitko, 2015; Laveault & Allal, 2016), central aspects of instruction were affected as shown by our data.

Differences Between Educational Levels

Next, we discuss whether the changes in assessment practices had more distinct effects on the different educational levels than the already existing differences (Brown, 2018). It must be emphasised that studies comparing assessment practices across

educational levels are extremely scarce. This fact strongly limits our ability to compare our findings with previous research, which is vital if stronger conclusions are to be drawn. In other words, without a direct contrast with other educational levels, it is difficult to determine whether a particular level had better adapted to the new context. However, we do not consider this difficulty to compare with previous research a negative characteristic of our study because precisely we attempted to address this shortfall of most research just focusing on one educational level.

Our results systematically showed that the primary education teachers made the most substantial adjustments to their assessment practices, followed by the secondary education teachers and finally by the higher education and vocational teachers. As an example, the primary education teachers (1) lowered their standards, (2) introduced more flexibility into criteria and grading, and (3) reduced the use of rubrics. These results are very different among our participants with older students, especially the higher education ones. Thus, we will next explore four possible reasons for these differences among the teachers working with youngest students and the ones with the oldest.

First, our participants who were used to working with more mature students, especially in higher education, had a longer tradition of employing online systems as instructional scaffolding for their classes (Cerezo et al., 2016). Therefore, having already deployed online systems with more technologically knowledgeable older students, higher education teachers were more able to adapt to the new circumstances with fewer changes and effort. Younger students did not know their way around online platforms, so a significant amount of instructional time was lost by their teachers (Holley & Oliver, 2010).

Second, there are vast differences in the pedagogical training across the different levels in Spain (e.g., Salazar Noguera & McCluskey, 2017), as is the case elsewhere. Whilst primary education teachers in Spain spend at least 4 years training, secondary education teachers can access teaching after a master's programme and higher education teachers can access teaching without any training at all. This has an impact on knowledge, preparedness, and willingness to perform changes in assessment design and implementation (Hill & Eyers, 2016; Laveault & Allal, 2016). Even further, universities worldwide still prefer to evaluate the professional quality of their teachers in terms of research output rather than teaching results (Laveault & Allal, 2016).

Third, it seems logical that the use of assessment instruments has changed less in higher and vocational education for reasons of accountability (Nichols & Harris, 2016). In Spain, teachers working at these levels must accredit students' knowledge for them to pass their subjects. Consequently, many teachers tried to be as faithful as possible to the syllabi and sought to maintain the same assessment and accreditation system despite the complex situation they were facing. Additionally, in higher education, regulatory requirements impede sudden changes in the syllabi because the latter are considered to be instructional contracts with the students (Lipnevich et al., 2021). In sharp contrast, 1 month after the first lockdown, the Spanish Minister of Education issued a new regulation in which early childhood, primary education, secondary education, and first-year pre-university students were permitted to advance to the next stage of their learning even if they had not reached the formerly requisite goals and standards (Torres Menárguez et al., 2020). This reduced the constraints of accountability and the attendance pressures.

And fourth, a crucial factor was the age of the learners: it is only logical that the use of assessment instruments changed more in primary and secondary education,

where students are generally less mature and more dependable, so the absence of face-to-face instruction might have affected them most (Gamage et al., 2020; Putri et al., 2020).

Practical Implications

We divided the practical implications in three categories: teacher training, institutional changes and preparation of students for online education. Regarding the first, it has been long acknowledged that we need to improve our teachers' assessment literacy (Allal, 2013; DeLuca et al., 2016; Laveault, 2016), and emergency remote teaching has confirmed this. First, teachers' assessment professional development should be conceptualised as involving pre-service and in-service processes (DeLuca et al., 2016). It is known that both are complementary and necessary and we should conceptualize teacher training as being important during their university training as when the teachers are already working. Also, we need to approach the professional development of teachers as a multidimensional and challenging task, using a regulation model in which they are proactive agents able to establish their own goals in their assessment education (Laveault, 2016).

Second, we need to provide institutional support and resources to offer greater opportunities for change in teachers' assessment practices. As an example, the institutional regulations need to support formative assessment practices such as self or peer assessment. In addition, teachers that are trying to improve their instructional strategies by using innovative practices, should have their efforts recognized by their institutions.

