

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/234169756>

The Use of Scoring Rubrics for Formative Assessment Purposes Revisited: A Review

Article in Educational Research Review · January 2013

DOI: 10.1016/j.edurev.2013.01.002

CITATIONS

695

READS

18,930

2 authors:



[Ernesto Panadero](#)

Dublin City University

119 PUBLICATIONS 9,768 CITATIONS

[SEE PROFILE](#)



[Anders Jönsson](#)

Kristianstad University

81 PUBLICATIONS 3,916 CITATIONS

[SEE PROFILE](#)

The use of scoring rubrics for formative assessment purposes revisited: A review

Ernesto Panadero¹ & Anders Jonsson²

Author note

¹ Departament de Psicologia Bàsica, Evolutiva i de l' Educació, Universitat Autònoma de Barcelona.

² Department of Learning and Environment, Kristianstad University, Sweden.

Recommended citation:

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9(0), 129-144. doi:<http://dx.doi.org/10.1016/j.edurev.2013.01.002>

This is a pre-print of an article published in Educational Research Review. Personal use is permitted, but it cannot be uploaded in an Open Source repository. The permission from the publisher must be obtained for any other commercial purpose. This article may not exactly replicate the published version due to editorial changes and/or formatting and corrections during the final stage of publication. Interested readers are advised to consult the official published version.

Correspondence concerning this article should be addressed to: Ernesto Panadero, Departament de Psicologia Bàsica, Evolutiva i de l' Educació, Universitat Autònoma de Barcelona, Barcelona, 08193, Spain. E-mails: ernesto.panadero@gmail.com, anders.jonsson@hkr.se

Acknowledgements

Support for this research was provided by grants from the Spanish Education Department (Ref. SEJ2005-00994) and Alianza 4 Universidades to Ernesto Panadero. The authors would like to thank to Dr. J. Alonso-Tapia and Dr. F. Dochy for their feedback and encouragement to conduct this study.

Abstract

The mainstream research on scoring rubrics has emphasized the summative aspect of assessment. In recent years, the use of rubrics for formative purposes has gained more attention. This research has, however, not been conclusive. The aim of this study is therefore to review the research on formative use of rubrics, in order to investigate if, and how, rubrics have an impact on student learning. In total, 21 studies about rubrics were analyzed through content analysis. Sample, subject/task, design, procedure, and findings, were compared among the different studies in relation to effects on student performance and self-regulation. Findings indicate that rubrics may have the potential to influence students learning positively, but also that there are several different ways for the use of rubrics to mediate improved performance and self-regulation. There are a number of factors identified that may moderate the effects of using rubrics formatively, as well as factors that need further investigation.

KEYWORDS: Rubrics; formative assessment; self-assessment; peer-assessment; self-regulation; learning; self-grading; self-efficacy.

The use of scoring rubrics for formative assessment purposes revisited: A review

Rubrics are documents that articulate the expectations for an assignment, or a set of assignments, by listing the assessment criteria and by describing levels of quality in relation to each of these criteria (Reddy & Andrade, 2010) (see Appendix for a rubric sample). Since rubrics – by definition – are quite detailed scoring guides, they can be used in order to validly assess multi-dimensional performances (Andrade & Valtcheva, 2009; Halonen et al., 2003) and rubrics have been the primary choice for many test developers as they allow for, at least moderately, reliable assessment of complex performances (Jonsson & Svingby, 2007; Moskal & Leydens, 2000). The use of rubrics as a classroom assessment instrument has also increased in higher education (Simon & Forgette-Giroux, 2001) and they are widely used at the school level (Reddy, 2007).

However, the unique features of rubrics do not only make them suitable instruments for enhancing the psychometric properties of performance assessments, but also for supporting in the process of formative assessment, where assessment information is used to inform students about their progress and aid them in their development (Black & Wiliam, 2009; Wiliam, 2011). A major problem for the research on the formative uses of rubrics is that this research has not been conclusive regarding whether or not the use of rubrics actually enhances student performance. In one review of research on the use of rubrics, Jonsson and Svingby (2007) noted that it was not possible to draw any conclusions about student improvement since the results were mixed. While some studies suggest that student performance can be improved by involving the students in the use and development of rubrics (e.g. Andrade & Du, 2005; Hafner & Hafner, 2003; McCormick, Dooley, Lindner, & Cummins, 2007), other studies showed no differences in the quality of the work done by students with

and without rubrics (Reitmeier & Vrchota, 2009). A similar situation is reported by Reddy and Andrade (2010) in a review on the use of rubrics in higher education. As a consequence, we do not know how the use of rubrics may facilitate in improving student performance or which factors are important in moderating the potential effect. This article therefore aims to re-review current research on the use of rubrics, but with a primary attention on using rubrics for formative purposes.

Research on the formative use of rubrics

Advocates for the use of rubrics for formative assessment assume that rubrics can promote student learning, as well as lead to positive changes in instruction. This could be done in several different ways, for example in either a teacher- or a student-centered way. Regarding the former, by making assessment criteria explicit, rubrics can be used by the teacher to enhance the alignment of learning, instruction, and assessment, something that is often referred to as “constructive alignment” (Biggs, 1996).

In a student-centered approach, the rubric could be shared with the students in order to support student learning (Jonsson, 2008; Jonsson & Svingby, 2007). As seen in some studies, there can be quite dramatic effects on student performance when a rubric is used as an assessment tool for learning. For example, a marked positive effect on student performance is demonstrated in a study by Andrade (1999), where students in a science class self-assessed their work with the assistance of a rubric. Results show that the treatment group considerably outperformed the control group (effect size = 0.99).

Similarly, the students in a study by Brown, Glasswell, and Harland (2004) showed quite large improvements (effect size = 1.6). Here the rubric was used in a training program for writing, involving guidance in “meta-cognitive monitoring”. In

yet another study reporting on student improvement, the context was student laboratory write-ups. The writing was supported by a rubric as well as peer-editing sessions and self-assessment. On an average, the scores of the write-ups improved by 17 percent in this study (Mullen, 2003). In a study about mathematical problem solving in upper-secondary school (Balan, 2012), the performance of students who had been working with peer assessment and rubrics was significantly higher as compared to the performance for students in a control group (effect size = 1.43). The last example is a study by Schamber and Mahoney (2006), where the combination of writing assignments and the use of a rubric improved the scores in an assessment of critical-thinking skills by 41 percent. Although few, these studies indicate that rubrics might be valuable in supporting student learning, at least in combination with different meta-cognitive activities (such as self-regulation, self- or peer assessment).

Some studies also show that students actually internalize the criteria in the rubric, making them their own, and use them while self-assessing (Andrade, 1999; Piscitello, 2001). There exist, however, also studies presenting results which are not as straightforward. Goodrich Andrade has conducted a couple of studies on writing performance, where the students were supported by a rubric and self-assessment, and she reports positive effects only for some of the groups investigated – and the results sometimes differ between gender (Goodrich Andrade, 2001; Goodrich Andrade & Boulay, 2003).

