Chapter 7

The Use and Design of Rubrics to Support AfL

Anders Jonsson and Ernesto Panadero

Abstract Rubrics are assessment instruments designed to assist in identifying and evaluating qualitative differences in student performance. Research into scoring rubrics has shown that they can serve two purposes: (1) aid assessors in achieving higher levels of consistency when scoring performance tasks, and (2) promote learning and/or improve instruction by making assessment expectations explicit and aiding the feedback process. In this chapter we summarize research on the formative use of rubrics, in order to identify how the use and design of rubrics may be optimized for the purpose of supporting student learning in an environment that often stresses independence and management of own learning. The presentation is organized around two different pathways through which rubrics may support student learning. These are through (a) facilitating the understanding and use of feedback and through (b) supporting students' self-regulated learning. We also analyze what is known about the implementation of rubrics in higher education, with a particular focus on more sustained and widespread implementations. The implications of these findings for both practice and future research on rubrics are discussed.

Introduction

Research on the use of scoring rubrics has shown that these assessment instruments can aid assessors in achieving acceptable levels of consistency when scoring performance tasks. Research has also documented positive educational consequences of rubric usage, such as supporting students' development towards independent learners and improved student performance. These effects seem to come from the fact that rubrics make expectations and criteria explicit, which in turn facilitate other processes, such as interpreting and using feedback. As a consequence, rubrics have been used as a tool for communicating expectations, as well as a support for other assessment-for-learning (AfL) practices (Jonsson and Svingby, 2007; Panadero and Jonsson, 2013).

Although most studies explicitly linking the use of rubrics to AfL are performed in school settings, there are a number of studies from higher-education contexts as well. A majority of these studies, however, are small scale and use short-term interventions (Brookhart and Chen, 2014; Panadero and Jonsson, 2013), which makes it difficult to get an overview of how to successfully use and design rubrics for formative purposes. The first aim of this chapter is

therefore to summarize research on rubrics, in order to identify how the use and design of rubrics may be optimized for the purpose of supporting student learning in an environment that often stresses independence and management of students' own learning. The second aim is to analyze what is known about the implementation of rubrics in higher education, with a particular focus on more sustained and widespread implementation.

What is a Rubric?

Assessing complex tasks, such as dental students communicating with patients or science students formulating arguments in socio-scientific issues, is difficult and a common recommendation for increasing the quality of teachers' assessment has therefore been to use detailed scoring protocols, preferably together with sampled responses which exemplify the points on the scoring scale (e.g. Linn, Baker, and Dunbar, 1994). An instrument that effectively matches these recommendations is the rubric. Rubrics are designed for assisting assessors in judging the quality of student performance and in order to accomplish this, all rubrics have three fundamental features in common. First, in order to assist in identifying the qualities to be assessed, the rubric includes information about which aspects or criteria to look for in student performance. Second, in order to assist in judging the quality of student performance, the rubric includes descriptions of student performance at different levels of quality. By combining the aspects to be assessed with the descriptions of quality into a two-dimensional matrix, along with a scoring strategy (i.e., the third feature), a rubric comes into existence.

How the Use of Rubrics may Support Student Learning

As established by reviews on research about rubrics, the transparency provided by rubrics may facilitate other AfL related processes, such as interpreting and using feedback or assessing the performance of peers (Jonsson and Svingby, 2007; Panadero and Jonsson, 2013). But there are also other less direct ways in which the use of rubrics have been shown to support student learning, such as reducing anxiety or supporting self-regulated learning (SRL) strategies. Below, these two pathways – i.e. the more direct pathway through facilitating the understanding and use of feedback, as well as the indirect pathway via SRL strategies – are outlined and illustrated with relevant studies.

Pathway 1: Facilitating the Understanding and Use of Feedback

According to literature reviews and meta-studies, such as the one by Hattie and Timperley (2007), feedback can have a strong influence on student learning. However, a number of students do not use the feedback they receive, and therefore do not realize the potential of feedback for learning (Jonsson, 2013). There may be several reasons for not utilizing feedback, but two of the main obstacles seems to be that students either do not understand the feedback or do not know how to use it. By using rubrics to support the feedback process, both of these difficulties may be mitigated.

