

YOUNG STUDENTS' USE OF EXPLICIT GRADING CRITERIA FOR SELF-REGULATION

Anders JÖNSSON

Publication version: June 2022
Open and collaborative peer review
Reviewers: Kari SMITH, Lesley WISEMAN

Abstract

The focus of this study is how young students (age 10-11 years) use explicit grading criteria as a means for self-regulating their learning and for improving their performance in school. Data comes from group interviews with students ($n = 64$), which were performed as part of an evaluation of a nationally mandated trial, where grades were introduced earlier in the Swedish school system. Findings from the analysis suggest that students do not use the grading criteria as such, but rely on colour markings, representing teachers' assessments. Findings also show that the students have several and specific strategies for self-regulation, but that these strategies are not connected to the grading criteria. Instead, the strategies are general and apply to all subjects. Some students also make reference to vague strategies, such as increasing effort. In this particular sample, references to specific versus vague strategies correlated with high-, and low-performing schools, but since there was only one low-performing school in the sample, this finding may not generalize to other schools. Changes in the pedagogical practice needed to accommodate a formative use of the grading criteria are discussed.

Keywords: explicit criteria, grading, self-regulated learning, summative assessment

Résumé

Cette étude se concentre sur la manière dont des jeunes élèves (âgés de 10-11 ans) utilisent des critères de notation explicites comme moyen d'autorégulation de leur apprentissage et d'amélioration de leurs performances scolaires. Les données proviennent d'entretiens de groupe avec des élèves ($n = 64$), qui ont été réalisés dans le cadre de l'évaluation d'un projet mandaté au niveau national, les notes ayant été introduites plus tôt dans le système scolaire suédois. Les résultats de l'analyse suggèrent que les élèves n'utilisent pas les critères de notation en tant que tels, mais se fient aux marques de couleur qui représentent les évaluations des enseignant-es. Les résultats montrent également que les élèves ont plusieurs stratégies spécifiques d'autorégulation, mais que ces stratégies ne sont pas liées aux critères de notation. Au contraire, les stratégies sont générales et s'appliquent à toutes les matières. Certain-es élèves font également référence à des stratégies vagues, telles que l'augmentation de l'effort. Dans cet échantillon particulier, les références à des stratégies spécifiques par rapport à des stratégies vagues sont corrélées avec les écoles à haut et à bas rendement, mais comme il n'y avait qu'une seule école à bas rendement dans l'échantillon, ce résultat peut ne pas être généralisé à d'autres écoles. Les changements dans la pratique pédagogique nécessaires pour permettre une utilisation formative des critères de notation sont discutés.

Mots-clés : critère explicite, notation, apprentissage autorégulé, évaluation sommative

1. Introduction

The focus of this study is how young students (age 10-11 years) use explicit grading criteria as a means for self-regulating their learning and for improving their performance in school. In Sweden, where the study is situated, national goals and performance standards were introduced during the 1990s as part of a major curriculum reform, thereby ending a prolonged period of norm-referenced grading. Several aspects of the reform were deemed unsatisfactory, however, and new reforms were suggested, for instance by further specifying the grading criteria (Statens Offentliga Utredningar, 2007). As a result of these reforms, Sweden has, since 2011, a grading scale with six levels (A-F) accompanied by extensive descriptions of so called "knowledge requirements" in all subjects. Students are graded from year 6 and onwards.

Although the guidelines from the Swedish National Agency for Education (2018) emphasize an integrative and holistic approach to grading, the digital educational platforms used by most schools, present portions of the grading criteria as separate items in a matrix format. As the educational platforms can be accessed by the students and their legal guardians through smartphones and other devices, it has become common practice to communicate assessment and progress information via these matrices. Most often, a specific item is marked with green if fulfilled, yellow if not quite fulfilled, and red if not fulfilled. By regularly checking their performance on assessment events, as communicated through the colour markings in the matrices, the students may keep track of their progress on a day-to-day basis.

Since the students are continuously provided with information about their performance in relation to the grading criteria, including extensive descriptions of performance standards, the matrices function as a feedback system, where students may take action based on the information provided. However, the details of this feedback process, and how it interacts with students' self-regulation strategies, are not known. The purpose of this study is therefore to explore how students navigate within this feedback system, by interpreting and using the information provided, to regulate their learning.

2. Background

According to the widely accepted definition by Sadler (1987), a criterion is "a distinguishing property or characteristic of anything, by which its quality can be judged or estimated, or by which a decision or classification may be made" (p. 194).

Following from this definition, using criteria for assessment purposes is a two-tier process. The first stage involves the discernment of the "distinguishing properties" in a text, a presentation, a product, or in any other format used, and the second involves making a judgement about the quality of the performance. Criterion-referenced assessments may result in a qualitative judgement about the potential of the particular piece of student work, which can be expressed in terms of strengths and suggestions for improvement according to the criteria.

