Using Criteria-Based Assessment Rubrics for Online Marking: Technological and Pedagogical Challenges

Samira El Boudamoussi
ETH Zurich

This paper discusses the challenges of developing and using assessment rubrics for digital marking in higher education. As part of a curriculum review process, the rubrics were designed to provide raters with valid tools to assess higher-order-thinking skills in a taught postgraduate program. With the online/hybrid modes of teaching and learning imposed by the COVID-19 restrictions, this timely solution proved to be very useful in digital settings. However, the development and implementation processes represent numerous challenges to developers and users while issues related to technical limitations of digital platforms are yet to be solved. This work was conducted at the Centre for Environmental Policy (CEP) at Imperial College London, within its Master's programme in Environmental Technology.

Keywords: assessment rubrics, digital assessment, higher education, curriculum transformations, curriculum design, assessment in higher education, marking authentic assessments, grading with rubrics, marking with rubrics, developing rubrics

INTRODUCTION

In the context of wider curriculum transformations in higher education, the shifts of emphasis taking place at the levels of curriculum design and teaching practices have significant implications in terms of assessment and marking, which in turn, help shape student learning experience and the quality of learning (Nordrum et al., 2013). One of these implications is the implementation of authentic assessments combined with the use of assessment rubrics as tools for marking and for improving feedback (Wiggins, 1990; Montgomery, 2002; Hack, 2013; Villarroel, 2018). On one hand, it is argued that the assessment is what determines the actual curriculum (Ramsden, 2003). On the other hand, "*criteria-based assessment has been pointed out as a means to meet current demands on curricula design*" (Nordrum et al., 2013, p.919), especially in the European context of the Bologna process.

Assessment rubrics are described in the literature as tools that are adapted to authentic assessments or assessments that focus on higher order thinking skills, complex thinking, and effective communication (Marzano et al, 1993).

There are two types of rubrics used in marking assessments: Analytic rubrics (Figure 1) and holistic rubrics (Figure 2).

FIGURE 1
ANALYTIC RUBRICS

Criteria	Levels of achievement							
	1	2	3	4	5	6	7	8
C1	Descriptors				Descriptors			
C2	Descriptors				Descriptors			
C3	Descriptors				Descriptors			

In their book "Introduction to Rubrics", Stevens & Levi (2012) refer to analytic rubrics as a grid that sets (1) in the first row the scale or the levels of achievement, (2) in the first column the dimensions of the assignment, which represent a breakdown of the skills/knowledge involved in the assignment, and (3) in the cells of the grid the descriptions of what constitutes each level of performance for each dimension. Those descriptions could be similar to specific feedback statements and can be used as feedback or as a basis for feedback.

FIGURE 2
HOLISTIC RUBRICS

Criteria	Levels of achievement							
	1	2	3	4	5	6	7	8
C1 Brief description	✓							
C2 Brief description				✓				
C3 Brief description								✓

Holistic rubrics (Figure 2) do not include such descriptions but also rely on a set of criteria or dimensions and a grading scale. Even though they are described in the literature as rubrics, holistic rubrics are essentially rating scales and they are generally very useful in assessing creative products such as in arts or in engineering.

Analytic rubrics can be used for grading a large variety of authentic assignments like posters, videos, different types of reports, group discussions and oral presentations. According to Hack (2015), there is evidence of the utility of the analytic rubrics for these types of assessments.

This paper starts by analysing the context of curriculum transformations that has led to the shifts in assessment strategies and how these shifts resulted in transforming the marking schemes into the development and implementation of analytic rubrics as part of a curriculum review process. Then, it will

discuss the challenges, pedagogical and technological, encountered by both the developers and the users of the rubrics. Drawing on practical experience and the lessons learned from this process, the paper concludes with recommendations on how to address some of the challenges met.

CURRICULUM TRANSFORMATIONS AND IMPLICATIONS FOR ASSESSMENT AND MARKING

Curriculum transformations in higher education are triggered in part by societal and economic demands (e.g. lifelong learning; sustainability education; employability and professional skills) as well as by a growing body of research findings on learning, teaching, assessing (Ramsden, 1985; Brown, 1997; Biggs, 2003b; Bloxham & Boyd, 2007; Carless, 2018), and on the relationships between approaches to teaching and learning approaches in higher education (Weiman and Gilbert, 2014; Trigwell and Prosser, 2004; Trigwell et al., 2005; Prosser and Trigwell, 2006).

These transformations are expected to trigger shifts in emphasis at various levels such as:

- Curriculum design
- Teaching methods and teaching practices
- Approaches to learning
- Assessment strategies and marking

Curriculum Design

In terms of curriculum design, it is expected to shift from a content-based curriculum towards a curriculum with more emphasis on skills and where the learning outcomes are not limited to knowledge and understanding but support the development of large sets of analytical and critical thinking skills, skills that support lifelong learning, as well as professional and transferrable skills, especially within Master's taught programmes.