To conclude, the use of rubrics has been shown to mediate improved student performance. However, the research on the use of rubrics is not unanimous. While some studies report on (very) large effects on student performance, others present only small changes or none at all. While some studies report on improvements for all students or for all criteria, others present changes only for some of the students,

assignments, or criteria investigated. These findings lead to questions about how rubrics may support student learning and which factors influence the effects of using rubrics for formative purposes. In order to guide future research in this area, research on the use of rubrics is (re-) reviewed in this study, with a particular focus on formative uses of rubrics. By comparing the results of different studies, in relation to different parameters (such as educational level, assignments, research design, etc.), factors that may influence the effects of using rubrics formatively – either positively or negatively – could be identified. Specifically, we aim to answer the following questions:

1. In what ways can the formative use of rubrics mediate improved student performance?
2. Which factors may moderate – either positively or negatively – the effects of using rubrics formatively?

Methodology

Criteria for inclusion

Studies from all kinds of disciplines were reviewed and included or rejected based on their relevance. First, a study was considered relevant for our research if it was based on empirical data on the use of rubrics. Theoretical arguments or articles presenting only practical guidelines were not included in the review, although they may have been used for the theoretical framework. Second, the use of rubrics had to be for formative purposes. Articles reporting on summative uses of rubrics (e.g. framework inter-judges reliability) were only selected if there was also a formative use (e.g. allowing the students' to reflect on different performances based on their grades). Third, the selection was limited to include only printed and peer-reviewed

material, such as articles in journals, edited books, research reports, and doctoral dissertations.

Search terms and databases

A literature search was conducted via the PsycINFO and ERIC databases, searching with no limitation on the year of publication. The following combinations of keywords and metacognitive activities were used: self-assessment & rubric/s; peer-assessment & rubric/s; self-regulation/self-regulated learning & rubric/s. These combinations produced 57 hits. After analyzing the abstracts and main conclusions of these articles only 17 were included. The main reasons for rejecting articles were that the studies focused exclusively on summative uses of rubrics (e.g. increasing inter-rater reliability), that they were non-empirical, or that the definition of rubrics diverged radically from the definition used in this study. Furthermore, by applying the so called “snow-balling method”, 4 additional studies were found and included in the review.

Method of analysis

The analysis performed in this review is of a qualitative nature. The main purpose is to identify a spectrum of potential effects of using rubrics formatively, as well as different factors, which may moderate these effects. We are therefore not primarily concerned with the relative strength of these effects or the amount of scientific evidence substantiating each of them.

The reason for taking this stance is partly the restricted number of studies on this topic in total (and for individual factors in particular), but also the great heterogeneity of the studies, which (together with the scarcity of data provided in most articles) excludes the possibility to perform a meta-analysis. The approach

adopted is therefore a narrative content analysis (Dochy, 2006), where it is the breadth of different categories that is of interest.

In a first step, the different articles were read in order to determine whether they contained relevant information and whether they fulfilled the inclusion criteria. In a second step, for each selected article the following information was extracted and entered into a table: sample, subject/task, research design, procedures, and findings (see Table 1). Thirdly, the information provided by the different studies was compared in order to explore possible patterns. This comparison was done for each of the parameters, for instance comparing the results of studies with respect to research design, intervention time, subject/task, etc. All steps were performed independently by each of the authors and then compared and discussed.

Results

In total, 21 studies that met the inclusion criteria were found, most of them investigating the effects of using rubrics for self-assessment purposes. The majority of these studies were published in journals focusing specifically on assessment issues, followed by journals on educational topics in general. A few studies were published in journals for a more specific audience, such as food-science education or library administration. Also, a couple of studies were included in a doctoral dissertation, where all of the individual studies had not yet been published. As suggested by these findings, the studies vary across subject (mostly English, but also algebra, educational psychology, social studies, etc.), and across educational level (predominantly 7th and 8th graders, but also students in grade 3-6 and under-graduate students). The performance assessed and supported by the rubrics was for the most part essay writing, but also different projects and oral presentations.

INSERT TABLE 1

How the use of rubrics may facilitate student performance

There are several ways for the formative use of rubrics to mediate improved student performance (see Fig. 1), such as increasing transparency, reducing anxiety, aiding the feedback process, improving student self-efficacy, or supporting student self-regulation.

Increasing transparency

An important prerequisite for improving students' performance is by letting them know what is expected from them (Good, 1987) and as can be seen from the studies reviewed, rubrics can aid in communicating expectations to the students. For instance, in the study by Andrade and Du (2005), students contrasted the frustration of not knowing their teachers' expectations with the relief or insight provided by a rubric. Similar results are reported by Reynolds-Keefer (2010), where the students claimed that they better understood teacher expectations when the assignment involved a rubric, and by Schamber and Mahoney (2006), where students' responses indicated that they found the rubrics useful for clarifying the components and requirements of the assignment. In the same study, a majority of the faculty members were also found to perceive the rubrics to be "very useful" in clarifying assignment requirements (cf. Reddy & Andrade, 2010). Furthermore, questionnaires administered at the end of a study by Goodrich Andrade (2001) revealed that students, after using a rubric, tended to identify more of the criteria by which their performance was evaluated.

As indicated above, the results from studies investigating the transparency provided by the use of rubrics are quite unanimous in suggesting that rubrics may aid

in clarifying expectations. Although acknowledging that there are results based on students' perceptions only (Andrade & Du, 2005; Reynolds-Keefer, 2010), there are also studies that investigated quantitative measurements of student performance together with either students' perceptions (Schamber & Mahoney, 2006) or student identification of teacher's expectations (Goodrich Andrade, 2001).

Reducing anxiety

Clarified expectations, as provided by rubrics, may also support student performance by lowering students' anxiety regarding assignments (Kuhl, 2000; Wolters, 2003). For instance, as reported by Andrade and Du (2005), the students spoke about increased confidence and making it easier to hand in assignments when having a rubric. These results are based on students' perceptions when asked about anxiety directly in relationship with rubrics.

The work by Panadero (2011), and Panadero, Alonso-Tapia, and Huertas (2012) has also shed some light on this particular topic. These authors used a performance/avoidance self-regulation scale measuring students' self-regulatory actions motivated by negative emotions, such as anxiety. In studies performed in higher education, rubrics significantly decreased this type of negative self-regulation (for instance that students do not finish a task because they get blocked), in two out of three studies. The authors therefore conclude that students' anxiety may decrease when implementing long-term interventions with rubrics, which is probably due to the fact that students know what is expected of their work and how it will relate to their grades (Panadero, 2011).

Aiding the feedback process

Another manner for the use of rubrics to contribute to student learning is by aiding the feedback process, which has been deemed useful by teachers and students

alike. In relation to the former, Schamber and Mahoney (2006) report that a majority of faculty members found the rubrics to be “very useful” for providing students with feedback on drafts. In relation to the latter, Andrade and Du (2005) report on how students reflected on their feedback by reviewing the scores they received from the teacher with the help of the rubric. Again, these results are based on students’ perceptions only.

Improving student self-efficacy

As has been suggested by previous research (Pajares, 2008), the performance of students with high levels of self-efficacy tend to be higher, as compared to those who report lower levels of self-efficacy. In the study by Andrade, Wang, Du, and Akawi (2009) generating a list of criteria from a model essay and using a rubric to self-assess drafts, was shown to improve students’ self-efficacy. The relation to performance was not, however, investigated in this particular study.