First, by making assessment criteria explicit, the understanding of feedback can be facilitated. As an example, Case (2007) wanted to reconfigure the feedback process through students' active engagement with explicit assessment criteria. To this end, he devised an electronic template for feedback, which incorporated the assessment criteria (in the shape of a rubric) as the main feedback element. Along with the feedback form, there was a bank of electronically-stored statements, which served as the basis of more specific feedback. This bank included comments relating to common student weaknesses, mistakes, recommended improvements and strengths. Findings show that the changes had a positive effect on student performance and that these improvements were largely attributable to the fact that the feedback process enhanced students' awareness and understanding of learning outcomes and assessment criteria. It is also noteworthy that there was a reduced marking time due to the use of statement banks and fewer student queries as a result of the explicit engagement with criteria. Similar results have been obtained in studies combining rubrics with exemplars (Jonsson, 2010).

Second, the fact that rubrics – by definition – include descriptions of student performance at different levels of quality, means that there is an inbuilt 'feed-forward' feature in all rubrics. This, in turn, may aid teachers when struggling to provide constructive feedback to students. For instance, Schamber and Mahoney (2006) report that most faculty members in their study found a rubric for assessing critical thinking to be very useful in order to provide students with feedback on drafts. Students also found the same rubric useful for clarifying expectations. Furthermore, field supervisors reported that a rubric for clinical supervision in counsellor education facilitated the provision of feedback on a continuous basis as well as giving concrete recommendations for improvements (Hanna and Smith, 1998).

Feedback need not only be provided by teachers, but also by peers or the student herself. There are, however, few scientific studies reporting on effects of self- and peer assessment using rubrics (Panadero and Jonsson, 2013). Also, in a number of such studies the main focus has not been on learning, but on accuracy (Jonsson and Svingby, 2007). One interesting example of research on self- and peer assessment using rubrics is a study by Reitmeier, Svensden, and Vrchota (2004). In their study, students worked with self- and peer-assessment of oral-communication skills in a food-preparation course and both the scores of the

presentations and the grades were higher as compared to the scores and grades attained the previous semester. These changes were attributed to the use of rubrics. In another study, dental students used a rubric for communication skills to provide feedback to each other during role play with a simulated patient. The evaluation of this exercise was very positive and most students thought they learned from the task and that the criteria were useful for them (Lucander, Knutsson, Salé, and Jonsson, 2012). Similar results have been reported for students using a rubric to monitor their task performance in statistics and epidemiology, as well as for real-estate-broker students (Jonsson, 2014).

Pathway 2: Facilitating Students' Self-regulated Learning

Rubrics have also been demonstrated to support student learning through different aspects of SRL. Some of these studies overlap somewhat with previously mentioned studies on self-assessment, where students were seen to use rubrics to monitor and evaluate their task performance (e.g. Jonsson, 2014). This is not surprising, given that self-assessment is considered an integral part of most models of SRL, since students need to monitor and evaluate their progress in order to regulate their work (Panadero and Alonso-Tapia, 2013). For instance, according to Zimmerman's (2013) cyclical model of self-regulation, which is one of the most cited in the literature, SRL consists of three phases (forethought, performance and self-reflection) and self-assessment is an important element in at least two of these phases (i.e., performance and self-reflection).

The forethought phase, on the other hand, is more often associated with goal-setting, planning and self-efficacy, because students do not necessarily self-assess in this part of the cycle. However, the forethought phase does provide important conditions for the self-assessment to come, for example by establishing the assessment criteria (Panadero and Alonso-Tapia, 2013). There seems to be no research where students have used rubrics as a tool for goal-setting, but there are a number of studies investigating students' planning. For instance, in a study by Andrade and Du (2005), students reported using the rubric to plan their approach to an assignment, much like a recipe or a map. Also, in a replication study by Reynolds-Keefer (2010), students' responses indicate that rubrics aided them in both planning and when performing the assignment. Most students claimed to read the syllabus and then start working on the assignment, using the rubric as a reference point. Several students also stated that they worked through the assignment by reading the rubric and working on one portion at a time, merging all the separate parts before submitting.

In Panadero and Jonsson (2013), rubrics were found to mediate improved student performance by for instance reducing anxiety. These effects too are contingent on the transparency provided by the rubrics. For instance, when asked about anxiety, pre-service

teachers spoke about increased confidence and making it easier to hand in assignments when having access to a rubric (Andrade and Du, 2005).