The abovementioned characteristics of criterion-referenced assessments are responsible for the potential that such assessments have for students' learning. First, by focusing on strengths and suggestions for development, criterion-referenced assessments are excellent material for formative feedback. Second, without a common scale,



criterion-referenced assessments are not easily comparable between students, which means that social comparisons may be avoided. Third, since the assessment focuses on the task at hand (i.e., the assessment is direct, see Frederiksen & Collins, 1989), the assessment may (over time) support students in learning to judge the quality of their own or others' performance.

Yet another possibility provided by criterion-referenced assessments is to communicate the criteria to the students prior to their performance. As suggested by for instance Panadero et al. (2016), students could benefit from being familiar with the criteria during all phases of the self-regulation cycle (e.g., Zimmerman, 2013). They can use criteria to set more realistic goals for the activity during the planning phase, monitor their work during the performance phase, and also self-assess their performance during the evaluation phase. However, in order to communicate the criteria to the students beforehand, the criteria have to be made explicit.

2.1 Explicit criteria

There are different ways to make criteria explicit, but here the focus will be on "scoring rubrics". This choice is made since rubrics are probably the most common way to communicate criteria to students, and also the most thoroughly investigated tool for such communication (Dawson, 2017).

Rubrics are instruments for assisting assessors in judging the quality of student performance on open and/or complex tasks, as opposed to drawing conclusions about student proficiency based on the quantity of correct answers. All rubrics have at least two 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. And by combining these features into a two-dimensional matrix, a rubric has come into existence (Jönsson & Panadero, 2017).

Brookhart (2018) suggests that "true rubrics" are distinguished by "listing criteria for the work and performance level descriptions across a continuum of quality" (p. 1). In contrast, "checklists" ask for dichotomous decisions (e.g., yes/no) and "rating scales" include numerical (e.g., 1-5), evaluative (e.g., Excellent-Good-Fair-Poor), or frequency scales (e.g., Always-Usually-Sometimes-Never). While such scales may be useful for rating purposes, they do not offer students any descriptions of the quality of their performance, which they need for taking the next step in their learning.

In a systematic review of research on rubrics, Jönsson and Svingby (2007) suggested that the use of rubrics has the potential of promoting learning and/or improving instruction by making expectations and criteria explicit, which facilitate feedback and self-assessment. This proposal was further corroborated by a later review on rubrics, which focused exclusively on the formative function of rubrics (Panadero & Jönsson, 2013). This review showed that the transparency provided by the use of rubrics may: (a) reduce student anxiety, (b) aid the feedback process, and (c) support student self-regulation; all of which may indirectly facilitate improved student performance. Brookhart and Chen (2014) also noted, in a follow-up review on both summative and formative uses of rubrics, that several studies reporting on the effects of rubric use on learning and performance used relatively rigorous designs, such as experiments and quasi-experimental studies.



Since then, a number of empirical studies reporting on positive effects on student performance from the use of rubrics have been published. Although most studies in this line of research are situated in a higher-education context (e.g., Greenberg, 2015; Lipnevich et al., 2014; Ritchie, 2016), there are several examples of studies in which younger students' use of rubrics have been investigated. For example, Vasileiadou and Konstantinos (2021) investigated primary students' self-assessment with rubrics as a means to improve student performance. The results showed that the intervention had a positive effect on students' school performance, especially in Writing. Similar findings have been reported elsewhere, both for students in elementary classes (Andrade et al., 2008) and with middle school students (Andrade et al., 2010). Another example is provided by Smit et al. (2014), where four 5th grade primary classes used a rubric for mathematical reasoning. Although the project time was too short for some students to learn how to use the rubric for self-regulation purposes, findings suggest, for instance, that the students were indeed able to use the rubric to guide their work and that it had a positive effect on motivation and self-efficacy (see also Smit et al., 2017).

2.2 The use of explicit criteria for self-regulation

As mentioned above, it has been suggested that students could benefit from being familiar with the criteria during all phases of the self-regulation cycle (Panadero et al., 2016). The most obvious support for this idea comes from studies that report on students using rubrics to plan, monitor, and evaluate their task performance. As an example, Jönsson (2014) performed case studies in different settings in professional education (public health, real estate brokers, and dental education), where rubrics were used. In these settings, students claimed to use the rubric as guidance when planning their assignments. Several students also reported that they used the rubrics to assess the progress of their work during task performance, using the criteria "as targets to make sure we included all the important aspects in the assignment" (p. 10), as well as to make a final check before submitting the assignment for summative assessment. Similar findings are reported in other studies (Andrade & Du, 2005; Reynolds-Keefer, 2010).

Studies such as these suggest that when students (at least in higher-education contexts) are provided with rubrics, they use them to support all phases of the self-regulation cycle (i.e., the forethought, performance, and self-reflection phases). During the forethought phase rubrics may support students by helping them to plan and structure how to perform a task, as well as to set more realistic goals for themselves. Rubrics may support students during the performance phase by facilitating the monitoring of own performance. During the reflective phase rubrics may support students by helping them to judge the quality of their work in relation to pre-set goals instead of relying on social comparisons.