Lastly, we need to pay special care to the preparation of students for online education. First, students and teachers should have their own electronic devices or, at least, access to so that their educational process is not interrupted. Second, high-quality

online learning may promote the enrolment of international students on programmes they would otherwise have not been able to join because they would have had to physically relocate. Other groups (e.g., hospitalised students, home-schooled students, and students travelling with their parents for various reasons) could also be accommodated. Third, distance grading and certification practices such as proctoring should be made more robust and ethical.

Conclusion

The main conclusion is that the sudden transition to emergency remote teaching affected teachers' assessment practices as they were compelled to alter their instruments, standards, and demands of students especially at the primary and secondary education. When it comes to formative assessment, the results were contradictory: (1) the delivery of feedback remained 'stable' in that the larger percentage of participants reported no changes, though the remainder reported either an increase or decrease; (2) rubrics remained a popular tool, though their formative use could be extended; and (3) self- and peer assessment declined considerably due to the challenges posed by distance learning. In sum, although some teachers have managed to adjust, the situation has, in general, deteriorated. Therefore, our teachers should receive more training on assessment accompanied by institutional changes that enable them to implement more formative activities as recommended above. Hopefully, the extraordinary situation thet emergency remote teaching posed might had served to emphasise the need to take action in both respects.

References

Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice, 20*(1), 20-34. https://doi.org/10.1080/0969594X.2012.736364

Almossa, S. Y., & Alzahrani, S. M. (2022). Assessment practices in Saudi higher education during the COVID-19 pandemic. *Humanities and Social Sciences Communications, 9*(5), 1-9. https://doi.org/10.1057/s41599-021-01025-z

Andrade, H. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching, 53*(1), 27-32. https://doi.org/10.3200/CTCH.53.1.27-31

Andrade, H. (2018). Feedback in the context of self-assessment. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 376-408). Cambridge University Press.

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. Jossey-Bass.

Area-Moreira, M., Bethencourt-Aguilar, A., Martín-Gómez, S., & San Nicolás-Santos, M. B. (2021). Análisis de las políticas de enseñanza universitaria en España en tiempos de Covid-19. La presencialidad adaptada. *Revista de Educación a Distancia (RED), 21*(65).

Balloo, K., Evans, C., Hughes, A., Zhu, X., & Winstone, N. E. (2018). Transparency isn't spoon-feeding: How a transformative approach to the use of explicit assessment criteria can support student self-regulation. *Frontiers in Education, 3*(69). https://doi.org/10.3389/feduc.2018.00069

Bartolic, S. K., Boud, D., Agapito, J., Verpoorten, D., Williams, S., Lutze-Mann, L., Matzat, U., Moreno, M. M., Polly, P., Tai, J., Marsh, H. L., Lin, L., Burgess, J.-L., Habtu, S., Mercedes Rodrigo, M., Roth, M., Heap, T., & Guppy, N. (2021). A multi-institutional assessment of changes in higher education teaching and learning in the face of COVID-19. *Educational Review*, 1-17.

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

Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., Egorov, G., Lambert, S. R., Al-Freih, M., Pete, J., Olcott, D., Jr., Rodes, V., Aranciaga, I., Bali, M., Alvarez, A. V., Jr., Roberts, J., Pazurek, A., Raffaghelli, J. E., Panagiotou, N., de Coëtlogon, P., … Paskevicius, M. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. *Asian Journal of Distance Education*, *15*(1), 1-126. https://doi.org/10.5281/zenodo.3878572

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education, 3*(22). https://doi.org/10.3389/feduc.2018.00022

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343-368. https://doi.org/10.1080/00131911.2014.929565

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, *86*(4), 803-848. https://doi.org/10.3102/0034654316672069

Brookhart, S. M., & Nitko, A. J. (2015). *Educational assessment of students* (7th ed.). Pearson.

Brown, G. T. L. (2018). *Assessment of student achievement*. Routledge.

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan
(Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393).
Sage.

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students'
LMS interaction patterns and their relationship with achievement: A case study
in higher education. *Computers & Education*, *96*, 42-54.
https://doi.org/10.1016/j.compedu.2016.02.006

COVID-19 pandemic in Spain: Revision history. (2021, May 17). In *Wikipedia.*
https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Spain

Dawson, P. (2017). Assessment rubrics: towards clearer and more replicable design,
research and practice. *Assessment & Evaluation in Higher Education*, *42*(3),
347-360. https://doi.org/10.1080/02602938.2015.1111294