According to the definition of the Bologna Working Group (2005, p.29), Learning Outcomes (LOs) are "*statements of what a learner is expected to know, understand and/or be able to do at the end of a period of learning*" or at the end of a learning process in general. Therefore, LOs can be defined at the level of a course unit, module, or programme level. LOs are also defined in reference to a qualification level within a framework of qualifications at the level of an individual institution, national, or international level. For example, the European Qualifications Framework (EQF) –a common European reference framework which aims at making qualifications more readable and understandable across different countries and systems– defines a total of eight reference levels of qualifications (Cedefop, ND). Out of those eight EQF levels, a Master's qualification corresponds to Level 7.

The requirements of setting the LOs at level 7 for a Master's taught programme imply greater emphasis on developing higher-order thinking skills and less emphasis on memorising facts or applying knowledge. As explained by Resnick (1989, cited by Moseley et al., 2006: p.29), "*It is widely accepted that learners need to become skilled in accessing and using knowledge productively rather than learning factual content as a memory-based exercise*".

With the shifts in curriculum design and in line with the constructive alignment framework (Biggs, 2003a), LOs have gained a crucial role as key elements that determine the teaching methods to be adopted and the assessment strategies to be implemented (Bloxham & Boyd, 2007).

Teaching Methods and Teaching Practices

At the level of teaching, the shifts are from what John Biggs refers to as Level 1 model of teaching (focused on what the student is) towards Level 2 and Level 3 models of teaching focused on what the student does and how it relates to teaching (Biggs, 1999). The purpose of this student-centered model of teaching is to support learning and to develop teaching effectiveness through well stipulated learning outcomes as well as well stipulated levels of understanding, all of which in alignment with the teaching and learning activities and with the assessment (Biggs, 1999). In this constructivist alignment approach to

curriculum design, the role of the teaching teams is to set the environment for students so that they mobilise their knowledge and skills, solve real-world problems, and "*construct meaning through relevant learning activities*" (Biggs, 2003a, p.3).

In addition to teaching approaches, research has identified more factors that may influence student success, such as student agency (Jääskelä et al., 2017) and students' identity and sense of belonging (Smith et al., 2005). Concepts such as self-efficacy and self-regulation (Bandura, 2006) have been also identified as key to improving student learning and student performance (Adie et al., 2018).

The sense of belonging "*recognises students' subjective feelings of relatedness or connectedness to the institution*" (Thomas, 2012, p.12). Various studies insist on the strong association between the sense of belonging and students' success and retention in higher education and its close alignment with the concepts of academic and social engagement (Kuh et al., 2005; Thomas, 2012).

Agency is seen as "*a core component of professionalism*" and "*a way of coping with uncertainty and changes in working life, thus, playing a key role in lifelong learning*" (Su 2011, cited by Jääskelä et al., 2017, p.2061). Studies reviewed by the authors point to the importance of professional identity arguing that its development requires the support of educators and that the role of higher education should focus not only on addressing the knowledge and skills but also on developing students' abilities to engage purposively with the changing and complex world (Jääskelä et al., 2017).

In the new curricula, the sense of belonging, student agency, self-efficacy, and self-regulation have been gaining increasing importance with the student-centered models of teaching and active learning (Smith et al., 2005) as well as the associated forms of formative and summative assessments (Lynam & Cachia, 2018). This has implications not only on what to assess and how to assess but also on the role of assessment itself.

Approaches to Learning

In terms of learning, the shifts are towards students being more engaged with learning and with the content. Learning occurs with and from the peers, while students take the responsibility of their own learning to become autonomous and independent learners (Smith et al., 2005).

Studies on teaching practice inventories conducted by Weiman and Gilbert (2014) confirm that there is sufficient evidence base from the research on the teaching practices that improve learning. Since the first studies on student learning initiated by Marton and Saljo (1976, cited by Trigwell & Prosser, 2004; Biggs, 1999; and Ramsden, 1985), more knowledge has been gained about students' approaches to learning and the relationships between the quality of students' learning outcomes and their learning approaches. Studies conducted on student learning approaches have identified "*three qualitatively different approaches by students to learning, that are now labelled deep, surface and achieving or strategic*" (Trigwell & Prosser, 2004, p.410). Research on approaches to teaching has attempted to find out whether certain perceptions about teaching or certain teaching practices are associated with a higher quality of student outcomes.

Findings on how students learn and how their approaches to learning could be influenced by the way teachers approach their teaching, resulted in the requirement of some form of initial training for higher education lecturers (in the UK and elsewhere) prompting lively debates around the professionalisation of academic staff in relation to learning, teaching and assessment (Bloxham & Boyd, 2007). The role of higher education lecturers is expected to shift from imparting subject knowledge to creating learning contexts and setting up learning environments for students to interact and ask questions while working on learning activities that are appropriately designed to be aligned with the assessment tasks (Biggs, 2003a). This way, learning is also part of the teaching and not something done separately by the student.

According to Lynam & Cachia (2018), there is evidence that learning approaches adopted by students also depend on the type of assessment used (Carless 2007; Raupach et al. 2013 cited by Lynam and Cachia, 2018). It is suggested that formative assessments are more likely to induce a deep learning approach while summative assessments are more associated with the surface approach (Lynam and Cachia, 2018). As Biggs (2003a, p.4) writes referring to the students: "*They'll learn what they think they'll be assessed on, not what's in the curriculum or what's been 'covered' in class*".