Investigating the relationship between self-efficacy and rubric use was one of the aims in the research performed by Panadero and colleagues (Panadero, 2011; Panadero, Alonso-Tapia, & Huertas, 2012). In this research, self-efficacy was impacted by the use of rubrics, but only in one of the studies. However, as suggested by van Dinther, Dochy, and Segers (2010), that students may need to be confronted with teachers’ feedback regarding their performance in order to have a realistic opinion about their advances (or lack of it). Which was only controlled in Panadero’s research in the study with significant results (Panadero et al., 2010).

INSERT FIGURE 1

Supporting student self-regulation

Developing an adequate sense of your own knowledge and skills is one aspect of owning your learning. Planning your work, as well as continuously self-assessing the work in progress, are other important aspects of self-regulated learning (Efklides, 2011; Zimmerman, 1990). Rubrics have been shown to facilitate both planning and self-assessment (Panadero, 2011). For instance, in relation to planning, the students in the study by Andrade and Du (2005) reported using the rubric to plan their approach to an assignment, much like a recipe or a map. Students' responses in the study by Reynolds-Keefer (2010), a replication of the study by Andrade and Du (2005), also indicate that rubrics aided the students in both planning and in the production of 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 (see also Reitmeier & Vrchota, 2009).

Indications that rubrics may aid the students in checking their work in progress (i.e. self-assess) are found in Andrade and Du (2005), where students reported doing a lot of informal self-assessment in addition to the formal rubric-referenced self-assessment required by the course. This self-assessment was facilitated by the rubric, which was used – as expressed by one of the students: “before, during and after to make sure I had everything covered” (p. 4). Reynolds-Keefer (2010) also reports that most students claimed that the rubric was helpful in reflecting on their work before submitting it (see also Andrade, Buff, Terry, Erano, & Paolino, 2009).

In a series of studies by Panadero et al. the effects of using rubrics were compared against other conditions on self-regulation. In a first study, Panadero et al. (2012) compared rubrics with self-assessment scripts (i.e. a list of specific questions

structured in steps to follow the expert model of approaching a task from beginning to end) in secondary education, reporting that the rubric and script groups outperformed the control group when self-regulation was measured through thinking-aloud protocols (i.e. verbalizations of thinking as students proceed through the task). In a second study, Panadero (2011) started from the same conditions as in the previous study, but in higher education. This time they found that the use of rubrics decreased the level of “performance/avoidance self-regulation”, which refers to actions motivated by negative emotions (e.g. anxiety), a positive result because decreasing this type of self-regulation typically increases students’ learning. Scripts, on the other hand, benefited “learning self-regulation”, such as focusing on correcting mistakes. In a third study, also with higher-education students, Panadero (2011) found that the use of rubrics once again had a positive impact by decreasing “performance/avoidance self-regulation”, while scripts enhanced “learning self-regulation”. These two studies were conducted in during a semester-long intervention. In a similar line of research, Reitmeier and Vrchota (2009) compared the use of a rubric with the use of a reflection sheet, but in this study no significant differences between these two groups were found.

Factors moderating the effects of using rubrics formatively

By comparing different parameters reported in the studies reviewed, a couple of factors have been identified, which seem to moderate the effect of using rubrics for formative purposes. There are also some factors identified which, according to the research reviewed, do not seem to influence the effects of using rubrics. Finally, there are factors identified which have not yet been systematically investigated or where the current research is still inconclusive.

Combinations of rubrics and meta-cognitive activities

A severe limitation to the conclusion that the use of rubrics may aid in improving student performance is the fact that most studies reporting on such improvements have combined the use of rubrics with other instructional interventions. However, although this means that there is limited evidence supporting the claim that the use of rubrics can in itself lead to improvements in performance, there are more substantial evidence suggesting that rubrics can facilitate improvements if combined with self-assessment or other meta-cognitive activities (Andrade, 2010; Halonen et al., 2003). Furthermore, the incapacity to provide such conditions is proposed, by the researchers, as explanations for the moderate or partial effects noted in some studies (Goodrich Andrade, 2001; Goodrich Andrade & Boulay, 2003). Examples of (more or less) successful combinations of rubrics and such activities are formal self-assessment lessons (Goodrich Andrade & Boulay, 2003), self-assessment of a first draft (Andrade et al., 2010; Andrade et al., 2008), peer and self-assessment (Andrade, Buff, et al., 2009; Panadero, 2011; Panadero et al., 2012; Sadler & Good, 2006), and self-assessment and exemplars (Jonsson, 2010). Furthermore, Brown et al. (2004) used a rubric in combination with explicit instruction and modeling, as well as training in meta-cognitive monitoring and scaffolding of writing.

Educational level and length of intervention

While the studies performed in higher education contexts tend to report on positive results when providing the students with rubrics (Jonsson, 2010; Petkov & Petkova, 2006; Schamber & Mahoney, 2006), longer and larger interventions are typically needed in order to produce positive results in schools. For instance, in studies where the rubric was introduced during one period only, or where the students got only a couple of lessons in self-assessment, the effects reported are small and only partial (Goodrich Andrade, 2001; Andrade & Boulay, 2003). In other studies, where

the interventions are somewhat longer (involving at least 5 periods), there is a clearer effect of the treatment (Andrade et al., 2008; Andrade et al., 2010). However, the largest gains are found in studies lasting for several weeks (Andrade, Buff, et al., 2009; Brown et al., 2004; Sadler & Good, 2006). Time devoted to work with the rubric therefore seems more crucial for younger students; a finding that might not be all too surprising.

Gender

Another investigated factor is student gender. While the interaction between treatment and gender is not significant in some studies (Andrade et al., 2010; Andrade et al., 2008), in other studies the use of rubrics has different effects on boys and girls (Goodrich Andrade, 2001; Goodrich Andrade & Boulay, 2003). Even the self-efficacy ratings of boys and girls are affected differently by long-term rubric use (Andrade, Wang, Du, & Akawi, 2009), where girls in the treatment group reported higher self-efficacy than boys. This effect was explained by girls being more confident when they created their own feedback, a conclusion supported by previous research (Pajares & Valiante, 1999). Future research is needed in order to clarify the effects of using rubrics in interaction with gender.

Topic/performance

Although the performance assessed in the majority of cases is student writing (Andrade, Buff, et al., 2009; Andrade, Du, & Mycek, 2010; Andrade, Du, & Wang, 2008; Brown et al., 2004; Goodrich Andrade, 2001; Goodrich Andrade & Boulay, 2003), there are several other performances that have been shown to improve through the use of rubrics. Such performances are students' ability to self-assess (Andrade, Buff, et al., 2009; Jonsson, 2010), pre-service teachers' analyses of classroom situations (Jonsson, 2010), design of conceptual maps and multimedia material for

pedagogical purposes (Panadero, 2011), project design (Petkov & Petkova, 2006), results on subject-matter tests (Sadler & Good, 2006; Schafer, Swanson, Bené, & Newberry, 2001), critical thinking (Schamber & Mahoney, 2006), and landscape analysis in secondary education students (Panadero et al., 2012). As suggested by the diversity of this research, the choice of topic does not seem to greatly influence the effects of using of rubrics.