Closely related to the issue of anxiety, is students' perceptions of their own capability and the impact on student self-efficacy from using rubrics has been investigated in a number of studies. The findings, however, are inconclusive. For instance, Panadero (2011) investigated the relationship between self-efficacy and rubric use in three different studies. In this research, self-efficacy was impacted by the use of rubrics in only one of the cases (with secondary students) and only in interaction with the type of feedback received. It should be noted that these studies did not control for whether the students had received feedback by the teacher, which is a factor suggested to significantly affect students' perceptions of their performance. Still, it is difficult to find studies where higher-education students' self-efficacy is affected by the use of rubrics. An interesting exception is a study by Yopp and Rehberger (2009), where low-performing mathematics students were given a rubric and told that they had to master the material (i.e., obtain a score of 100% based on the rubric). While instruments for measuring both general and subject-specific self-efficacy were administrated to the students, only the mathematics self-efficacy changed during the intervention. According to the authors, a possible explanation for this could be that academic attitudes and beliefs are difficult to change in a single specific course. This explanation is corroborated by Balan (2012), but since the self-efficacy questionnaires in the research by Panadero (2011) were task specific, the difference between general and subject- or task-specific measures does not seem to explain the inconclusive results in relation to self-efficacy and rubrics.

Panadero and his colleagues have also done a number of studies relating to students' learning orientations and SRL. In one of their investigations, they found that the level of SRL strategies, as measured through think-aloud protocols, was higher in a group of secondary education students using rubrics as compared to students in a control group (Panadero, Alonso-Tapia, and Huertas, 2012). In another study, it was found that scores on a performance- and avoidance-oriented SRL scale decreased for pre-service teachers using rubrics (Panadero et al., 2013). In yet another study, Panadero and Romero (2014) found that a group of pre-service teachers using rubrics scored higher on a learning-oriented SRL questionnaire, as compared to students who were asked to self-assess their work without any instrument to facilitate the self-assessment. Performance- and avoidance-oriented SRL scores also decreased significantly in the rubric group.

The findings above are indications of positive effects on students' SRL, but there are other findings as well. The students using rubrics in the study by Panadero and Romero reported higher levels of stress while performing the task as compared to the control group. Also, the learning-oriented SRL scores decreased for first year psychology students who used rubrics (Panadero, Alonso-Tapia, and Huertas, 2014). This means that while the use of rubrics may decrease performance- and avoidance-oriented SRL strategies, which are often detrimental for learning, they do not necessarily increase learning-oriented SRL.

In summary, the use of scoring rubrics has been shown to facilitate student learning through two (more or less) distinct pathways. The first pathway involves aiding students in understanding and using feedback from teachers, peers or themselves. The fact that rubrics contain different levels of quality also means that it is easier to provide – and possibly to interpret – constructive feedback. The second pathway involves an impact on SRL and related factors such as self-efficacy, student anxiety, and learning-orientation; factors that may in turn affect student performance. Again, an important requirement seems to be the transparency provided, which makes it possible for students to estimate their own capability, as well as to plan, monitor and evaluate their work according to the explicit criteria. This means that students can exert more control of their own learning, which potentially reduces anxiety and negative SRL strategies. At the same time, however, this control and responsibility can create feelings of stress and also make students more performance oriented. An extreme version of this performance orientation is reported by Torrance (2007), where an unfortunate combination of explicit criteria and coaching in post-secondary education resulted in a situation where 'criteria compliance' tended to replace productive learning.

How the Design of Rubrics can Support Student Learning

As noted by Dawson (2015), the term 'rubric' is sometimes used with different meanings. Rubrics may therefore be 'embraced and resisted based on often imprecise and inconsistent understandings of the term' (p. 1). Consequently, Dawson proposes a framework for the categorisation of rubrics, encompassing fourteen design elements that can be used to make distinctions between different rubrics. This framework, with Dawson's nomenclature, is used in this section.