Assessment Strategies and Marking

In the context described above, the shifts in terms of curriculum design, teaching methods and learning approaches have great implications for what is assessed, how and why it is assessed. The shifts at the level of assessment strategies are towards a variety of assessments that fulfil different purposes and roles, such as learning-oriented assessment, assessment as learning, assessment for learning, and assessment of learning (Bloxham & Boyd, 2007). More emphasis is being put on assessment as a feedback and learning tool rather than –or at least in addition to–a judgmental tool (Carless, 2007). While feedback should provide students with useful and constructive information about their work and learning progress, efforts to promote feedback literacy among faculty and students have become necessary (Carless & Boud, 2018; Molloy et al., 2020).

By putting the student's learning at the centre of curriculum design and in order for the assessment used to shape students' learning experience and transform their learning approaches, it is expected that new forms of assessing progressively replace traditional or paper-and-pencil exams. For example, more divergent assessments are implemented, which anticipate a range of responses rather than convergent assessments which include questions having wrong or right answers and test the learner's knowledge of facts and procedures (Torrance & Pryor, 2001). Other forms of assessments, such as performance-based assessments-PBA or authentic assessments (Baker et al., 1993; Wiggins, 1990), are considered to be more adapted to assess higher order thinking processes and students' learning outcomes in general (Resnick & Resnick, 1992; Wiggins, 1989, cited by Baker et al., 1993). According to Baker et al. (1993, p.1212), "*Wiggins (1989) saw the class of performance-based assessments that he calls authentic performance-based assessment useful as a curriculum reform strategy*".

Authentic assessments are referred to as tasks or activities designed to mimic the tasks, problems, or situations faced by professionals and experts in the real world (Frey et al., 2012). Therefore, authentic assessments usually use contemporary, ill-structured, and open-ended problems to assess a broad range of skills, knowledge, and competencies, involving complex thinking, information processing, and effective communication (Lynam & Cachia, 2018). Undergraduates and postgraduates are expected to actively engage in classroom activities that are rooted in community, some of those activities need to be finalised, planned or executed without or with minimum intervention from the tutor or lecturer (Newell, 2010; Brundiers et al., 2010). In this kind of settings, students are active learners who work independently to seek information, mobilise knowledge, and integrate skills and knowledge to produce artefacts, portfolios, or reports in the form of group assignments that are then assessed (Hack, 2015).

In order to evaluate students' performance across a wide variety of learning outcomes as those mentioned above by using authentic assessment tasks, it is suggested to use criteria-based rubrics as the most common and adapted way of marking (Hack, 2015). While fulfilling the requirements of transparency and clarity in the assessment process, assessment rubrics are believed to address various issues around the quality of marking and consistency, especially in the case of multiple markers (Hack, 2015).

This paper focuses on the challenges faced by the developers (curriculum review team) and users (markers) of the assessment rubrics implemented as part of the curriculum review process of a taught Master's programme in Environmental Technology in the UK.

CHALLENGES OF ASSESSING AT MASTER'S LEVEL

Within the Learning and Teaching Strategy being implemented since 2017 at Imperial College London (Imperial College London, 2017), the Centre for Environmental Policy (CEP) initiated in 2018 the curriculum review process for the taught Master's programme in Environmental Technology taking into account various shifts in terms of the Graduate Attributes, the adoption of a modular architecture with well-defined learning outcomes for each module, and the shifts in teaching approaches as well as in the assessment strategies. However, this was not without being challenging.

First, all the LOs were set at the right level for a Master's degree, which is level 7 defined by the European Qualifications Framework (Bologna Working Group, 2005). Level 7 learning outcomes raised great concerns among teaching staff about the most effective way to assess and mark them. On one hand,

the programme's LOs covered a wide range of learning categories including knowledge and understanding but also more specific skills and competencies. On the other hand, level 7 LOs generally focus on higher order and complex thinking skills as well as on a broad variety of skills such as information processing, analysis and evaluation skills, and effective oral and written communication skills. While the majority agreed that closed book exams and even open book exams were not fit for purpose to assess this type of LOs, opting for authentic assessments was not completely unanimous.

Second, there was a great awareness among the experienced teaching staff that the existing marking schemes were not adapted to mark such assessments as those proposed in the new curriculum. Those assessments formed an integral part of the teaching and learning process and included, for example, students producing posters, videos, and different types of reports; participating in assessed group discussions; and oral presentations in groups and individually. The shift towards analytic rubrics was based on various literature reviews that acknowledge a growing evidence base for the utility of rubrics to mark authentic assessments, Performance-based assessments, and assignments designed to measure this kind of higher order learning outcomes. According to Katherine Hack, the rubrics also improve inter-rater reliability and thus they are useful for assessing large cohorts and where multiple or inexperienced assessors are employed (Hack, 2013). The MSc in environmental technology enrols cohorts of more than 150 students every year and up to 26 to 30 markers with different levels of experience ranging from null to 20-25 years of experience with teaching and marking.

Third, even though the use of assessment rubrics in higher education has been largely associated with their integration in the virtual learning environments (Hack, 2015), many markers expressed their preferences for using pen and paper to mark the assessments. However, as soon as the redesigned curriculum was deployed in the academic year 2019-2020, the national lockdown came into force boosting the need for digital assessments, online marking, and the remote provision of feedback.