Other factors

There is also a number of other factors investigated, but each of these are represented by one or two studies only, making it impossible to identify any patterns or draw any conclusions. Such factors are: different subject matter (Schafer et al., 2001), contextual factors (e.g. testing requirements interfering with the research; Andrade, Buff, et al., 2009; Goodrich Andrade, 2001; Goodrich Andrade & Boulay, 2003), students' time to revise (Goodrich Andrade, 2001); socio-economic factors (Andrade, Buff, et al., 2009), and how accessible the language in the rubric is (Goodrich Andrade, 2001).

Discussion

The aim of this literature review was to investigate how the formative use of rubrics may mediate improved performance, as well as to identify factors that may moderate the effects of using rubrics formatively in order to guide future research in this area.

As suggested by the findings in this review, the use of rubrics may mediate improved performance through (a) providing transparency to the assessment, which in turn may (b) reduce student anxiety. The use of rubrics may also (c) aid the feedback process, (d) improve student self-efficacy, and (e) support student self-regulation; all of which may indirectly facilitate improved student performance. Consequently, there

are a number of different, but not necessarily unconnected, ways in which the use of rubrics may support student learning.

Starting with transparency, studies in which scoring rubrics are shared with the students, the students tend to appreciate the transparency provided. This is an interesting finding in itself, because there are a number of other studies (i.e. studies not involving rubrics) stressing the difficulties students encounter when trying to decipher assessment criteria and feedback (e.g. Orsmond, Merry, & Reiling, 1996, 1997; Price, Handley, Millar, & O'Donovan, 2010). So why would this be easier with rubrics, as compared to for instance feedback, where many students have been shown to have great problems understanding the academic discourse (Jonsson, 2013)? One possible explanation could be that the format of rubrics, with both criteria and levels of attainment, more easily conveys information to students and therefore provides greater transparency. This explanation is, however, inconsistent with studies where rubrics have been handed out to the students, but without any specific training (e.g. Duke, 2003; Toth, Suthers, & Lesgold, 2002). In such studies, there are typically no significant effects on student performance, which indicates that it is not the format that facilitates improved student performance. Instead, the findings of this review suggest that the transparency provided by the rubrics (as perceived by the students), can reduce students' anxiety and influence students' self-regulation strategies, for instance by decreasing the level of "performance/avoidance self-regulation". Since "performance/avoidance self-regulation" means that students try to avoid negative consequences, such as being perceived as ignorant, the transparency provided by rubrics may reduce the need for such approaches, which in turn may have an effect on student performance. However, this effect is also linked to the time factor, since high

expectations (as by the rubric) have been shown to cause stress if there is not enough time to meet the standards.

Taken together, the use of rubrics may – by making the expectations explicit – provide some kind of confidence to the students, which in turn have an effect on their performance. This is verified by studies investigating students' perceptions of using rubrics. Other instruments, such as scripts, may have effects on students' self-regulation, which are indeed more learning oriented as compared to the effects by using rubrics. Still, the extrinsic focus on assessment and grading, which is often associated with rubrics (although this does not need to be the case), seems to exert a greater influence on students' self-regulation than the use of self-assessment scripts. Possible implications of using rubrics may therefore be, on the one hand, to provide a feeling of confidence and control, but on the other hand make students more performance oriented.

An extrinsically driven performance orientation, together with increased transparency, may help students to more effectively deliver what the teacher wants, instead of relying on trial-and-error. This means that the students may focus on what is considered important and hence improve their performance. However, this is also the very effect of rubrics that is criticized by some authors, for instance by arguing that transparency encourages instrumentalism (e.g. Torrance, 2007). Although a relevant argument, this instrumentalism need not be a consequence of transparency, but rather an attitude that is always present among students to some extent (cf. Biggs, 1999), even if it might be more apparent when the assessment criteria are known to the students. A more optimistic interpretation of this instrumentalism is therefore that the transparency allows some (performance oriented) students to strategically reach the minimum level of performance, but that this does not necessarily affect other

(learning oriented) students negatively. This could be compared to some research in which formative assessment has been seen to particularly favor low-achieving students (Balan, 2012; Black & Wiliam, 1998). It would therefore be of great interest to perform studies on the use of rubrics, where differences between students with different learning orientations are investigated.

Besides reducing students' anxiety and decreasing the level of "performance/avoidance self-regulation" through perceived transparency, there are indications that rubrics may aid students in regulating their learning. Examples are studies showing that students use rubrics to reflect on their feedback, to plan their assignments, to check their progress, and to review their work before submission, which may also have an effect on students' performance. This means that rubrics may be a helpful tool for student self-regulation.

The second question guiding this review was which factors that may moderate the effects of using rubrics formatively. As suggested by the findings, the effects of using rubrics may be moderated by using rubrics in combination with meta-cognitive activities, by educational level, and by the time devoted to learn how to use the rubric. In some studies, there are also gender differences. However, as suggested by the review, the choice of topic does not seem to greatly influence the effects of using of rubrics. Furthermore, positive effects have reported for rubrics of different formats, which suggest that rubric design is not a crucial feature. Still, this factor has not been systematically investigated.

That the use of rubrics is combined with meta-cognitive activities, such as self-assessment, in almost all studies reviewed, makes it difficult to draw any conclusions about the influence of such activities. Whether the use of rubrics has to

involve some kind of meta-cognitive activity, or whether any activity which allows for more time on task would suffice, therefore remains unclear.

What does seem clear, though, is that it may take quite comprehensive interventions in order to (so to speak) get the most out of using rubrics. In principle, the effect of using rubrics seems to increase the longer the implementation is. This is true at least at the school level, while the performance of students in higher education does not seem to be quite as dependent upon the length of the intervention. Still, there are too few studies to get a clear picture.

Futures lines of research

As expressed above, this article re-reviewed current research on the formative use of rubrics, with the main aim to guide future research in this area. As a consequence, we outline a number of topics for future research, which have emanated from the review.

First, although a number of hypothetical features of rubric design could most probably be deducted from research on formative assessment, there seems to be no studies on the effect of different designs (such as holistic versus analytical, few levels versus several levels, or task specific versus generic). This could be addressed by studies using an experimental, or a quasi-experimental, research design, where different rubric formats are compared.

Second, effects of personal differences on the use of rubrics, such as goal orientation, prior academic performance, and self-regulation strategies, remain largely unknown. For example: Are students with a learning orientation hampered by the use of rubrics, for instance by becoming more instrumental and less creative? Does the use of rubrics differ among students with different goal orientations? Do students with higher levels of self-regulation benefit more from the use of rubrics as compared to

students with lower levels? These questions could be addressed by studies using a mixed methods research design, where the goal orientations of students using rubrics are measured with questionnaires, and their perceptions and study strategies are explored using qualitative methods (for example thinking-aloud protocols and interviews). Such a line of research could add considerable information about how to successfully implement the use of rubrics for formative purposes.