In another line of research, Panadero and his colleagues investigated dimensions of self-regulated learning (SRL) associated with negative emotions and stress, actions directed by anxiety, external pressure to perform, and task avoidance. In these studies, students' scores on performance- and avoidance-oriented SRL scales were shown to decrease for students using rubrics (Panadero et al., 2012; Panadero & Romero, 2014). These are positive findings, but the students' using rubrics in the study by Panadero and Romero (2014) also reported higher levels of stress while performing the task as compared to the control group. Furthermore, learning-oriented SRL scores may decrease for students using rubrics (Panadero et al., 2014). This means that while the use of rubrics may decrease



performance- and avoidance-oriented SRL strategies, typically unfavourable for learning, they do not necessarily increase the use of learning-oriented SRL strategies.

Student self-efficacy is yet another aspect of supporting students' SRL. By using rubrics to improve performance and facilitate self-assessment, rubrics should (ideally) increase student self-efficacy and favour a mastery orientation among the students. This ideal picture, however, does not really appear in empirical research. For example, in a study by Panadero et al. (2012) no significant effects were found for students' self-efficacy. Furthermore, Balan and Jönsson (2018) investigated the effects of explicit assessment criteria on primary students' performance, self-efficacy, and self-regulation. Successive levels of explicitness, from feedback based on (implicit) criteria to a combination of exemplars and explicit criteria, were implemented in eight classes during four teaching sequences. Findings show that there was an increase in both performance and self-efficacy, but not in relation to explicitness. These changes were instead assumed to be an effect of the formative feedback provided as part of the intervention, which rapidly and substantially increased low-performing students' performance and self-efficacy. The only change related to the level of explicitness, was an increase in self-regulation scores by high-performing students when having access to both exemplars and explicit criteria. This study therefore provides a clear indication of the need to differentiate between low-, and high-performing students in relation to SRL.

A recent meta-analysis on the effects of self-assessment interventions on self-regulation strategies, also shows that students using rubrics (or similar instruments), reported lower self-efficacy after the intervention than participants not using them (Panadero et al., 2017). This could be an effect of students becoming aware of the complexity of high-quality performance and therefore reporting lower self-efficacy. These findings correspond to Andrade and Du (2005), where some students found it overwhelming to read all the levels of the rubric before receiving feedback from the teacher. However, there were also students aiming for higher performance, who did not read the levels describing lower quality in the rubric, suggesting that students may react differently to rubrics, depending on their previous performance and academic self-esteem.

2.3 Critique against the use of explicit criteria

Making criteria explicit undoubtedly has its advantages. However, there are also perils of transforming implicit criteria to explicit. Some problems with explicit criteria have been meritoriously discussed by Sadler (e.g., 2009, 2014). For instance, Sadler points out that it does not matter how many criteria you define, they will still not be able to represent the richness and complexity of real-world performance. This means that teachers always run the risk of encountering student performance that is judged as high quality, but that does not fit into the predefined set of criteria. That explicit criteria cannot fully represent the richness and complexity of real-world performance also means that assessments of different parts or aspects of performance does not necessarily add up to the whole. This is particularly evident in cases where sub-scores from analytical assessments are arithmetically added together into a summary score, possibly resulting in a score not in line with a holistic assessment of overall quality. It should be noted, however, that scoring criterion-referenced assessments (as defined here) is questionable, since it means placing qualitatively different dimensions of performance on the same scale and also making the assessment compensatory (Sadler, 1987). It would be more reasonable to express the outcome of criterion-referenced assessments in terms of strengths and suggestions for improvement (i.e., a qualitative assessment). In such cases, there does not have to be any



conflict between analytic and holistic assessments; rather they may complement each other (Tomas et al., 2019).

Another potential peril of transforming implicit criteria to explicit ditto, is that the use of explicit criteria may turn students' attention away from productive learning and instead focus on surface strategies and "criteria compliance" (Torrance, 2007). This means that when being informed about the criteria, learners will focus on meeting these criteria with minimal effort and also limit their performance to what is explicated by the criteria, leaving other things aside. As noted by Panadero and Jönsson (2020), in a critical review of the arguments against the use of rubrics, the empirical support for "criteria compliance" is relatively weak, since most critics base their claims on either anecdotal evidence and personal experiences, or theoretical assumptions. More research is therefore needed in order to understand how students use explicit criteria to support their learning.

2.4 Assessment criteria versus grading criteria

The focus of this study is how young students use explicit grading criteria as a means for self-regulating their learning and to improve their performance in school. Consequently, there are at least two important aspects that differ from previous research. First, several of the studies reporting on the use of rubrics for formative purposes are situated in a highereducation context. As remarked by Panadero and Jönsson (2013), while studies performed in higher-education contexts tend to report on positive results when providing the students with rubrics, longer and larger interventions are typically needed in order to produce positive results in schools. Time devoted to work with the rubric therefore seems more crucial for younger students and studies only investing a few lessons typically report no, small, or mixed results (Panadero & Jönsson, 2013; Smit et al., 2017).