In the following sections, four categories of challenges are described. In addition to the challenges faced by the developers during the process of designing and developing the rubrics, on one hand, and those encountered by the users during the implementation phase, on the other, a distinction is made between two types of challenges: pedagogical and technological.

CHALLENGES TO THE DEVELOPMENT OF ANALYTIC RUBRICS

Pedagogical Challenges to Developers

In the process of developing rubrics for the MSc environmental Technology assignments, the pedagogical challenges faced by the developers were mainly related to the key features of analytic rubrics:

1. Dimensions of the assignment: How to decide on which evaluation criteria to include in a rubric?
2. Levels of achievement: Which marking scale to adopt and how to determine the granularity of the scale?
3. Descriptions: How to make sure that there is a good progression in the formulation of quality descriptions for each criterion at each level of performance and that the descriptions are clearly discernible from one level to the other?

Dimensions of the Assignment

Following the constructive alignment framework, the assessments are meant to assess the intended learning outcomes (ILOs) and the rubrics are tools that enable markers to make judgements on the student's performance in each ILO. So, the first challenge faced by the developers was to make sure that the dimensions of the rubric, reflect the intended learning outcomes of the module assessed. In a sense the use of a criteria-based rubric should enable markers to actually focus their judgments on the ILOs of the module being assessed as the main evaluation criteria. For example, in a poster assignment, markers would be tempted to consider such aspects as the visual impact of the poster, the originality or uniqueness of the design, and any "personal touch" as evaluation criteria. Based on the information collected from the teaching staff, this kind of dimensions, rather subjective and not relating to any of the module's ILOs were

discarded. Instead, the dimensions that relate to the actual components of the assignment and, at the same time, could be mapped to the ILOs were retained as evaluation criteria.

Table 2 shows the steps followed to come up with the evaluation criteria for the assessment rubric for a poster assignment.

TABLE 2
PROCESS OF DECIDING ON THE DIMENSIONS OF A RUBRIC TO ASSESS A POSTER ASSIGNMENT

What do you look for in the poster?	**Module ILOs**	**Evaluation criteria**	**Descriptions in terms of what the student does**
Visual Impact: first impression The poster's design is original (new, unique, surprising or shows a personal touch).	No mapped ILO	No evaluation criteria	No description
Content: information logically organised & coherent, e.g. introduction & conclusion | aims, objectives & results clearly conveyed | appropriately referenced.	ILO3. critically analyse, synthesise and evaluate knowledge relating to environmental governance and management. ILO4. identify trade-offs in problem solutions and understand how to mitigate them.	1) Problem statement	1) States the problem, announces the aim and objectives.
		2) Background analysis	2) Analyses institutional, social, technological and historical backgrounds: Shows critical thinking skills such as evaluation and synthesis.
		3) Policy recommendations	Makes recommendations based on the background analysis, evaluation, and synthesis. Justifies the resources needed, the benefits to society and the pathway to change.
		4) Feasibility	Outlines the practicality, cost-effectiveness and social acceptance for each recommendation. Supports reasoning with arguments based on analysis, evaluation, and synthesis of evidence.
		5) Trade-offs & Mitigation	Describes and discusses the trade-offs for each recommendation. Justifies and evaluates mitigation steps based on evidence.
Design: layout | clarity | originality | balance of images, graphics & text | appropriateness of images used**.**	ILO5. …communicate high quality outputs using a range of media	6) Structure and presentation	The poster is clear, structured, and organised. Covers all aspects and is appropriate for the target audience.

Once the evaluation criteria were defined, the next step consisted of writing down a generic statement that described the performance of the student in that specific criterion. The generic statement was then used to develop a series of progressive statements that translate the levels of achievement along the rating scale.

Levels of Achievement

The second challenge is agreeing on a marking scale. The discussion on which scoring system to adopt was made simple by the 0-100% scale ranging from fail to distinction and commonly used in higher education in the UK. Common marking scales assign marks to levels of achievement ranging from beginner to expert (Figure 3), from poor to excellent (Figure 4), or from fail to distinction (Figure 5). However, the agreement on which granularity to adopt for the scale was not that easy because markers had different opinions and would generally prefer more granularity, especially around the upper middle of the scale (60%). This was justified by the gaussian distribution of marks. Furthermore, while most scales in the literature are represented in a descending way, markers expressed their preference for an ascending scale mainly because it was "more intuitive" and "more practical".