Dawson, P., Henderson M., Ryan, T., Mahony, P., Boud, D., Phillips, M., & Molloy, E.
(2018). Technology and feedback design. In J. M. Spector, B. B. Lockee, & M.
D. Childress (Eds), *Learning, Design and Technology: An International
Compendium of Theory, Research, Practice and Policy* (pp. 1-45). Springer.
https://doi.org/ 10.1007/978-3-319-17727-4_124-1

DeLuca, C., Lapointe-Mcewan, D., & Luhanga, U. (2016). Teacher assessment literacy:
A review of international standards and measures. *Educational Assessment,
Evaluation and Accountability*, *28*(3), 251-272. https://doi.org/10.1007/s11092-
015-9233-6

DeLuca, C., Valiquette, A., & Klinger, D. A. (2016). Implementing assessment for
learning in Canada: The challenge of teacher professional development. In D.
Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of
implementation* (pp. 145-160). Springer.

Double, K.S., McGrane, J.A. & Hopfenbeck, T.N. (2020). The impact of peer
assessment on academic performance: A meta-analysis of control group
studies. *Educational Psychology Review, 32*(2)**,** 481–509.
https://doi.org/10.1007/s10648-019-09510-3

Gamage, K. A., Silva, E. K. D., & Gunawardhana, N. (2020). Online delivery and
assessment during COVID-19: Safeguarding academic integrity. *Education
Sciences*, *10*(11), 301. http://.doi.org/10.3390/educsci10110301

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational
research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hebebci, M. T., Bertiz, Y., & Alan, S. (2020). Investigation of views of students and
teachers on distance education practices during the coronavirus (COVID-19)
pandemic. *International Journal of Technology in Education and Science
(IJTES)*, *4*(4), 267-282. https://doi.org/10.46328/ijtes.v4i4.113

Hill, M. F., & Eyers, G. E. (2016). Moving from student to teacher: Changing
perspectives about assessment through teacher education. In G. T. L. Brown &
L. R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment*
(pp. 57-76). Routledge.

Hill, G., Mason, J., & Dunn, A. (2021). Contract cheating: an increasing challenge for
global academic community arising from COVID-19. *Research and Practice in*

Technology Enhanced Learning, 16(24), 1-20. https://doi.org/10.1186/s41039-021-00166-8

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020, March 27). *The Difference between Emergency Remote Teaching and Online Learning*. EDUCAUSE Review. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Holley, D., & Oliver, M. (2010). Student engagement and blended learning: Portraits of risk. *Computers & Education*, *54*(3), 693-700. https://doi.org/10.1016/j.compedu.2009.08.035

Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., & Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior*, *119*, 106713.

Jiang, L., & Yu, S. (2021). Understanding changes in EFL teachers' feedback practice during COVID-19: Implications for teacher feedback literacy at a time of crisis. *The Asia-Pacific Education Researcher, 30*(6), 509-518. https://doi.org/10.1007/s40299-021-00583-9

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational research review, 2*(2), 130-144. https://doi.org/10.1016/j.edurev.2007.05.002

Jung, J., Horta, H., & Postiglione, G. A. (2021, in press). Living in uncertainty: the COVID-19 pandemic and higher education in Hong Kong. *Studies in Higher Education*. https://doi.org/10.1080/03075079.2020.1859685

Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E., & Cooper, H. (2021). A meta-analysis on the impact of grades and comments on academic motivation and achievement: a case for written feedback. *Educational Psychology, 41*(7), 922-947. https://doi.org/10.1080/01443410.2019.1659939

Laveault, D. (2016). Building capacity: Professional development and collaborative learning about assessment. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 131-143). Springer.

Laveault, D., & Allal, L. (2016). *Assessment for Learning: Meeting the challenge of implementation*. Springer

Linden, K., & Gonzalez, P. (2021). Zoom invigilated exams: A protocol for rapid adoption to remote examinations. *British Journal of Educational Technology, 52*(4), 1323-1337. https://doi.org/10.1111/bjet.13109

Lipnevich, A. A., Panadero, E., Gjicali, K., & Fraile, J. (2021). What's on the syllabus? An analysis of assessment criteria in first year courses across US and Spanish universities. *Educational Assessment, Evaluation and Accountability*. https://doi.org/10.1007/s11092-021-09357-9

Lipnevich, A. A., & Smith, J. K. (2018). *The Cambridge Handbook of Instructional Feedback.* Cambridge University Press.

Moliner, L., Lorenzo-Valentin, G., & Alegre, F. (2021). E-Learning during the COVID-19 Pandemic in Spain: A Case Study with High School Mathematics Students. Journal of Education and e-Learning Research, 8(2), 179-184.