Second, there is a difference between "assessment criteria" and "grading criteria". While the term "criterion-referenced assessment" refers to judging or estimating the quality of student performance on individual tasks according to criteria, "grading", as used here, means making a decision about students' overall performance based on accumulated data throughout a semester, course, or other period of time. Grading criteria are therefore likely to be more overarching and abstract, as compared to assessment criteria, and grading typically also involves summarizing student performance into a standardized format, expressed along a scale. As a consequence, grading criteria lack several of the characteristics that make criterion-referenced assessment suitable for supporting student learning and self-regulation. Still, as outlined in the introduction, the students in this study are continuously provided with information about their performance in relation to the grading criteria, which include extensive descriptions of performance standards (for an example, see Figure 1), and the students are expected to take action based on this information.



Figure 1

Excerpt from the Swedish national curriculum (Lgr11, English version, p. 61) showing part of the knowledge requirements for grade E at the end of year 6 in mathematics

Pupils can solve simple problems in familiar situations in a **basically** functional way by choosing and applying strategies and methods with **some** adaptation to the type of problem. Pupils describe their approach in a **basically** functional way and apply **simple and to some extent** informed reasoning about the plausibility of results in relation to the problem situation, and can also **contribute** to making **some proposals** on alternative approaches.

Note. The words in bold are the ones that differ for different grades. For grade C, for example, students are expected to “solve simple problems in familiar situations in a **relatively well** functional way” instead. It should also be noted that the excerpt shows approximately one third of the knowledge requirements for this grade. As there are knowledge requirements for several grades in each subject, the Swedish syllabi are very extensive documents.

2.5 Grading criteria and subject matrices

As outlined in the introduction, the students in this study are continuously provided with information about their performance in relation to grading criteria. Since grading criteria are likely to differ between countries, both in design and focus, examples of the Swedish criteria are provided here, in order to aid the understanding of students' statements.

The Swedish national curriculum for compulsory school (Lgr11) includes separate syllabi for all individual subjects. These syllabi, in turn, consist of a description of the aim for the specific subject, followed by core content and performance standards (called “knowledge requirements”). Core content is provided for three intervals (i.e., years 1-3, 4-6, and 7-9) and the knowledge requirements declare what the students are expected to master at the end of each interval.

The knowledge requirements include performance standards for three grade levels (E, C, and A) at the end of years 6 and 9. Since students are not graded in year 3, there is only one level at the end of this interval. The grades D and B have no explicit standards. Instead, these grades are used when all the knowledge requirements for the lower grade level are satisfied, and most of the higher.

The knowledge requirements for the three grade levels are mostly identical, and differ only with respect to some specific words that signal which grade level student performance belongs to. For example, Figure 1 shows an excerpt from the knowledge requirements for grade E at the end of year 6 in mathematics, where the words in bold are the ones that differ for different grades. For grade C, for example, students are expected to “solve simple problems in familiar situations in a **relatively well** functional way” instead of in a “**basically** functional way”. It should also be noted that the excerpt shows only about one third of the knowledge requirements for this grade. The whole text for grade E at the end of year 6 is approximately 230 words in length.

As is obvious from the description of the Swedish grading criteria, the syllabi are quite extensive documents, making the communication regarding student performance in relation to these criteria a challenge for teachers. It has therefore become common practice to use a matrix format to present the criteria, so that the differences between grade levels become more apparent. In these matrices, the knowledge requirements are split into separate segments and fitted into different cells. As there are three grade levels (E, C, and



A), these are placed in separate columns, where similar segments (differing only in words in bold, as exemplified in Figure 1) from each grade level constitute the rows in the matrix (Figure 2).

Figure 2

Example of how the knowledge requirements from the Swedish national curriculum (Lgr11, English version, pp. 61-63) are presented in a matrix format

Grade E	Grade C	Grade A
Pupils have basic knowledge of mathematical concepts and show this by using them in familiar contexts in a basically functional way. ...	Pupils have good knowledge of mathematical concepts and show this by using them in familiar contexts in a relatively well functioning way. ...	Pupils have very good knowledge of mathematical concepts and show this by using them in new contexts in a well functioning way. ...
Next segment	Next segment	Next segment

Note. Only a small part of the text is shown in the cells.

Even if the matrix format facilitates the comparison between requirements for different grade levels, the matrices are still very extensive and therefore cumbersome as tools for communication. Another common practice is therefore to use colour markings, so that a segment is marked with green if satisfactorily fulfilled by the student, yellow if not quite satisfactorily fulfilled, and red if not fulfilled at all. In the most extreme case, students (or legal guardians) may be presented with matrices with colour markings, but without any actual text in the cells. In these cases, students may still be able to extract information from the matrices by being familiar with the text that has been left out, or by comparing with a complete version of the knowledge requirements. However, the colour markings in an empty matrix also provide some information, even without any knowledge about the text that has been left out. For example, the number of red, yellow, or green cells give an overview of overall progress in a subject, which—albeit coarse—is still somewhat more detailed than a single letter. A red or yellow cell in an otherwise green matrix also signals that something is not handled satisfactorily. And although it may not be possible to see exactly what, the signal still provides an incentive to look more closely elsewhere, for instance by going through written feedback or asking the teacher.