FIGURE 3
MARKING SCALE FROM BEGINNER TO EXPERT

Performance level	Expert	Proficient	Competent	Novice	Beginner
Score %	90-100	70-89	60-69	50-59	40-49

FIGURE 4
MARKING SCALE FROM POOR TO EXCELLENT

Performance level	Excellent	Very good	Good	Intermediate	Poor level
Score %	90-100	70-89	60-69	50-59	40-49

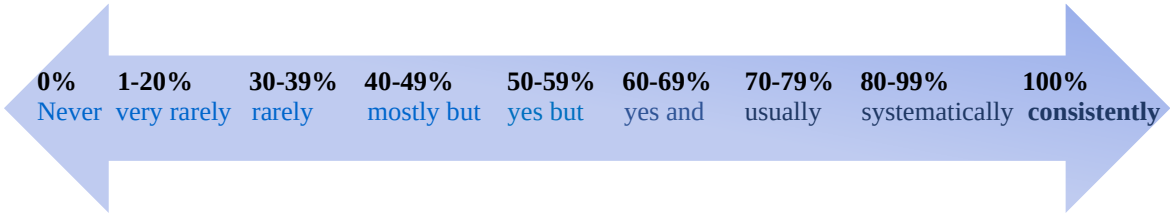
FIGURE 5
MARKING SCALE FROM FAIL TO DISTINCTION

Performance levels	Poor (F-Fail)		Poor (D-Fail)		Satisfactory (C-Pass)	Good (B-Merit)	Excellent (A-Distinction)		
Score %	0	1-29	30-39	40-49	50-59	60-69	70-79	80-99	100

Descriptions

The third challenge is the formulation of qualitative descriptions for each level of achievement in each dimension of the rubric. These statements should describe what an excellent, a very good, a good, a satisfactory and an unsatisfactory performance look like. Therefore, it is not very helpful to use these same words to describe the corresponding levels. For our MSc rubrics, we adopted a systematic progression based on the technique suggested by Andrade (2000), which starts with "no" and then "no but" followed by "yes", "yes but" and "yes and…". This technique was complemented by considering whether the student's performance in a given dimension was systematic and consistent throughout the assignment or whether it was observable most of the time, sometimes, or only rarely or very rarely (Figure 6).

FIGURE 6
TECHNIQUE TO WRITE RUBRIC DESCRIPTIONS INSPIRED BY ANDRADE (2000)



We were also guided by the general framework based on the SOLO Taxonomy and described by Biggs (2003a) as being helpful in writing verbs for the intended learning outcomes. This framework applies to the category of ILOs related to "knowledge and understanding" and describes four levels of "understanding" starting from a minimal understanding followed by a more descriptive understanding, then progressing to integrative understanding and finally to an extended understanding (see table 3).

TABLE 3
LEVELS OF UNDERSTANDING BASED ON THE SOLO TAXONOMY

Levels of understanding	Description
Minimal understanding	Sufficient to deal with terminology, basic facts. Verbs: memorize, identify, recognize
Descriptive understanding	Knowing about several topics. Verbs: classify, describe, list.
Integrative understanding	Relating facts together and understanding basic theory. Verbs: apply to known contexts, integrate, analyse, explain the aetiology.
Extended understanding	Being able to go beyond what has been taught, deal creatively with new situations. Verbs: apply to novel contexts, hypothesize, reflect, generate.

By applying these fours levels of understanding, we were able to generate progressions for other types of ILOs which are not necessarily within the category of "knowledge and understanding" (Figure 7).

FIGURE 7
TECHNIQUE TO WRITE RUBRIC DESCRIPTIONS INSPIRED BY GIBBS (2003A)

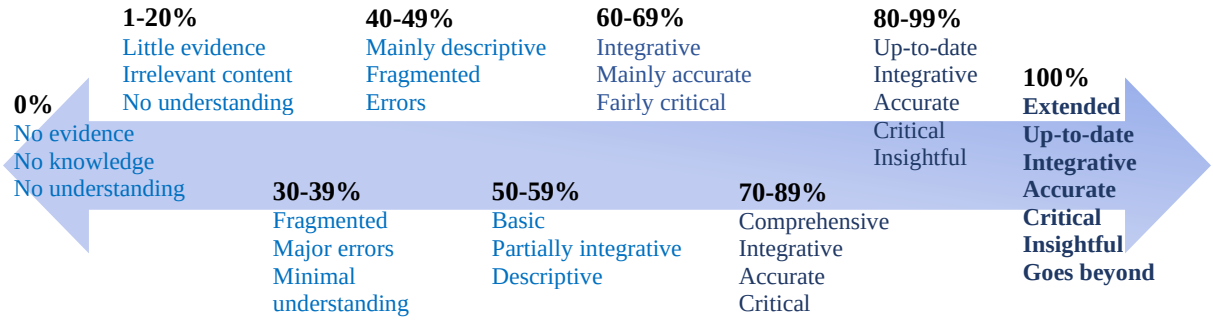


Figure 7 illustrates the technique used to write descriptions for multiple rubrics and for various types of rubric dimensions at each level of achievement. This technique considers the level of "descriptive understanding" (table 3), for example, as falling below the pass mark in a Master's level assignment.

By integrating both techniques described in figure 6 and figure 7, we followed a systematic way to write progressive descriptions from lower to higher levels of achievement for every dimension. This presented the benefit of providing a certain consistency -a priori- across the dimensions of the rubrics (the same progression of levels is applied systematically to every dimension), across the levels of achievement (a given level of achievement is defined in the same way for all dimensions), and across rubrics (the same technique is applied to all rubrics). For example, the pass level ranging from 50 to 59% is awarded to students who demonstrate integrative skills but only partially, as their work may still include passages that indicate descriptive levels of understanding.