Third, there are indications of rubrics being used as a tool for self-regulation, for instance by students planning and reviewing their work with the aid of rubrics. This research, however, is based almost entirely on students' statements except for Panadero et al. (2012), which means that there is a need for research to explore how students actually use rubrics for such purposes. Again, this could be addressed through the use of thinking-aloud protocols, but also by analyzing students' work at different stages before submission.

Fourth, more research is needed on the confounded effects of using rubrics together with different meta-cognitive activities. It is still unclear, which activities are needed in order to influence the effects of using rubrics. Specifically, the effects of using peer-assessment in combination with rubrics have not been much explored (see Table 1), and mainly for summative purposes. This question too, could be addressed by studies using an experimental, or a quasi-experimental, design, where the use of rubrics is combined with different activities (such as feedback, self- and peer assessment) and then compared.

Fifth, gender is an important individual factor that has been investigated, but with mixed results. Whereas Andrade, Wang, et al. (2009) found significant differences related to gender, Andrade et al. (2010) did not. Thus it is important to continue investigating the influence of gender, in order to further clarify the

relationship between gender and the use of rubrics. This question could be addressed by studies using an experimental, or a quasi-experimental, design, where the results of boys and girls are compared. However, in order to investigate the reasons for the gender differences, possible moderating variables (such as motivation and confidence) also need to be investigated.

Finally, a particularly important issue to be addressed in future research is research design and the data reported. About half of the studies reviewed have used a quasi-experimental design (which may be considered high-quality in this context, as it keeps the research close to classroom practice), one study has used an experimental design (Panadero et al., 2012), and the rest has used descriptive designs. Whereas a descriptive design is not a flaw *per se*, it is very difficult to draw any firm conclusions about rubric efficiency from such research. Another problematic issue is the way data is reported. In several cases there are insufficient information about the participants, the procedure, or (which is especially troublesome) about the analyses made. For instance, effect-size data is often missing, which is crucial for evaluating the impact of the intervention and the absence of such data prevented us from performing a meta-analysis. We therefore strongly recommend future research on the use of rubrics to pay special attention to these matters, in order to improve the quality of the research in this field.

References

- Andrade, H. (1999). The role of instructional rubrics and self-assessment in learning to write: A smorgasbord of findings. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.
- Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. J. Andrade

- & G. J. Cizek (Eds.), *Handbook of formative assessment*. New York: Routledge.
- Andrade, H., Buff, C., Terry, J., Erano, M., & Paolino, S. (2009). Assessment-driven improvements in middle school students' writing. *Middle School Journal*, 40(4), 4–12.
- Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, 10(3), 1–11. Retrieved from <http://pareonline.net/getvn.asp?v=10&n=3>.
- Andrade, H., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice*, 17(2), 199–214. <http://dx.doi.org/10.1080/09695941003696172>.
- Andrade, H., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices*, 27(2).
- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice*, 48(1), 12–19.
- Andrade, H., Wang, X. L., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research*, 102(4), 287–301.
- Balan, A. (2012). *Assessment for learning: A case study in mathematics education*. Doctoral dissertation. Malmö University, Malmö, Sweden.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <http://dx.doi.org/10.1007/bf00138871>.

- Biggs, J. (1999). *Teaching for quality learning at University*. Buckingham: SRHE and Open University Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7–74.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment Evaluation and Accountability*, 21, 5–31.
- Brown, G. T. L., Glasswell, K., & Harland, D. (2004). Accuracy in the scoring of writing: Studies of reliability and validity using a New Zealand writing assessment system. *Assessing Writing*, 9(2), 105–121.
<http://dx.doi.org/10.1016/j.asw.2004.07.001>.
- Dochy, F. (2006). A guide for writing scholarly articles or reviews for the Educational Research Review. *Educational Research Review*, 4, 1–2. Retrieved from <http://www.journals.elsevier.com/educational-research-review/>.
- Duke, B. L. (2003). The influence of using cognitive strategy instruction through writing rubrics on high school students' writing self-efficacy, achievement goal orientation, perceptions of classroom goal structures, self-regulation, and writing achievement. Unpublished doctoral dissertation. University of Oklahoma, USA.
- Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, 46(1), 6–25.
- Good, T. L. (1987). Two decades of research on teacher expectations: Findings and future directions. *Journal of Teacher Education*, 38(4), 32–47.
<http://dx.doi.org/10.1177/002248718703800406>.

Goodrich Andrade, H. (2001). The effects of instructional rubrics on learning to write.

Current Issues in Education, 4(4). Retrieved from

<http://cie.ed.asu.edu/volume4/number4/>.

Goodrich Andrade, H., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research*, 97(1), 21–34.

Green, R., & Bowser, M. (2006). Observations from the field: Sharing a literature review rubric. *Journal of Library Administration*, 45(1), 185–202.
http://dx.doi.org/10.1300/J111v45n01_10.

Hafner, O. C., & Hafner, P. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education*, 25(12), 1509–1528.
<http://dx.doi.org/10.1080/0950069022000038268>.

Halonen, J. S., Bosack, T., Clay, S., McCarthy, M., Dunn, D. S., Hill Iv, G. W., & Whitlock, K. (2003). A rubric for learning, teaching, and assessing scientific inquiry in psychology. *Teaching of Psychology*, 30(3), 196–208.
http://dx.doi.org/10.1207/s15328023top3003_01.

Jonsson, A. (2008). Educative assessment for/of teacher competency. Doctoral dissertation. Malmö University, Malmö, Sweden.

Jonsson, A. (2010). The use of transparency in the “Interactive Examination” for student teachers. *Assessment in Education: Principles, Policy & Practice*, 17(2), 183–197.

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

- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144.
- Kuhl, J. (2000). A functional-design approach to motivation and self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 111–169). San Diego, CA: Academic Press.
- McCormick, M. J., Dooley, K. E., Lindner, J. R., & Cummins, R. L. (2007). Perceived growth versus actual growth in executive leadership competencies: An application of the stair-step behaviorally anchored evaluation approach. *Journal of Agricultural Education*, 48(2), 23–35.
- Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10). Retrieved from <http://PAREonline.net/getvn.asp?v=7&n=10>.
- Mullen, Y. K. (2003). Student improvement in middle school science. Master's thesis, University of Wisconsin.
- Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education*, 21(3), 239–250. <http://dx.doi.org/10.1080/0260293960210304>.
- Orsmond, P., Merry, S., & Reiling, K. (1997). A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment & Evaluation in Higher Education*, 22(4), 357–368. <http://dx.doi.org/10.1080/0260293970220401>.
- Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning. Theory, research and applications* (pp. 111–168). New York: Lawrence Erlbaum Associates.

- Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology*, 24(4), 390–405. <http://dx.doi.org/10.1006/ceps.1998.0995>.
- Panadero, E. (2011). Instructional help for self-assessment and self-regulation: Evaluation of the efficacy of self-assessment scripts vs. rubrics. Doctoral dissertation. Universidad Autónoma de Madrid, Madrid, Spain.
- 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. <http://dx.doi.org/10.1016/j.lindif.2012.04.007>.
- 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. doi:10.1016/j.stueduc.2013.04.001
- 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. doi:10.1080/02103702.2014.881655
- Petkov, D., & Petkova, O. (2006). Development of scoring rubrics for IS projects as an assessment tool. *Issues in Informing Science and Information Technology*, 3, 499–510.
- Piscitello, M. E. (2001). Using rubrics for assessment and evaluation in art. Master's thesis. Saint Xavier University, Chicago (USA).

- Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, 35, 277–289.
- Reddy, Y. M. (2007). Effects of rubrics on enhancement of student learning. *Educate*, 7(1), 3–17.
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435–448.
<http://dx.doi.org/10.1080/02602930902862859>.
- Reitmeier, C. A., & Vrchota, D. A. (2009). Self-assessment of oral communication presentations in food science and nutrition. *Journal of Food Science Education*, 8(4), 88–92.
- Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment Research & Evaluation*, 15(8). Retrieved from <http://pareonline.net/getvn.asp?v=15&n=8>.
- Sadler, P. M., & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, 11(1), 1–31.
- Schafer, W. D., Swanson, G., Bené, N., & Newberry, G. (2001). Effects of teacher knowledge of rubrics on student achievement in four content areas. *Applied Measurement in Education*, 14(2), 151–170.
- 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.
<http://dx.doi.org/10.1353/jge.2006.0025>.

- Simon, M., & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. *Practical Assessment Research & Evaluation*, 7(18).
Retrieved from <http://PAREonline.net/getvn.asp?v=7&n=18>.
- 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, 281–294.
- Toth, E. E., Suthers, D. D., & Lesgold, A. M. (2002). “Mapping to know”: The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education*, 86(2), 264–286.
<http://dx.doi.org/10.1002/sce.10004>.
- van Dinther, M., Dochy, F., & Segers, M. (2010). Factors affecting students’ self-efficacy in higher education. *Educational Research Review*, 6(2), 95–108.
<http://dx.doi.org/10.1016/j.edurev.2010.10.003>.
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37, 2–14.
- Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist*, 38(4), 189–205.
- Zimmerman, B. J. (1990). Self-regulated learning and academic-achievement. An overview. *Educational Psychologist*, 25(1), 3–17.
http://dx.doi.org/10.1207/s15326985ep2501_2.

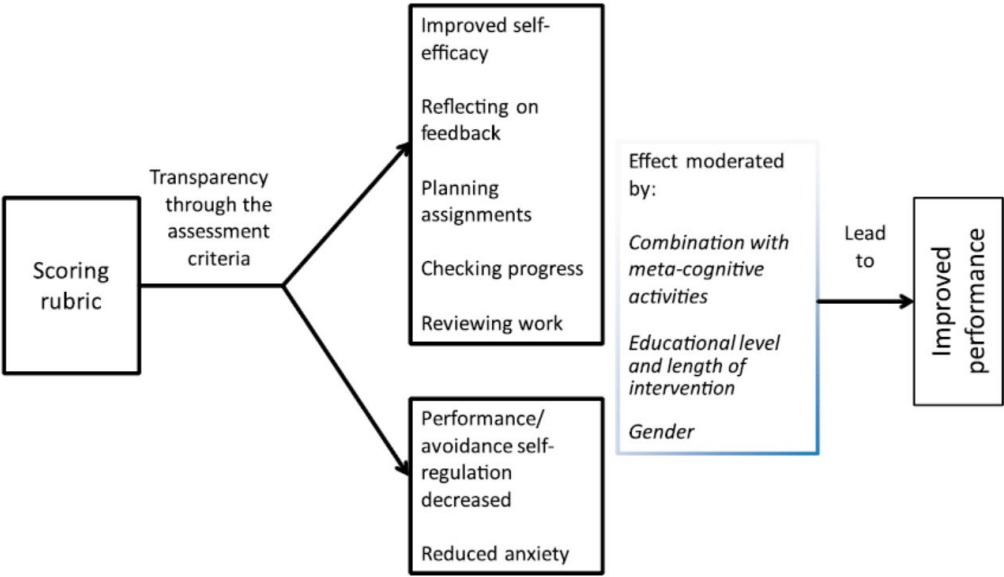


Fig. 1. Rubric and moderating effects that improved performance.

Table 1. Studies using rubrics for formative purposes

Study	Sample	Subject	Design	Type of aid	Procedure	Findings and interpretations	Outcomes
Goodrich Andrade, 2001 J. Current Issues in Education	9 groups of 8 th grade students in two middle schools in California. Total of 242 students. Unknown control group N.	Writing	Quasi-experimental (treatment and control groups)	Writing rubric	Three essays were written. Treatment group worked with a rubric. After writing the essays, students were asked to explain what criteria they think their teachers used to evaluate their writings. Then, the three essays were scored.	Rubric only had a significant effect on the second essay score. In the third essay there was a significant interaction effect with gender: girls in the treatment group performed less well than girls in the control group. The lack of significant effect in 1 st essay was explained in terms of novelty: was the first time for many of the participants and the rubric was complicated. Students in the treatment group showed a better understanding of the assessment criteria. However, this did affect to their scores as the intervention effect was not significant. Thus handing out and explaining a rubric can increase students' knowledge of the criteria for writing, while transforming that knowledge into actual writing is more demanding	Essay scores Assessment criteria
Schafer et al. 2001 J. Applied Measurement in Education	46 dyads (92 teachers). No age data. In each dyad one member was control, so a total of 46.	Algebra, biology, English, and government	Quasi-experimental (treatment and control groups)	Specific rubric for the four subjects	All teachers participated in a first general session. Only the treatment group participated in the second session, where they were taught about rubrics. Not clear if there was a control when the rubric had been used in the classroom. Afterwards, students' learning was evaluated through two exams.	Data was analyzed through a meta-analysis. Each teacher was considered an independent study. The rubric effect only was significant in algebra (for both exams) and biology (one exam). The use of a rubric was positive for the teachers, as they had a clearer idea of their own assessment criteria.	Students' learning, but based on the teacher's knowledge of rubrics and not on students' use of the instrument.
Andrade & Boulay, 2003 J. Journal of Educational Research	N=397. 183 7 th graders, 214 8 th graders. 191 boys, 201 girls (5 unidentified).	Writing (English Language Arts)	Quasi-experimental (treatment & baseline)	Rubric for making essays	The two groups received a rubric and wrote two essays. Control group were just handed the rubric, while the treatment group had two sessions with the researcher in which they were taught how to self-assess their work and look for the criteria with color pencils.	The purpose of this study was to compare conditions where students were just handed rubrics and where they received specific rubric training. No significant effect of the training, but the tendency favored treatment group. There was a significant interaction effect of gender, where girls in the treatment group had better scores than girls in the control.	Outcome: essay scores
Hafner & Hafner, 2003 J. International	107 university biology students.	Biology oral presentation	Quasi-experimental with post measure.	Rubric for an oral presentation	3 consecutive academic years using the rubric. It was handed out at the beginning of the year and students were encouraged to use it for their	According to the researchers there was a better learning, especially for the last groups due to watching the others presentations. Peer-assessment grades were highly correlated to teacher's grades, especially for	Mainly peer-assessment Self-assessment only counted in an informal way