3. Aim and research questions

The aim of this study is to explore how young students (10-11 years old) regulate their learning within a feedback system that involves the continuous provision of information in relation to grading criteria. The specific research questions are:

- How do students use the grading criteria?
- What strategies do the students have for self-regulation?



4. Method

4.1 Sample and data

This study is based on data from an evaluation of an officially mandated trial (Löfgren et al., 2021), where grades were introduced in grade 4 (students aged 10-11 years) at a limited number of Swedish schools (n = 11). The evaluation was performed in order to provide the Swedish government with information about the consequences of introducing grades earlier in the educational system¹.

Of the eleven schools taking part in the trial, five schools were included in the qualitative part of the evaluation, which was based mainly on semi-structured interviews with teachers, principals, special-education teachers, and students (n = 135), collected at two occasions one year apart (Löfgren et al., 2021). The selection of schools was made to achieve maximal variation within the small sample, taking ownership, geographical location, students, socio-economic background, and performance measures into consideration (see Table 1). Data for this particular study comes only from the student interviews (n = 64).

The students were interviewed in groups of 3 to 5 individuals (Table 2). Due to students' age, they were interviewed in groups, which was thought to help them feel safe in the interview situation and less discouraged from sharing their thoughts and experiences. The students' teachers were explicitly asked to select appropriate students and create groups, where the students would feel safe with each other, as well as to include students at different levels of performance.

The schools who took part in the trial are not representative for Swedish schools. As can be seen in Table 1, most schools in the sample were independent, although on a national level they constitute less than 20 percent of all schools (while the rest are public schools). Furthermore, most schools in the sample have a higher percentage of parents with higher education, as well as a lower percentage of students with migratory background, as compared to the country as a whole. These differences are reflected in students' grades (in Table 1, mathematics is used as an example), which are typically higher than the national average.

Table 1

Official statistics (2017-2018) for schools participating in the evaluation

School	Public or independent school	Geographical location	Parents with higher education (%)	Students with migratory background (%)	Grade average in mathematics (0-20)
Sweden	N/A	N/A	58	25	12.7
A	Independent	City	89	5	16.3
B	Independent	City	44	90	14.8
C	Public	Urban	79	6	13.5
D	Independent	Urban	75	27	15.0
E	Independent	City	44	92	9.0

¹ Before 2011, grades were introduced in grade 8 (students aged 14-15 years) in the Swedish educational system. However, some political parties have argued for introducing grades earlier and as of 2011, along with the reforms described in the introduction, grades were introduced in grade 6. To evaluate the consequences of introducing grades even earlier, an officially mandated trial was performed with volunteering schools. However, due to political bargaining, the government allowed schools to introduce grades in grade 4 before the trial was finished.



Table 2*Overview of student interviews*

School	Groups interviewed	Number of students
A	2	7
B	4	15
C	3	14
D	6	20
E	2	8

A couple of schools in the sample (i.e., schools B and E) depart from the general pattern by having a lower percentage of parents with higher education, as well as a higher percentage of students with migratory background, as compared to the whole country and the other schools. Among these, students at school B still have relatively high grade averages, while students at school E have relatively low grade averages. Overall, students at most schools in the sample perform substantially higher than the national average, while students at one school perform substantially lower.

4.2 Procedure

Due to large geographical distances between the schools, four different researchers performed the interviews, using a shared and agreed upon interview guide. The guide contained questions about students' experiences of being graded, as well as a number of vignettes, which were used if students found it hard to talk about their own experiences. For example, one vignette asked the students to imagine that a new student, coming from another school, joined their class. The new student had never received a grade before, and the interviewees were asked to explain to the new student what grades are and how grading works. Another vignette asked the interviewees to give advice to the newcomer about how to improve her/his grades.

Both students and their legal guardians had to give consent to participation. Written informed consent was obtained from the latter. The consent form was available in both Swedish and Arabic.

The interviews were recorded with portable mp3-players and then transcribed. These transcriptions have been used as data in this study.

4.3 Analysis

The interviews were analysed (or "re-analysed" in relation to the original evaluation) by the author with conventional thematic analysis, which is a method for identifying, analysing, and interpreting patterns of meaning within qualitative data (Clarke & Braun, 2017). The analysis was mainly inductive in nature and followed the procedure outlined by Braun and Clarke (2006). First, the transcripts were read and re-read, marking interesting passages with different colours. For example, students' statements about what they do in order to improve their performance or grades (i.e., their use of self-regulation strategies) were marked with one colour, while statements about matrices, progression, or teacher feedback were marked with other colours. The statements were then collected in a spreadsheet and further analyzed in relation to the research questions. This was done in relation to each research question separately, for instance by looking at what kinds of strategies the students claimed to use, as well as by looking for potential interactions between the use of strategies and grading criteria. From this analysis, typical ways to talk about the use of grading criteria and



self-regulation strategies were identified, and compelling extracts from the data were selected to exemplify these themes. The selected extracts were then checked against the research questions and literature, and descriptions of the themes were made. Finally, the extracts were translated into English by the author.