Technological Challenges to Developers

Regarding the technological challenges to developers, we'll focus on the challenges presented by the use of rubrics for online marking. There are various virtual learning environments and platforms that suggest setting assignments with added rubrics. In addition to the well-known plagiarism software Turnitin (TII), virtual learning environments like Google Classroom, Microsoft Teams, and Blackboard Learn also provide the possibility of adding rubrics to assignments which are set on their platforms. However, all these platforms present technical issues, especially when it comes to importing rubrics which have been already created and agreed upon outside the platform (Table 4).

TABLE 4
CREATING AND IMPORTING RUBRICS INTO DIGITAL PLATFORMS

Digital platforms	Can rubrics be created?	Can rubrics be imported?
Turnitin (TII)	Yes	Yes, if the rubric is created as a spreadsheet and properly formatted in order to be easily imported into TII. Instructions need to be followed regarding the size of the title and rubric's dimensions. More adjustments are required upon import. If created in TII, the rubric can be exported as .rbc file and reimported into TII but no editing is possible.
Google Classroom	Yes	Yes, if created in Google Sheets only. If created in Google Classroom, the rubric can be exported to Google Sheets, edited, and reimported.
MS Teams	Yes	Yes, if created in MS Teams only. The rubric can be exported as .CSV file, edited, and reimported.
Blackboard Learn	Yes	Yes, if created in Blackboard Learn. The rubric can be exported and reimported as it is with no editing possible: the exported file is a zip folder that contains multiple files.

Table 4 shows that in most cases the rubric must be created within the platform and then exported. The exported file can be reimported into the platform –to be used for another assignment—but not all the platforms allow a format file that can be edited before reimporting.

In all platforms, the editing interface is not as user-friendly as a commonly used editing software or spreadsheet and, generally, does not offer the same possibilities of co-editing, reviewing, and commenting. Instructions and video tutorials on how to import rubrics seem to be based on the assumption of the

individual teacher rather than a teaching team and do not acknowledge the collaborative and iterative process of developing well thought-through rubrics.

As discussed before, the development of rubrics is rather a cyclical process that goes through various loops either with the students or with the teaching staff or both and it would be more practical if the rubric could be co-edited and commented upon outside the platform to be then imported as a ready file. Also, it must be possible to modify a rubric to be reused for the same assignment or adapted for another assignment or another course.

TII offers the possibility of importing spreadsheets, but, once imported, some adjustments have to be made within the platform. These adjustments affect the weightings of the criteria and the marking scale, which have to be coded manually after having uploaded the rubric. If the marking scale is set as a succession of marking ranges, it is required to set a point value for each range to enable the automatic calculations of the final score. But, this applies to all platforms.

CHALLENGES TO THE IMPLEMENTATION OF ANALYTIC RUBRICS

Pedagogical Challenges to Users

In terms of pedagogical challenges to users, it is important to consider the work of Sadler (2009) and the variations of marking approaches among markers. In addition to the multiple issues of marking in general, the use of rubrics raises questions as to whether the markers adopt a criterion-based approach while using the rubrics and judge students on how well they demonstrated progress towards the desired LOs or whether they rely on common academic norms despite the criteria-based rubrics.

A shared understanding of the criteria and a shared interpretation of the quality descriptions are considered in the literature as key for reliability (Andrade, 2000; Hack, 2015; Sadler, 2009). Therefore, the provision of training alone might not be sufficient to achieve this shared understanding. Involving markers in the process of developing the rubrics is essential for consistency, validity and reliability purposes while also mitigating some of the technological challenges.

Technological Challenges to Users

The Technological challenges are mainly related to the general challenges of navigating the digital world, the familiarity or non-familiarity of markers with the digital interface chosen to host the assignment and the rubric. Many markers expressed their preferences for paper marking and would rather use a paper version of both the rubric and the assignment before they "*encode everything in the system*". To support staff, we organised various training sessions, we produced guidelines, and we offered one-to-one support sessions.

Other technological challenges are related to the adaptability of the digital platforms hosting the rubrics to the virtual learning environments (VLEs) used for teaching and learning and to the institutional marking regulations.

For example, double blind marking on TII is only possible through the setting of three submission boxes: one for each marker and one for the moderated marking. While all markers are enrolled in the same course, they are enabled to see each other's marking and should rely on their academic integrity not to do so. Furthermore, the feedback comments made in the first and second boxes can only be transferred manually to the third box to be then released to students together with their final mark.

This kind of challenges can be worked around resulting in rather complicated workflows and time-consuming processes.

SUMMARY

In this paper, four categories of challenges raised by the assessment rubrics are discussed: pedagogical and technological challenges. A distinction is made between the challenges related to the process of developing the rubrics and those related to their implementation by users.

The pedagogical challenges to developers are mainly focused on the 3 main features of a rubric. First, it is important to make sure that the evaluation criteria are mapped to the LOs and do not focus on assessing the components of student work especially if those components are not taught, such as creativity or originality, for example. The quality descriptions should reflect a clear progression towards integrative learning and critical thinking skills. And finally, the marking scale should be familiar to markers and should reward higher order learning outcomes.