It should be noted that Dawson (2015) does not express support or preference for any decisions regarding any particular design. In his case this has to do with his research approach, but generally it also has to do with the fact that many research articles do not include complete information about the rubrics used and that most of the design elements of rubrics have not been systematically investigated (Panadero and Jonsson, 2013). Still, there are some tentative recommendations that can be made, based on existing research. From their review of 75 studies of rubrics, Jonsson and Svingby (2007) identified a number of important design features of rubrics, such as 'Specificity' (i.e., whether the rubric is more task-specific or more general), 'Scoring strategy' (i.e., whether the scoring is holistic or analytic) and 'Quality levels' (i.e., the number and type of levels of quality). Importantly, these recommendations differ depending on the main purpose of using rubrics. For summative purposes, Jonsson and Svingby suggest using task-specific rubrics with few levels in order to increase reliability. Such a design would not, however, be appropriate for formative purposes since fewer quality levels would make the rubric less useful for providing and understanding

constructive feedback. Furthermore, a rubric that can only be used for one particular task is of less use for formative assessment. As it may take a while for students to learn how to use a rubric, the rubric needs to be applicable to at least a group of similar tasks (called 'task-type rubrics' in Dawson's terminology). There are examples of successful use of generic rubrics, for instance a study by Balan (2012), where a generic rubric for mathematical problem-solving in upper-secondary school was used. On the other hand, it took the students several weeks to comprehend and use the rubric for self- and peer assessment purposes. Learning to use this rubric was therefore an investment, which paid off since the students were able to use the rubric during the remaining part of the course, whenever they encountered a mathematical problem-solving situation. Generally, however, it is doubtful whether such generic rubrics provide enough guidance to students in higher-education contexts, since the students often work more independently and within shorter courses or modules.

Another important design feature, which is affected by the purpose of the assessment, is the scoring strategy. Students' strengths and weaknesses are the raw material for formative assessments and this information becomes visible through an analytic scoring strategy. But in contrast to summative assessments, the assessment of separate criteria does not have to be summarized into a total score. A rubric does not even have to include any numerical scores. As noted by Dawson (2015), the quality levels in a rubric may come from taxonomies like the SOLO taxonomy, but they can also be expressed through grade levels or statements of student proficiency. Consequently, there is nothing inherent in the design or use of rubrics that has to do with quantification of knowledge.

The distinction between identifying strengths and weaknesses versus scores is important for several reasons. While scores may also be used to identify strengths and weaknesses, their main merit is that the addition of scores is compensatory. This means that weaknesses in some areas can be compensated for by strengths in other areas. On the other hand, as soon as the aggregated score is calculated, the original pattern of strengths and weaknesses disappears (Sadler, 2005). A summary score therefore has little value for formative assessment purposes, which means that formative assessment should primarily focus on an analytical assessment of strengths and weaknesses, not on aggregated scores or holistic judgments.

A design feature of rubrics that has become more or less common practice is to have the same number of quality levels for all criteria. Even if this might give an impression of coherence and logic, it has been shown that such an arrangement can affect the scoring process, so that assessors to a greater extent give the same score in relation to several criteria (i.e. a kind of halo effect). As a result, validity is affected negatively since the variance of what is actually assessed becomes narrower than in the original construct. This problem may disappear, however, by designing rubrics with a different number of quality levels for different criteria (Humphry and Heldsinger, 2014).

Jonsson (2014) has done an in-depth study on students' use of rubrics in professional education, where he investigated three different assessment situations, which included the use of rubrics. In one case, public-health students in a course on statistics and epidemiology constructed a questionnaire and a database with fictitious data. In the second case, real-estate-broker students systematically reviewed a house. And in the third case, dental students were assessed by peers when communicating with a simulated patient. As shown by results from questionnaires and interviews with students, students in all cases perceived the criteria as both comprehensible and useful. They also actively used the rubric. For instance, the real-estate-broker students used the rubric for planning, as well as to monitor and evaluate their task performance. The dental students used the criteria to discuss each other's performances and give each other feedback, as well as to reflect about their performance as professionals beyond the scope of the assignment. From these cases, Jonsson identified two important factors, which seemed to facilitate the communication of expectations to the students. These factors are called 'accessibility' and 'alignment'.