5. Findings

The findings are presented according to the research questions, first how students use the grading criteria and then which strategies they have for self-regulation. Letters in parentheses after the quotes refer to the schools in Table 1, while numbers refer to the specific group of students at that school (Table 2).

5.1 Students' use of grading criteria

This section describes how students claim to use the grading criteria. Although all students know about the grading criteria, there are differences in how available they are, for instance whether students may access them on paper or digitally, and, in the latter case, if the criteria are readily available through students' smartphones or not. This difference is most pronounced between the schools, while students tend to express similar attitudes and experiences within each school. For example, students at some schools (schools A, B, and D, the most high-performing schools in the sample) have access to the grading criteria through their mobile phones or other digital devices, and these students claim to check their progress regularly, sometimes on a daily basis. At other schools, the teachers do not update the criteria regularly or do not share their assessments with the students through the digital platforms. At school E, for instance, progress reports are made on paper rather than digitally. As the progress reports on paper do not include the actual criteria (only grades or colour markings), the criteria are not as readily available to the students in these cases. As a consequence, the criteria are not present to the same extent in all interviews.

One of the most notable features of how the students talk about the grading criteria in the interviews, is that there are very few references to the actual content of the grading criteria. Instead, students mainly refer to the colour markings, which appear in the "subject matrices":

.../ matrix is after a test, so that after a test we can see if it is green or yellow or red. (B:1)

These colour markings indicate to the students what grade their performance represent and if they have made progress:

And then you press subject matrix like this, "show the grades", and they have filled in here, so there are three cells, and two cells are green and then it's a C. If one cell is green and one is yellow, then it's "on the right track" or something like that. We use them to see if we have improved. (D:2)

If the students have not made appropriate progress, the colour markings signal that more effort is needed:

Here's something called subject matrix and then you can see what you ... if it's green, then you are doing good, really good. But if it's yellow, you have to practice some more, but if it's red, you have to kind of ... then you should really practice some more. (D:2)



Even if this information may be considered coarse, the students still appreciate the fact that they receive more detailed information about their performance than before:

Yes, but then you got like ... at the end of the semester you only got a piece of paper and /.../ then each subject was checked and above, in a table, it said "not passed", "passed", "more than acceptable knowledge". (C:1)

As can be seen from these quotes, however, students do not refer to the actual content of the grading criteria, only to the cells and the colour markings. But unlike explicit criteria, the colour markings do not necessarily provide any specific or task-related information. The students therefore have to piece together information from different sources in order to "puzzle out" what the markings may mean:

When I feel that things are not going very well in class, I check the grades I have received and what it is that has not gone quite as well, and then I try to figure out what it is. And then I try to ... improve this particular thing. (A:1)

Figuring out what to improve from this, sometimes scant, information is difficult for the students. Several students therefore wish for more detailed feedback:

/.../ it's good, although I would like ... that they would write "you have become good at this" so "you need to practice more on this", so explain more. (B:3)

And the ... yes, but I didn't really know in math which ... I mean in which area I needed to improve, because they only write in the subject matrices when we've done a test. /.../ So I didn't really know what to improve, but I knew I had to improve in math. (A:1)

Formative feedback is sometimes provided by the teachers, alongside the colour markings, which is appreciated by the students:

Yes, sometimes they write something like "well done, you've been good at explaining this particular thing and this and you've discussed this" and things like that, and then ... and then it's, well, because you have ... I mean sometimes they write things like that and ... and sometimes not, but ... sometimes it only says "well done" and then you don't know... (D:1)

Taken together, students' statements suggest that they do not use the grading criteria as such but rely mainly on the colour markings, representing teachers' assessments. Since the markings do not include specific information about strengths and areas in need of improvement in student performance, it is difficult for the students to figure out exactly what to do with the information. In some cases, students realize that they have to improve, but without knowing what or how. In other cases, the teachers provide additional information, making it easier for the students to use the information formatively.

5.2 Students' self-regulation strategies

This section describes the strategies students claim to use when regulating their learning. Students at all schools have something to say about strategies for self-regulation, although the descriptions of their strategies may be more or less elaborate. For example, one of the students at school B describes her strategies like this:



I have different [strategies], I have a lot of things. So, the first thing is to ask the teacher about an extra copy of the homework or about the tests, for example if the parent can help at home and do an extra test for you. Or if you ... when you have a test or an assessment, you can take your cousin, or even your dolls or something, and sit and explain to them what you have done ... I mean, in relation to the goals, what you have done and so on. Then you... then it comes into your brain and gets stuck there, because then you have, like the teacher, you have explained to others, so you must understand it. And you have to make an effort, you have to listen during class, and you have to ask ... do your homework and ask your parents for help if you need it ... You always have to ask for help if you need it, otherwise there is no point. (B:1)