The pedagogical challenges to users are mainly related to the variations in marking approaches and the reliability requirements. It is not obvious to mark against LOs even with a criteria-based rubric, reliance on norm-referencing as mentioned by Sadler (2009) is still present. To overcome these challenges, we decided to involve a group of markers in the process of developing the rubrics in addition to organising training sessions and conducting sample moderations.

In terms of technological challenges to developers, it is important that the VLEs provide adapted features to easily upload rubrics that have been created elsewhere and that the interface to create new rubrics is more user-friendly. Other challenges related to second marking, anonymous marking, and double blind marking need to be addressed. In this sense, not only the required features should be enabled on the digital platforms, but also the institutional requirements and all related policies should be accounted for.

As to users, the technological challenges do not differ much from the challenges of navigating the digital spaces in general. The familiarity with and access to those spaces and their features are key and thus training and guidelines are essential but, for busy academics, one-to-one tutorials are highly recommended.

TABLE 5
SUMMARY OF PEDAGOGICAL AND TECHNOLOGICAL CHALLENGES OF DEVELOPING AND USING RUBRICS

Challenges	**Pedagogical**	**Technological**
Developers	Evaluation criteria to map the LOs Quality descriptions to map the evaluation criteria and thus the LOs Quality descriptions to reflect clear progression towards integrative learning Marking scale: familiar to markers; rewards higher order thinking skills	Features of the VLEs to be adapted to the needs of pedagogical teams Facility of importing or uploading ready rubrics created in commonly used editing software or spreadsheet User-friendly interface for creating new rubrics User-friendly interface for reusing and editing/adapting old rubrics
Users	Norm-referenced or criterion-referenced marking? Marking to level 7 LOs? Shared understanding of the quality descriptions Shared interpretation of the quality descriptions Involvement in the development and testing of rubrics Training sessions Pedagogical guidelines	Screen or paper? Online or offline? Familiarity with and access to the digital interface Familiarity with the features of the interface Training sessions One-to-one support Technical guidelines

CONCLUSIONS

This paper shows how the implementation of assessment rubrics could respond to the needs of new forms of assessing and to the assessment of a broad variety of intended learning outcomes such as those involving higher order thinking skills, complex thinking, and effective communication among others. The rubrics are pointed out as a means to meet current demands on curricula design, especially in the European context of the Bologna process, but also in the context of wider transformations taking place in higher education, namely in terms of curriculum design, teaching practices, learning approaches, and assessment strategies.

The use of analytic rubrics to assess student assignments submitted online is also adapted to digital settings allowing the assessment of a variety of assignments involving large cohorts and multiple markers with variable experience. Nevertheless, the development of rubrics and their implementation inherent to the use of digital platforms, add up to the challenges they represent to both the developers and users of such rubrics. If many of the challenges met can be addressed by the pedagogical teams and the educational technologists (e.g. involving staff and/or students in the development of rubrics; organising training on how to mark with rubrics, how to integrate rubrics into a VLE, or how to use the digital platforms hosting the rubrics), other challenges can only be addressed by collaborative efforts between the developers of digital platforms and the developers of VLES, as well as between the two and the pedagogical teams, curriculum designers, and learning technologists.

This suggests that if we were to bring about change in the way we assess and how we assess student learning, and thus achieving some of the desired shifts in curriculum design, the technological solutions have to be adapted to the needs of education and not the other way around.

REFERENCES

Adie, L.E., Willis, J., & Van der Kleij, F.M. (2018). Diverse perspectives on student agency in classroom assessment. *Australian Education Research, 45*, 1–12.

Allen, S., & Knight, J. (2009). A Method for Collaboratively Developing and Validating a Rubric. *International Journal for the Scholarship of Teaching and Learning*, *3*(2), 10.

Andrade, H.G. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, *57*(5), 13–18.

Azmi Dandis, M. (2014). Benefits and criticism. Should we support or neglect using rubrics? Evidences from a literature review. *Edmetic*, *3*(2), 91–113.

Baker, E.L., O'Neil, H.F., & Linn, R.L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, *48*(12), 1210.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan, & F. Pajares (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Information Age Publishing,

Biggs, J. (1999). What the Student Does: Teaching for enhanced learning, *Higher Education Research & Development*, *18*(1), 57–75. DOI: 10.1080/0729436990180105

Biggs, J. (2003a). Aligning teaching and assessment to curriculum objectives. *Imaginative Curriculum Project, LTSN Generic Centre, 12*.

Biggs, J. (2003b). *Teaching for quality learning at university*. Buckingham: The Open University Press.

Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher education: A practical guide*. Maidenhead: Open University Press.

Bologna Working Group. (2005). *A Framework for Qualifications of the European Higher Education Area*. Bologna Working Group Report on Qualifications Frameworks. Copenhagen: Danish Ministry of Science, Technology and Innovation.

Brown, G.A., Bull, J., & Pendlebury, M. (2013). *Assessing student learning in higher education*. Routledge.