Montenegro-Rueda, M., Luque-de la Rosa, A., Sarasola Sánchez-Serrano, J. L., & Fernández-Cerero, J. (2021). Assessment in higher education during the

COVID-19 pandemic: A systematic review. *Sustainability, 13*(19), 10509.

https://doi.org/10.3390/su131910509

Nichols, S. L., & Harris, L. R. (2016). Accountability assessment's effects on teachers

and schools. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and*

Social Conditions in Assessment (pp. 40-56). Routledge.

Panadero, E., & Brown, G. T. L. (2017). Teachers´ reasons for using peer assessment:

positive experience predicts use. *European Journal of Psychology of Education*,

32, 133–156. https://doi.org/10.1007/s10212-015-0282-5

Panadero, E., Brown, G. T. L., & Courtney, M. (2014). Teachers' reasons for using self-

assessment: a survey self-report of Spanish teachers. *Assessment in Education:*

Principles, Policy & Practice, *21*(4), 365–383.

https://doi.org/10.1080/0969594X.2014.919247

Panadero, E., Lipnevich, A. A., & Broadbent, J. (2019). Turning self-assessment into

self-feedback. In D. Boud, M. D. Henderson, R. Ajjawi, & E. Molloy (Eds.),

The Impact of Feedback in Higher Education: Improving Assessment Outcomes

for Learners. Springer.

Popham, W. J. (1994). Educational assessment's lurking lacuna: The measurement of

affect. *Education and Urban Society*, *26*(4), 404-416.

https://doi.org/10.1177/0013124594026004007

Putri, R. S., Purwanto, A., Pramono, R., Asbari, M., Wijayanti, L. M., & Hyun, C. C.

(2020). Impact of the COVID-19 pandemic on online home learning: An

explorative study of primary schools in Indonesia. *International Journal of*

Advanced Science and Technology, *29*(5), 4809-4818.

Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., & Koole, M. (2020). Online university teaching during and after the Covid-19 crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education, 2*(3), 923-045. https://doi.org/10.1007/s42438-020-00155-y

Salazar Noguera, J., & McCluskey, K. (2017). A case study of early career secondary teachers' perceptions of their preparedness for teaching: lessons from Australia and Spain. *Teacher Development*, *21*(1), 101-117. https://doi.org/10.1080/13664530.2016.1204353

Senel, S., & Senel, H. C. (2021). Remote assessment in higher education during COVID-19 pandemic. *International Journal of Assessment Tools in Education*, *8*(2), 181-199. https://doi.org/10.21449/ijate.820140

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565-600. https://doi.org/10.1037/bul0000098

Tejedor, S., Cervi, L., Pérez-Escoda, A., Tusa, F., & Parola, A. (2021). Higher education response in the time of coronavirus: perceptions of teachers and students, and open innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(1), 43. https://doi.org/10.3390/joitmc7010043

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249-276. https://doi.org/10.3102/00346543068003249

Torres Menárguez, A., Zafra, I., Lucas, B., & Urra, S. (2020, April 15). Spanish students to get a passing grade due to coronavirus confinement. *El País.*

https://english.elpais.com/spanish_news/2020-04-15/spanish-students-to-get-a-

passing-grade-due-to-coronavirus-crisis.html

Wiliam, D. (2011). What is assessment for learning?. *Studies in educational*

evaluation, *37*(1), 3-14. http://doi.org/10.1016/j.stueduc.2011.03.001

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it

take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping*

teaching and learning (pp. 53-82). Lawrence Erlbaum Associates.

Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic

review on factors influencing teachers' intentions and implementations

regarding formative assessment. *Assessment in Education: Principles, Policy &*

Practice, 28(3), 228-260. https://doi.org/10.1080/0969594X.2021.1884042

Appendix 1

Use of Self- and Peer-Assessment

	Early childhood		Primary education		Secondary education		Higher education		Vocational education		Other level		Total	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Experience rating of the use of self-assessment in the face-to-face learning	51		184		283		40		75		46		679	
Positive	33	65	150	82	171	60	29	73	51	68	30	65	464	68
Neutral	18	35	32	17	107	38	10	25	22	29	13	28	202	30
Negative	0	0	2	1	5	2	1	3	2	3	3	7	13	2
Main challenges of self-assessment	45		182		329		181		78		42		857	
Reliability	32	71	120	66	212	64	102	56	47	60	22	52	535	62
Other	9	20	37	20	41	12	45	25	14	18	13	31	159	19
Amount of time needed	3	7	12	7	42	13	18	10	8	10	1	2	84	10
Creates tension with the teachers' authority	0	0	9	5	21	6	5	3	5	6	5	12	45	5
Does not foster students' learning	1	2	4	2	13	4	11	6	4	5	1	2	34	4
Main advantages of self-assessment	97		400		594		289		130		84		1594	
Fosters students' awareness about their learning	53	55	184	46	265	45	131	45	59	45	39	46	731	46
Identifies and corrects mistakes	21	22	106	27	164	28	63	22	42	32	24	29	420	26
Students learn through their use	22	23	95	24	131	22	61	21	24	18	18	21	351	22
Saves teachers' time	1	1	13	3	29	5	14	5	5	4	2	2	64	4
Other	0	0	2	1	5	1	20	7	0	0	1	1	28	2
Peer-assessment modality	24		185		254		121		58		38		680	
There was interaction between the assessor and assessee	13	54	89	48	89	35	30	25	14	24	15	39	250	37
Group feedback	5	21	40	22	60	24	27	22	10	17	13	34	155	23
Individual feedback	2	8	31	17	42	17	19	16	15	26	8	21	117	17
The assessee was anonymous	1	4	15	8	31	12	18	15	6	10	1	3	72	11

The assessor was anonymous	0	0	5	3	27	11	12	10	8	14	0	0	52	8
Other	3	13	5	3	5	2	15	12	5	9	1	3	34	5
Changes in peer-assessment	4		20		25		26		7		7		89	
No change	2	50	9	45	8	32	10	38	2	29	2	29	33	37
Do not use it	2	50	7	35	11	44	4	15	3	43	3	43	30	34
Others	0	0	0	0	3	12	3	12	0	0	2	29	8	9
Unclassifiable answers	0	0	2	10	0	0	3	12	1	14	0	0	6	7
Organised by groups now	0	0	1	5	0	0	3	12	1	14	0	0	5	6
Yes	0	0	1	5	2	8	2	8	0	0	0	0	5	6
Had to stop using it	0	0	0	0	1	4	1	4	0	0	0	0	2	2
Experience rating of the use of peer-assessment in the face-to-face learning	19		122		175		41		89		29		475	
Positive	15	79	101	83	122	70	30	73	54	61	22	76	344	72
Neutral	4	21	20	16	48	27	9	22	30	34	6	21	117	25
Negative	0	0	1	1	5	3	2	5	5	6	1	3	14	3
Main challenges of peer-assessment	26		202		353		229		65		33		908	
Students do not trust on their peers' grading	3	12	82	41	128	36	57	25	22	34	11	33	303	33
Reliability	14	54	45	22	100	28	58	25	18	28	9	27	244	27
Creates tension within the group	4	15	53	26	73	21	50	22	11	17	6	18	197	22
Other	4	15	9	4	17	5	38	17	6	9	6	18	80	9
Amount of time needed	0	0	6	3	15	4	14	6	6	9	0	0	41	5
Does not foster students' learning	0	0	4	2	15	4	11	5	2	3	1	3	33	4
Creates tension with the teachers' authority	1	4	3	1	5	1	1	0	0	0	0	0	10	1
Main advantages of peer-assessment	44		347		386		265		78		48		1168	
Fosters students' awareness about their learning	15	34	100	29	115	30	103	39	25	32	19	40	377	32
Promotes teamwork	11	25	88	25	89	23	37	14	13	17	13	27	251	21
Students learn through their use	9	20	79	23	87	23	39	15	19	24	7	15	240	21
Identifies and corrects mistakes	8	18	69	20	73	19	43	16	19	24	8	17	220	19
Saves teachers' time	0	0	10	3	19	5	11	4	0	0	0	0	40	3
Other	1	2	1	0	3	1	32	12	2	3	1	2	40	3