Journal of Science Education	Only 61 submitted the self-assessment part. No age data.		Empirical with no control group.		work. Students peer-assessed each other's oral presentations. Students also self-assessed their work, but this did not count towards their grades.	presentations with good grades. Also, self-assessment grades were similar to teacher's assessment. Students said they had a better understanding with rubric.	Focus on reliability and validity issues of peer assessment as compared to teacher assessment.
Brown, Glasswell, & Harland, 2004 J. Assessing Writing	Study 2 "Just over 60 students, in four classes in three schools"	Argumentative writing	Mainly descriptive	It is unclear how the rubric was used during the intervention. There was also "metacognitive training and modeling", so the intervention results are not based solely on the effects of using the rubric.			
Andrade & Du, 2005 J. Practical Assessment, Research & Evaluation	14 pre-service teachers No age data	Educational psychology	Descriptive study with focus group discussions about using rubrics. Qualitative data. No control.	Rubric created with the students	Students worked with a rubric they created in an educational psychology course. Afterwards, they discussed the use of the rubric in focus groups.	Students perceived that the rubric helped them to set the task goals, to plan, to supervise their work, to reflect about feedback, to achieve better grades, to improve the quality of their work, and to be less anxious.	Students' perceptions about rubrics and its role for their learning No outcome measures
Green & Bowser, 2006 J. Journal of Library Administration	No sample data	Literature review	Descriptive study No control group	Rubric for a literature review project	Training students to do a good literature review.	Descriptive study in which the authors develop a model to help students write better literature reviews.	See procedure
Petkov & Petkova, 2006 J. Issues in Informing Science and Information Technology	40 undergraduate students. No age data.	Management of business information	Quasi-experimental design. One "control" & one treatment.	Rubrics for project assessment in "Management of Business Information" and rubrics for project assessment in "Systems Analysis and Design"	One group worked with rubric, the other without. They presented the final project.	The mean percentage grade for the group using rubrics was higher than the mean for the group that was not using rubrics.	Outcome: project score Problems with the design and conclusions.

Sadler & Good, 2006 J. Educational Assessment	Four 7 th grade general science classrooms. The 1 st test was taken by 101 students, the second by 95.	Biology unit on the classification of organisms.	Quasi-experimental design. Control group, self-assessment & peer-assessment.	Rubric created with the students. Not clear if all the students participated in rubric construction and whether this is in fact what affected their learning.	The four groups did a biology test and two groups peer assessed, one self-assessed and the other was control. One week later the same test was performed and the four groups peer assessed. The teacher for the 4 groups was the 2 nd author. No more teachers involved.	There was a high correlation between teacher's and students' self-assessment grades. Teacher and peer-assessment grades also showed a high correlation, but lower than self-assessment. When self-assessing, students with lower grades tended to overestimate their performance. One week later test results showed that only self-assessment students performed significantly better. There was no difference between peer-assessment groups and control group. In conclusion, self and peer assessment can be used to save teacher time because they seem to be reliable. Furthermore, students' performances improve if they self-assess, but not if they peer-assess.	Outcome variables: test scores and accuracy of self- and peer grading.
Schamber & Mahoney, 2006 J. The Journal of General Education	No sample data. No control group.	General education course: "Today's decisions"	Quasi-experimental. Two groups, both treatment but intervention improved from year to year.	5 rubrics for each part of the paper	Students worked in groups to write a paper with five different parts. In the 2002 students did "prewriting" for the two first parts. In the 2003, students did "prewriting" for the five parts and used rubrics. Not clear if the rubric was used by the students or only by the assessor.	Prewriting and rubric were good strategies to promote students' critical-thinking skills.	Outcome variables: paper scores.
McCormick et al., 2007 J. Journal of Agricultural Education	44 college seniors. No age data.	Leadership development program	Quasi-experimental design (only one group). No control group.	Rubric	Students used a rubric to self-assess their performance in the course. The rubric score was compared to the final score on the exam to test the accuracy of their estimations.	Students showed a better understanding of the course material. Half of the students were accurate when estimating their final scores and half of them underestimated.	Outcomes variable: knowledge and accuracy of self-assessment scores. Problems in the design.
Andrade, Du & Wang, 2008 J. Educational Measurement: Issues and Practices	116 3 rd grade (N=46) & 4 th grade students (N=70). Four classrooms in the treatment, &	Writing	Quasi-experimental design with control group.	Writing rubric	Both conditions had to write an essay. Treatment group studied an exemplar, discussed strengths and weaknesses, and made a list of a good essay features. Then both groups wrote a first draft and received feedback from the teacher. Before the final submission both groups self-assessed their work, the treatment group with a rubric. In	No effect of time used for writing. No effect of prior rubric knowledge. Significant effect of using the rubric on the essay scores.	Outcomes variables: time for writing, essay grades or scores. Moderating variable: previous rubric experience.

	3 in the control. Unknown control group N.				sum, all students self-assessed, but treatment group used a rubric and analyzed a model essay.		
Andrade, Buff, Terry, et al., 2009 J. Middle School Journal	No specific information (Middle school students)	Writing	Descriptive. No control group.	Writing rubric	Intervention using two rubrics with students from middle school to improve their writing. Students were trained to self-assess and peer-assess their writing during several months.	Improvement in the state tests for ELA and teachers perceived that their students improved their writing skills.	Outcome variables: teachers' perceptions on students' learning and ELA results.
Andrade, Wang, Du & Akawi, 2009 J. Journal of Educational Research	268 students 3 rd grade (N=72), 4 th grade (N=105), 5 th grade (N=44), 6 th grade (N=60), 7 th grade (N=26). 3 schools, 18 classrooms. Control: 9 groups (N=137).	Writing	Quasi-experimental design with control group.	Writing rubric	Both conditions had to write an essay. Both groups studied an exemplar, discussed strengths and weaknesses, and made a list of a good essay features. Then, both groups wrote a first draft and reviewed their work (treatment group with a rubric). Afterwards the teacher gave them feedback and the students revised their essays before final submission. In sum, all students self-assessed, but the treatment group used a rubric and analyzed a model essay. Self-efficacy was measured three times: after discussing the exemplar, after handing out the rubric, after reviewing the first draft.	(A) There was no significant effect of the short-term use of the rubric (1 st & 2 nd measures) on student self-efficacy. Instead both groups improved. (B) There was no significant effect of the long-term use of the rubric by itself, but there was an effect in interaction with gender: girls who reported more previous exposure to rubrics tended to show higher levels of self-efficacy for writing. (C) There was a significant effect of gender: girls showed higher levels of self-efficacy, but only on the first measure. (D) There was a significant interaction between gender and rubric use: girls in the treatment group showed higher levels of self-efficacy than boys. Girls were more confident when they created their own feedback instead of being exposed to social judgment.	Outcome variables: self-efficacy.
Reitmeier & Vrchota, 2009 J. Journal of Food Science Education	35 senior students. Rubric N=17 Reflection N=18 No age data.	Oral presentation in a food science course	Quasi-experimental 2 treatment groups. No control group.	Oral presentation rubric	At the beginning of the course students were divided into two groups and were also given a self-assessment sheet: in one group a rubric, in the other a reflection sheet. Students self-assessed their oral presentations. Teacher and peers also assessed the presentations.	There were no significant differences between groups, hence no difference in the use of a rubric or the reflection sheet. No significant difference based on the assessor (self, peer, or teacher), even though peer-assessment scores were the highest. Students' own scores and teacher's scores were close to each other. Based on an analysis of students' comments, the authors concluded that the rubric was resourceful in order to communicate concepts, skills, etc., but once the students had that baseline knowledge, the reflection sheet encouraged them to go beyond and be more	Outcome variables: scores for oral presentation.