The teachers made the rubrics *accessible* to the students, both in terms of understanding and availability. First, the teachers explained the meaning of the criteria in the rubric, which was done criterion by criterion, holistically or by letting the students use the criteria during an instructional event. Second, the teachers presented the rubrics to the students *before* they carried out their assignments. This means that the students could use the rubrics as guides when planning, monitoring and evaluating their performance. Third, the teachers made the rubrics available to the students by publishing the documents digitally or by handing them out on paper. The students did not, therefore, have to rely on teachers' oral description and their own interpretations and notes. Instead, the criteria could be reviewed and discussed both individually and among peers. The students could also have the rubrics beside them when they performed their assignments or, in the case of the dental students, assessed their peers.

Alignment refers to how the rubrics were aligned with the tasks. In these cases, the rubrics were analytic and of task-type specificity. But they were also 'direct' by focusing on the performance of skills that students were expected to master. The directness of rubrics has the advantage of potentially aiding in the improvement of the skills sought to assess. This is in opposition to indirect assessments, where the connection between the skills sought for and what is actually assessed may not be clear to the students, due to the sometimes complex transformation of scores to interpretations of student performance. In this case, the directness of the criteria was seen to facilitate student engagement with the rubrics; to guide their performance and as tools for self-assessment and reflection.

Taken together, the possibility to provide recommendations on how to design and use rubrics is limited by the lack of a common terminology. There is also still much work to be done in systematically investigating different aspects of design and use (Dawson, 2015; Panadero and Jonsson, 2013). Still, there are some recommendations that can be made for designing and using rubrics to support AfL practices:

- Use an *analytic scoring strategy without summarizing into a total score*, so that the aspects to be assessed are explicitly spelled out and – most important – strengths and weaknesses in relation to individual criteria are discernible.
- Use *several quality levels*, so that the quality sought becomes visible to the students and for aiding in producing and understanding constructive feedback. It should be noted that having the same number of levels for all criteria may compromise the validity of scoring.
- Use *task-level specificity*, so that rubrics are neither too closely tied to the particular task nor too generic. Instead, rubrics need to be applicable to several, but similar, tasks assessing the same competency.
- Make the rubrics *accessible* to the students by (a) explaining the criteria and quality levels, (b) making the rubric available, digitally or on paper, and (c) providing the students with the rubric before they perform the task.
- Use *direct criteria*, so that they may guide student performance and facilitate self-assessment and reflection.

Towards a Sustained and Widespread Implementation of Rubrics in Higher Education

As suggested by the research reviewed above, a wider implementation of rubrics in higher education could have great potential, both for students' short-term performance and for their development towards independent learners. To date, however, research on larger scale implementations of rubrics for formative use is lacking and most studies are based on small samples and short-term interventions. This is particularly true in the higher-education context, whereas in school settings there are a number of studies with larger samples and longer interventions (Brookhart and Chen, 2014; Panadero and Jonsson, 2013).

Although few, there are some studies in a higher-education context, with at least medium-sized samples. For instance, besides the previously mentioned study by Jonsson (2014) in which 166 students participated, there is another study by the same author with a large sample size (Jonsson, 2010). In this study, a rubric was shared with pre-service teachers before performing an on-line examination of complex teaching skills. The same examination was studied during three consecutive years with a cohort of pre-service teachers in science and mathematics (n = 170, 154, and 138). Some changes were implemented after the first year, which were thought to increase the transparency of the assessment. For instance, there were some clarifications in the rubric and the students could access exemplars illustrating the quality levels in the rubric. These changes in transparency lead to major improvements in students' performance on the examination, especially for those students who claimed to have used the rubric actively during the examination. In both of these studies, however, the rubric intervention is limited. In Jonsson (2010), students were provided the rubric (along with exemplars) before the examination, and in Jonsson (2014) the students were provided the

rubric before they performed an authentic task. In the latter case, the criteria and levels of quality were explained to the students, but students were not offered any training in using the rubric in either of the cases.