Brundiers, K., Wiek, A., & Redman, C.L. (2010). Real-world learning opportunities in sustainability: from classroom into the real world. *International Journal of Sustainability in Higher Education*, *11*(4), 308–324.
Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in Education and Teaching International*, *44*(1), 57–66.
Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325.
Cedefop-European Centre for the Development of Vocational Training. (n.d.). *European Qualifications Framework*. Retrieved from https://www.cedefop.europa.eu/en/events-and-projects/projects/european-qualifications-framework-eqf
Frey, B.B., Schmitt, V.L., & Allen, J.P. (2012). Defining Authentic Classroom Assessment. *Practical Assessment, Research, and Evaluation*, *17*(1), 2.
Hack, C. (2013). Using rubrics to improve marking reliability and to clarify good performance. *The Higher Education Academy STEM Conference*. London, 2013.
Hack, C. (2015). Analytical rubrics in higher education: A repository of empirical data. *British Journal of Educational Technology*, *46*(5), 924–927.
Imperial College London. (2017*). Innovative teaching for world class learning. Learning and Teaching Strategy*. Retrieved from https://www.imperial.ac.uk/learning-and-teaching-strategy
Jääskelä, P., Poikkeus, A-M., Vasalampi, K., Valleala, U.M., & Rasku-Puttonen, H. (2017). Assessing agency of university students: Validation of the AUS Scale. *Studies in Higher Education*, *42*(11), 2061–2079, DOI: 10.1080/03075079.2015.1130693
Kuh, G.D., Kinzie, J., Schuh, J.H., & Whitt, E.J. (2005). *Assessing conditions for student success: An inventory to enhance educational effectiveness*. San Francisco: Jossey-Bass
Marzano, R., Pickering, D., & Mctighe, J. (1993). *Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model*. Alexandria, VA: Association for Supervision and Curriculum Development.
Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*, *45*(4), 527–540.
Montgomery, K. (2002) Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching. *College Teaching*, *50*(1), 34–40. DOI: 10.1080/87567550209595870
Moseley, D., Baumfield, V., Elliott, J., Gregson, M., Higgins, S., Miller, J., & Newton, D. (2006). *Frameworks for Thinking. A Handbook for Teaching and Learning*. New York: Cambridge University press.
Newell, W.H. (2010). Educating for a Complex World: Integrative Learning and Interdisciplinary Studies. *Liberal Education*, *96*(4), 6–11.
Nordrum, L., Evans, K., & Gustafsson, M. (2013). Comparing student learning experiences of in-text commentary and rubric-articulated feedback: strategies for formative assessment. *Assessment & Evaluation in Higher Education*, *38*(8), 919–940. DOI: 10.1080/02602938.2012.75822
Prosser, M., & Trigwell, K. (2006). Confirmatory factor analysis of the Approaches to Teaching Inventory. *British Journal of Educational Psychology*, *76*, 405–419.
Ramsden, P. (1985). Student learning research: retrospect and prospect. *Higher Education Research and Development*, *4*(1), 51–69.
Ramsden, P. (2003). *Learning to Teach in Higher Education*. London: Routledge.
Reddy, Y.M. & Andrade, H. (2010). A review of rubric use in higher education. *Assessment and Evaluation In Higher Education*, *35*, 435–448.
Resnick, L.B., & Resnick, D.P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B.R. Gilford, & M.C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37–75). Boston: Kluwer.
Romero-González, M. (2021). Mapping skills for curriculum design using a programme level approach in a transnational education context. In S. Norton, & R. Dalrymple (Eds.), *Employability: Breaking*

the mould A case study compendium. York: Advance HE. Retrieved from: https://www.advance-he.ac.uk/news-and-views/shifting-emphasis-content-skills
Sadler, D.R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179. DOI: 10.1080/02602930801956059
Smith, K.A., Sheppard, S.D., Johnson, D.W., & Johnson, R.T. (2005). Pedagogies of engagement: Classroom-based practices. *Journal of Engineering Education*, *94*(1), 87–101.
Stentoft, D. (2017). From saying to doing interdisciplinary learning: Is problem-based learning the answer? *Active Learning in Higher Education*, *18*(1), 51–61.
Stevens, D.D., & Levi, A.J. (2012). *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning*. Sterling, Virginia: Stylus Publishing.
Thomas, L. (2012). *Building Student Engagement and Belonging in Higher Education at a Time of* Change. Final report from the What Works? Student Retention & Success programme. Paul Hamlyn Foundation. Retrieved from: https://www.phf.org.uk/wp-content/uploads/2014/10/What-Works-reportfinal.pdf
Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, *27*(5), 615–631.
Trigwell, K., & Prosser, M. (2004). Development and Use of the Approaches to Teaching Inventory. *Educational Psychology Review*, *16*(4).
Trigwell, K., Prosser, M., & Ginns, P. (2005). Phenomenographic pedagogy and a revised Approaches to teaching inventory. *Higher Education Research & Development*, *24*(4), 349–360. DOI:10.1080/07294360500284730
Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: creating a blueprint for course design, *Assessment & Evaluation in Higher Education*, *43*(5), 840–854. DOI: 10.1080/02602938.2017.1412396
Wieman, C., & Gilbert, S. (2014). The teaching practices inventory: A new tool for characterizing college and university teaching in mathematics and science. *CBE—Life Sciences Education*, *13*(3), 552–569.
Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation*, *2*(1), 2.