Other students are not as elaborate in their descriptions, but they can give quite specific examples of strategies that they find helpful:

Really participate during class and raise your hand if you don't understand, because then you need to have it explained to you one more time /.../ but if you maybe get a D, for example, then you say this to your parents, just like "yes, but I got a D" and then maybe they will give you a little extra help, and if you don't understand your homework then maybe they'll explain it to you one more time, and things like that. (A:1)

You listen during class and try to keep up and don't draw or do something else when someone is talking. (C:1)

Borrow books. /.../ To be able to read and write, that's the most [important]. (E:1)

In the interviews, several "classic SRL strategies" (e.g., Nussbaumer et al., 2015) are frequently mentioned by the students, such as strategies for managing resources (e.g., help-seeking, time management, working with peers), metacognitive strategies (e.g., paying attention, concentrating), and cognitive resources (e.g., rehearsing, elaborating, organizing). Strategies for managing attitudes and feelings were also mentioned:

You should make an effort, you should listen during class, you should be like ... you should not be like this "ugh, this is so boring". You should be kind of happy, you should feel like this is fun. Otherwise it's ... otherwise there is no point in sitting and listening. (B:1)

Not just look through it briefly in school, but take the books home and work with your parents, ask if there's something you don't ... I mean, you should not be ashamed, for instance, I'll ask my parents if I can't ... (B:4)

One student even suggests that you should try not to be ill:

You have to be ... try not to be ill, not be away from school. (C:2)

In addition to these specific strategies, a number of students mention persistence and increased effort as strategies. These descriptions are typically very vague:

/.../ you sort of do your own thing on tests and stuff ... yes, you ... yes, you just practice and continue to practice ... yes, that's all. (D:1)

Work harder. Work better. (E:1)

Then maybe you should do some more work at home... (C:1)



In this particular sample, the students at high-performing schools were able to describe both a greater number of different strategies, and more specific ones, as compared to students at the low-performing school, who mainly referred to vague strategies about increased effort.

In summary, students' statements suggest that they have several and specific strategies for self-regulation, but that these strategies are not connected to the grading criteria. In fact, no students, in any school, referred to the grading criteria when describing their self-regulation strategies. Instead, the strategies are general and applies to all subjects. Some students also make reference to vague strategies, such as "working harder". In this particular sample, references to specific versus vague strategies correlated with high-, and low-performing schools.

6 Discussion

The aim of this study has been to explore how students regulate their learning and optimize their performance within a feedback system, which involve the continuous provision of information in relation to grading criteria. This has been done by interviewing students about their experiences of being graded and analysing their answers in relation to how they use the grading criteria and what strategies they have for self-regulation. The interviews have been transcribed and analysed with conventional thematic analysis. Below, the findings from the analysis are discussed in relation to previous research on the use of explicit criteria for formative purposes.

6.1 Using explicit criteria for formative purposes

Criterion-referenced assessments have characteristics that make them suitable for supporting student learning, such as focusing on strengths and suggestions for development in relation to shared criteria. This means that students can use the information from such assessments in order to improve their performance on particular tasks (i.e., short-term use), but also - over time - learn to identify the important qualities that the criteria represent and self-assess their performance (i.e., long-term use).

Although grading criteria are typically more abstract and less task-specific than assessment criteria, they could - at least in principle - serve the same purpose by focusing on progress (including strengths and needs for improvement) over time. The positive effects of assessment criteria come primarily from the transparency they provide (Panadero & Jönsson, 2013). If students understand the expectations better, they are in a better position to deliver what the teacher asks for, but also to interpret and use feedback, and to engage in self-, and peer assessment (e.g., Smit & Birri, 2014). As can be seen from the interviews, however, this is not the situation with the Swedish grading criteria.

Instead of engaging with the criteria, and trying to comprehend the qualities they represent, students engage only with the colour markings. This means, for one thing, that even if the markings provide some crude information about what the students need to improve (e.g., math), they do not provide more specific information about strengths or areas in need of improvement, or guidance in how to improve. Students therefore need to figure out what the colour markings may mean and what they are supposed to improve,



which is quite the opposite of the suggested transparency achieved by explicit assessment criteria.

Another important point is that the colour markings do not provide sufficient information about the qualities the criteria represent, which means that they are unable to empower the students to interpret and use these qualities for formative-assessment activities, such as self-, and peer assessment. Instead of supporting students to become self-sustained in their learning process, the students become dependent upon their teachers to provide a continuous input of external feedback. As indicated by students' statements, this feedback seems to come primarily after tests or other summative assessment events, and not as an integral part of teaching.

6.2 Using explicit criteria for self-regulation

Previous research has suggested that students could benefit from being familiar with the criteria during all phases of the self-regulation cycle (e.g., Panadero et al., 2016), for instance by using criteria to plan, monitor, and evaluate their task performance.