In a study by Kocakülah (2010), on the other hand, physics students (n = 153) took active part in constructing a rubric. First, the students were introduced to the concept and design of rubrics and were then asked to design a rubric on their own in groups. The students were free to select the type of rubric they wanted, but afterwards they had to present their rubric to the whole class and justify their choices. From the presentations, the best rubric was chosen, collectively modified and finally used by the students while doing the test. Another example of a more extensive intervention is the study by Reitmeier et al. (2004) mentioned previously. Here rubrics were used for teacher, self- and peer-assessments and all students were required to self- and peer assess repeatedly over the course of a semester. As can be seen, these interventions are not as limited as in the abovementioned research. Interestingly, however, positive effects are reported in a number of studies in higher education, seemingly regardless of the extent of the intervention. One exception is a study by Green and Bowser (2006), where there were no significant differences between scores from students who had access to a rubric and those who did not. As pointed out by Reddy and Andrade (2010), however, the sample size in this study was very small (n = 16), which – together with the limited intervention – makes it difficult to draw any conclusions from this comparison. But this is not the only example. In Panadero et al. (2013), investigating 69 pre-service teachers, no significant effects for students' performance or self-efficacy were documented. On the other hand, students preferred the use of rubrics as compared to so called 'scripts' and there was a positive effect on students' SRL strategies.

Taken together, the general picture is that the use of rubrics is appreciated by the students and that rubrics often contribute to improved performance or students' SRL strategies. As noted by Panadero and Jonsson (2013), this is in contrast to studies performed in school settings, where longer interventions are often needed in order to produce a clear effect. Obviously, many students in higher education are capable of using rubrics productively, for instance to monitor and evaluate their work, with only minimal guidance.

Even if the findings presented above may seem promising, it should be kept in mind that there are still relatively few studies addressing formative aspects of using rubrics and that most of these are based on small samples and short-term interventions. Furthermore, in most studies students could be characterized as 'rubric neophytes', which means that not much is known about how rubrics are used by students more accustomed to such instruments or how students are affected when they are exposed to rubrics during longer periods of time. For instance, fears have been voiced about the limiting effects of rubrics (e.g., Sadler, 2009; Wilson, 2006). By providing criteria beforehand, we may guide some students, helping them to focus their efforts on what is considered important and worthwhile, but we may at the same time restrict others.

With reference to the fears of hampering students by providing them with criteria specified beforehand, Sadler (2009) makes a case for 'emerging criteria'. This means that assessors should address criteria that surface in the moment of assessing a particular piece of work – much like the appraisal by connoisseurs of art, wine, etc. The main argument is that when breaking down holistic judgments into more or less discrete components, these components – no matter how many they are and no matter how carefully they are selected – cannot sufficiently represent the full complexity of the multi-criterion qualitative judgment made by the connoisseur. Furthermore, qualities not represented by the criteria might be filtered out and not taken into account by the assessors. Instead of relying on analytic assessment and pre-set criteria as a vehicle for transparency in assessment, Sadler therefore argues that students need to develop a conceptualisation of what constitutes 'quality' by repeatedly evaluating authentic work.

A problem, however, is that novices may not know what to look for in authentic work. This is evident in a number of studies. An illustrative example is provided by Orsmond and Merry (1996), where students were asked to assess each other's work. Even though all criteria were explained to the students, they were unable to recognise some of these criteria in the work by their peers. For instance, a majority of students had actually drawn a 'clear and justified conclusion' (which was a criterion), but did not know it.

In our understanding, the question of using pre-set criteria or developing a conception of quality through evaluating authentic work is therefore not a question of either one or the other. Rather, what seems to be needed is an integration of both. Students need explicit criteria to know what to look for in authentic work, but they also need to experience authentic work in order to know how the criteria may be realised. Rubrics can provide a scaffolding structure for students when learning to identify indicators of quality, but like other scaffolding structures it can be disregarded if not needed and gradually phased out as the students become more independent.

This is exemplified in the study by Jonsson (2014), where the rubrics provided an important support for low-performing/low-confidence students, whereas some confident students actually refused to use the rubric because they wanted to manage on their own. The findings from this study thus suggest that rubrics do not necessarily limit high-performing/high-confidence students, since they may choose not to use it.

Conclusion

The main idea permeating this chapter is that rubrics can be used not only by teachers to increase the reliability and validity of summative assessments, but also shared with the students. Rubrics then become a tool for communicating expectations, which is a fundamental

requirement for successful AfL implementation. As a tool for communication, rubrics can be used to support different AfL processes, such as giving and understanding feedback, as well as self- and peer-assessment. Perhaps even more important, the transparency provided by the use of rubrics has been shown to support students' SRL, so that students are able to plan, monitor and evaluate their performance with the aid of a rubric. In order to enhance these effects, rubrics should be designed to promote formative-assessment practices and also made accessible to the students.