						aware of their work and make deeper reflections.	
Andrade, Du & Mycek, 2010 J. Assessment in Education: Principles, Policy & Practice	Total N=162 in 11 classes: 5 classes (N=89) treatment, 6 classes (N=73) comparison 75% girls. No age data.	9 classes English-language arts 2 classes history-social studies	Quasi-experimental 2 groups (treatment & comparison)	Writing rubric	The task was writing an essay. Students in the comparison group created a list of good essay criteria, wrote a first draft, received feedback from the teacher and submitted the final version. In addition, students from the treatment group analyzed a model essay, were given a rubric and used it to self-assess.	Treatment group had better final scores on the essay and in each of the seven criteria that were evaluated. Non-significant effect of gender.	Outcome variables: essay score, rubric criteria
Jonsson, 2010 J. Assessment in Education: Principles, Policy & Practice	3 cohorts of pre-service teachers (n= 170, 154, 138)	Educational psychology	Comparison between cohorts	Students in cohort 2 and 3 had access to rubric and exemplars prior to task, while cohort 1 only had access to rubric.	Students analyzed simulated classroom situation and self-assessed their analyses. Rubric and exemplars were used to provide transparency to the assessment.	There was a large difference in scores between cohort 1 (no exemplars) and cohort 2 and 3 (with exemplars).	Student scores
Reynolds-Keefer, 2010 J. Practical Assessment Research & Evaluation	45 undergraduate students in educational psychology	Writing	Descriptive study Qualitative data No control	Writing rubric	Students used a rubric for two writing assignments. The rubric was discussed two weeks before the due date. Graded papers were returned with the rubric attached. Students completed a 9 items questionnaire about rubric use.	Students considered the rubric useful and some of them used it as a checklist. Students claimed to understand their scores and felt that they were fairly graded. They also perceived that their work was better when using the rubric. Results on decreased anxiety are vaguely reported. 13 out of 42 claimed that their anxiety decreased, but there is no data reported on the other 29 participants. Students who reported a decrease in anxiety based their claim on improved communication with the teacher. 60 percent of the participants reported that they would use rubrics when they started working as teachers. 40 percent claimed that they would not because: (a) they thought they did not know enough about rubrics; (b) rubrics are not suitable for some age levels or subjects.	Students' perceptions on the use of rubrics.

Panadero, Alonso-Tapia & Huertas, 2012.	120 secondary students Age M=15.9. 52.5% females.	Landscape analysis	Experimental 2x3x2 (Total of 12 conditions) 1 st IV: type of instructions. 2 nd IV: rubric, script & control. 3 rd IV: type of feedback.	Rubric and script about landscape analysis	Participants analyzed 3 landscapes and received different types of instruction, self-assessment tools, and feedback. Self-regulation was measured through questionnaires and thinking aloud protocols.	Self-regulation: the use of scripts produced higher levels of self-regulation followed by the use of rubrics, which in turn outperformed the control group. Performance: the use of rubric and scripts improved performance as compared to the control group. Self-efficacy: the combination of the use of rubrics and mastery feedback produced the highest effect.	Self-regulation measures, student performance, and self-efficacy.
Panadero, Alonso-Tapia & Reche, 2013	69 pre-service teachers enrolled in a New technologies course. Age M=20.6. 98.6% females.	Design of multimedia material (power point, WebQuest) for educational purposes.	Quasi-experimental: control, rubric, and script groups.	Rubric and scripts to help design multimedia material	During a semester the participants designed multimedia material in one of three conditions. Along the semester they received feedback and training for a correct used of the tools. Before the intervention self-regulation, self-efficacy, and goal orientation was evaluated as a pre-control measure.	Self-regulation: the use of rubrics decreased the negative self-regulated actions oriented to performance/avoidance. No difference found with the control group on the positive type of self-regulation. Performance: no difference. Self-efficacy: no difference. Perception of the tool: the rubric was perceived as more useful for self-assessment as compared to scripts.	Self-regulation, performance, self-efficacy, and students' perceptions
Panadero, Alonso-Tapia & Huertas, 2014	85 first year psychology students enrolled in a Motivation and emotion course. Age M=18.9. 88.2% females.	Summarizing and conceptual maps design	Quasi-experimental: rubric and script groups. The control group was rejected due to invalid data.	Rubrics and scripts for summarizing and designing conceptual maps	During a semester the participants wrote six summaries and designed six conceptual maps with curricular content using one of two self-assessment tools. Along the semester they received feedback and training for a correct use of the tools. There was a pre-measure for self-regulation, self-efficacy, and goal orientation to test equality among groups.	Self-regulation: the use of rubrics decreased the negative self-regulated actions oriented to performance/avoidance. However the scripts enhanced more positive self-regulation. Performance: the rubric group outperformed the scripts group. Self-efficacy: no difference. Perception of the tool: the rubric was perceived as more useful for self-assessment as compared to scripts. Motivational effects: students using the rubric reported having more learning goals activated during the task.	Self-regulation, performance, self-efficacy, motivational effects, and students' perception

Appendix A**RUBRIC: HOW TO MAKE A SUMMARY**

Ernesto Panadero, Jesús Alonso Tapia & Juan Antonio Huertas

Score Criteria	4	3	2	1
The main idea from the original text is included	The summary starts with the main idea of the original text (not to be mistaken with the theme)	The summary contains the main idea, but it is not at the beginning	What is proposed in the summary as the main idea is mistaken (for example, theme is presented instead of the main idea)	The summary does not contain the main idea
The secondary ideas that support the main idea are included	The rest of the important ideas are expressed in the summary with an economy of words	The rest of the important ideas are expressed in the summary, but with an excess of words	Some, but not all, of the important ideas are expressed in the summary. An excess of words is used.	The summary contains a number of elements that are superficial and an excess of words is used.
Order of ideas	The progression follows a logical order which can be easily identified (connective words are used)	The progression follows a, more or less, logical order, which it is not easily identified (wrong use, or absence of, connective words)	The progression follows a, more or less, logical order, but there are unjustified leaps or omissions	The summary lacks an identifiable structure and it is difficult to follow the ideas expressed
Non-necessary information	There are no unnecessary details or personal reflections	There are no details or personal reflections, but the main idea is repeated redundantly	There are details and personal reflections, but the main idea is not repeated	There are details and personal reflections and the main idea is repeated redundantly
Precision of the terms used	The words chosen are efficient and concise	The words chosen are to some extent concise	The words chosen are ambiguous and repetitive	The words chosen are confusing