Challenges of self-assessment specific to the emergency remote teaching	20		89		131		90		36		21		387	
No	7	35	30	34	46	35	31	34	23	64	10	48	147	38
Others	3	15	12	13	18	14	24	27	2	6	5	24	64	17
Unclassifiable answers	1	5	8	9	16	12	13	14	1	3	1	5	40	10
Reliability and objectivity	1	5	13	15	11	8	1	1	2	6	1	5	29	7
Difficult to know how students self-assess	1	5	3	3	19	15	3	3	0	0	0	0	26	7
Not knowing if another person (e.g. parents) does it	2	10	11	12	1	1	2	2	0	0	1	5	17	4
Do not use it	0	0	4	4	5	4	2	2	5	14	1	5	17	4
Lack of students' participation, attention, and dedication	1	5	1	1	6	5	4	4	1	3	1	5	14	4
Lack of interaction with students and time to comment on the self-assessment	0	0	1	1	2	2	8	9	1	3	0	0	12	3
Lack of immediacy	1	5	1	1	3	2	1	1	0	0	1	5	7	2
Students' lack of confidence	0	0	3	3	2	2	0	0	1	3	0	0	6	2
Not feasible in early childhood and elementary school	3	15	2	2	0	0	0	0	0	0	0	0	5	1
Inflation of scores	0	0	0	0	2	2	1	1	0	0	0	0	3	1
Advantages of self-assessment specific to the emergency remote teaching	16		72		114		23		31		21		277	
No	8	50	34	47	57	50	16	70	23	74	10	48	148	53
Promotes students' autonomy, responsibility, and honesty	3	19	4	6	16	14	3	13	2	6	3	14	31	11
Foster students' learning	1	6	9	13	6	5	1	4	3	10	3	14	23	8
Unclassifiable answers	1	6	4	6	11	10	2	9	0	0	3	14	21	8
Promotes students' motivation and self-regulation	0	0	10	14	6	5	0	0	1	3	1	5	18	6
Others	2	13	7	10	8	7	1	4	0	0	0	0	18	6
Provides information for the teacher	1	6	1	1	6	5	0	0	1	3	0	0	9	3
Reduces teachers' time dedicated to evaluating	0	0	3	4	4	4	0	0	1	3	0	0	8	3
Do not use it	0	0	0	0	0	0	0	0	0	0	1	5	1	0

Challenges of peer-assessment specific to the emergency remote teaching	6		38		44		47		17		6		158	
No	1	17	12	32	15	34	19	40	10	59	3	50	60	38
Do not use it	1	17	8	21	6	14	5	11	3	18	1	17	24	15
Others	1	17	5	13	2	5	9	19	1	6	0	0	18	11
Technical problems	0	0	3	8	5	11	3	6	0	0	1	17	12	8
Unclassifiable answers	0	0	0	0	3	7	7	15	1	6	1	17	12	8
More difficult to manage/create groups	0	0	2	5	7	16	1	2	1	6	0	0	11	7
Students do not have the maturity	3	50	3	8	0	0	0	0	0	0	0	0	6	4
Worsened relationships among peers	0	0	2	5	3	7	0	0	0	0	0	0	5	3
Had to stop using it	0	0	2	5	0	0	2	4	0	0	0	0	4	3
Reliability, objectivity, and lack of honesty	0	0	1	3	2	5	0	0	1	6	0	0	4	3
Yes	0	0	0	0	1	2	1	2	0	0	0	0	2	1
Advantages of peer-assessment specific to the emergency remote teaching	6		32		40		36		11		4		129	
No	3	50	20	63	27	68	27	75	7	64	3	75	87	67
Do not use it	0	0	8	25	3	8	5	14	3	27	1	25	20	16
Others	2	33	1	3	4	10	1	3	1	9	0	0	9	7
Unclassifiable answers	1	17	1	3	1	3	2	6	0	0	0	0	5	4
Fosters students' learning	0	0	2	6	2	5	0	0	0	0	0	0	4	3
Promotes students' autonomy, responsibility, and honesty	0	0	0	0	2	5	0	0	0	0	0	0	2	2
Yes	0	0	0	0	1	3	1	3	0	0	0	0	2	2
Use of self-assessment and peer-assessment	64		207		337		192		85		51		936	
Use of Self-Assessment before	4.79	2.85	6.05	2.26	5.12	2.36	4.15	2.91	5.04	2.53	5.52	2.47	5.12	2.59
Use of Self-Assessment now	3.76	2.88	5.01	2.65	4.64	2.56	3.85	2.87	4.8	2.52	5.23	2.6	4.55	2.7
Use of Peer-Assessment before	2.39	2.01	4.24	2.62	3.77	2.66	2.86	2.77	3.37	2.7	3.78	2.54	3.56	2.69
Use of Peer-Assessment now	1.62	1.49	1.63	1.55	1.81	1.69	2.02	2.23	2.22	2.29	1.64	1.33	1.83	1.82