Taken together, a wider implementation of rubrics in higher education could have great potential, both for students' short-term performance and for their development towards independent learners. However, since research on long-term use of rubrics is lacking, as well as research on the effects of rubrics for students with different performance backgrounds or learning-orientations, a general recommendation could be to advance slowly towards a sustained and widespread implementation of rubrics in higher education.

References

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research and Evaluation, 10*(3), 1-11.

Balan, A. (2012). *Assessment for learning. A case study in mathematics education* (Unpublished doctoral dissertation). Malmö University, Malmö.

Brookhart, S. M., & Chen, F. (2014). The quality and effectiveness of descriptive rubrics. *Educational Review, 67*(3), 343-368.

Case, S. (2007). Reconfiguring and realigning the assessment feedback processes for an undergraduate criminology degree. *Assessment and Evaluation in Higher Education, 32*(3): 285-299.

Dawson, P. (2015). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment and Evaluation in Higher Education*. doi:10.1080/02602938.2015.1111294

Green, R., & Bowser, M. (2006). Observations from the field: Sharing a literature review rubric. *Journal of Library Administration, 45*(1), 185-202.

Hanna, M. A., & Smith., J. (1998). Using rubrics for documentation of clinical work supervision. *Counselor Education and Supervision, 37*(4), 269-278.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Humphry, S. M., & Heldsinger. S. A. (2014). Common structural design features of rubrics may represent a threat to validity. *Educational Researcher, 43*(5), 253-263.

Jonsson, A. (2010). The use of transparency in the 'interactive examination' for student teachers. *Assessment in Education: Principles, Policy & Practice, 17*(2), 185-199.

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*(1), 63-76.

Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment and Evaluation in Higher Education, 39*(7), 840-852.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144.

Kocakülah, M. S. (2010). Development and application of a rubric for evaluating students' performance on Newton's laws of motion. *Journal of Science Education and Technology, 19*(2), 146-164.

Linn, R. L., Baker, E., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Lucander, H., Knutsson, K., Salé, H., & Jonsson, A. (2012). 'I'll never forget this': Evaluating a pilot workshop in effective communication. *Journal of Dental Education, 76*(10), 1311-1316.

Orsmond, P., & Merry, S. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education, 21*(3), 239-250.

Panadero, E. (2011). *Instructional help for self-assessment and self-regulation: Evaluation of the efficacy of self-assessment scripts vs. rubrics* (Unpublished doctoral dissertation). Universidad Autónoma de Madrid, Spain.

Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and practical connotations. When it happens, how is it acquired and what to do to develop it in our students. *Electronic Journal of Research in Educational Psychology, 11*(2), 551-576.

Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2012). Rubrics and self-assessment scripts: Effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences, 22*(6), 806-813.

Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2014). Rubrics vs. self-assessment scripts: Effects on first year university students' self-regulation and performance. *Infancia y Aprendizaje: Journal for the Study of Education and Development, 37*(1), 149-183.

Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts: Effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation, 39*(3), 125-132.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited. *Educational Research Review, 9,* 129-144.

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy and Practice, 21*(2), 133-148.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*(4), 435-448.

Reitmeier, C. A., Svendsen, L. K., & Vrchota, D. A. (2004). Improving oral communication skills of students in food science courses. *Journal of Food Science Education, 3*(2), 15-20.

Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment Research and Evaluation, 15*(8), 1-9.

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education, 30*(2), 175-194.

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment and Evaluation in Higher Education, 34*(2), 159-179.

Schamber, J. F., & Mahoney, S. L. (2006). Assessing and improving the quality of group critical thinking exhibited in the final projects of collaborative learning groups. *The Journal of General Education, 55*(2), 103-137.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice, 14*(3), 281-294.

Wilson, M. (2006). *Rethinking rubrics in writing assessment*. Portsmouth, NH: Heinemann.

Yopp, B. D., & Rehberger, R. (2009). A curriculum focus intervention's effects on prealgebra achievement. *Journal of Developmental Education, 33*(2), 28-30.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135-147.