In the interviews, students describe many different, detailed, and specific strategies for self-regulation, such as paying attention during class, asking for help, managing their time, elaborating on their answers, and studying together with peers. Some students also refer to less specific strategies, focusing on perseverance, diligence, and effort. It is therefore no doubt that at least some of the students have a battery of self-regulation strategies to use.

Regardless of how specific the strategies are, however, there is a complete absence of references to the grading criteria when describing self-regulation strategies. Although it is obvious from the descriptions about how the students use the colour markings for evaluating their progress, they do not mention the grading criteria when asked about self-regulation strategies. The idea to use the criteria for planning, monitoring, and self-assessing does therefore not seem to exist among the students. Instead, they rely on external feedback from the teacher in order to evaluate their progress.

Of course, even if students do not actively use the grading criteria for self-regulation, the external feedback in the form of colour markings may still have an impact on students' goal-orientations and self-efficacy (Smit & Birri, 2014). However, since the students do not seem to use the content of the criteria, it is difficult to see how the markings could facilitate a more learning-oriented approach. On the contrary, most of the students appear to have a very strong performance-orientation, where they strive towards more green cells and higher grades. In a sense, students may be considered to display the symptoms of "criteria compliance" (Torrance, 2007), but without using the grading criteria.

6.3 Criteria not appropriate for formative purposes

As suggested by the findings, students do not seem to use the grading criteria for self-regulation or other formative purposes. Instead, they use only the colour markings and rely on self-regulation strategies that are not explicitly connected to the criteria. There may be several reasons for this, but one plausible explanation is the design of the criteria.

According to the distinction between "true rubrics" and rating scales made by Brookhart (2018), the Swedish knowledge requirements might be classified as hybrids between these



two instruments. Similar to rubrics, the requirements include performance level descriptions, but the different levels are distinguished from each other only by the logic of rating scales, such as basic-good-very good (see Figures 1 and 2 above). Such relative terms do not, however, provide any tangible information to the students about the qualities sought for. For example, what does “basic knowledge of mathematical concepts” look like, and how can this be distinguished from “good knowledge”?

Although the Swedish knowledge requirements may be adequate instruments for teachers' grading, they are not necessarily appropriate for formative purposes, at least not in the hands of students. In order to be used formatively by the students, the requirements may need to be transformed from rating scales to “true rubrics”, including descriptions of performance levels across a continuum of quality. Under such circumstances, students are in a better position to use the criteria to identify specific strengths and areas in need of improvement (e.g., Panadero & Jonsson, 2013) and not, as currently seems to be the case, realizing that something is wrong without knowing exactly what.

7 Conclusions

Students' statements suggest that they do not use the grading criteria as such, but rely on the colour markings, representing teachers' assessments. Their statements also suggest that they have several and specific strategies for self-regulation, but that these strategies are not connected to the grading criteria. Instead, the strategies are general and applies to all subjects. Some students also make reference to vague strategies, such as increasing effort. In this particular sample, references to specific versus vague strategies correlated with high-, and low-performing schools, but since there was only one low-performing school in the sample, this finding may not generalize to other schools.

7.1 Pedagogical implications

Findings from this study suggest that students did not make use of the grading criteria to support their learning and self-regulation, only the colour markings. This is not to say that grading criteria cannot serve the same purpose as assessment criteria, by focusing on progress over time. However, to accommodate such a formative purpose, some changes in the pedagogical practice are needed. In particular, findings from this study point to at least two such changes.

First, as suggested by research on the formative use of rubrics (e.g., Panadero & Jönsson, 2013), the positive effects of assessment criteria come primarily from the transparency they provide. Criteria may support learners in, first, discerning the “distinguishing properties” in a performance, and second, making a judgement about the quality of the performance. This, in turn, helps the students to understand the qualities sought for, which puts them in a better position to deliver what the teacher asks for, but also to interpret and use feedback, and to engage in self-, or peer assessment (Smit & Birri, 2014). However, if the criteria do not help the students to discern these “distinguishing properties”, for instance because the criteria focus on evaluative rather than descriptive levels of quality (Brookhart, 2018), the criteria lose most of their formative potential. In order to be used formatively by the students, the requirements may therefore need to be transformed from rating scales to “true rubrics”, so that students can use the grading criteria to identify strengths and areas in need of improvement.



Furthermore, and similar to the situation with assessment criteria, students have to engage actively with the grading criteria in order to - over time - develop an understanding of their meaning and usefulness. However, judging the quality of their performance on individual tasks is not enough, they also need to evaluate and judge their progress over time, with the aid of the grading criteria. Since the students are relatively young, and grading criteria are typically more abstract as compared to more task-specific criteria, it is likely that considerable time needs to be invested in order for students to be able to understand and use the criteria (cf. Balan, 2012; Panadero & Jönsson, 2013). One established way of making such investments is by working with portfolios, which may provide a record of student performance that can be used to evaluate progress over time (cf. "the learning portfolio" in Klenowski et al., 2006).

Second, from the statements made by students in this study, a strong performance-orientation can be inferred. The students focus on the number of green cells in the matrices, which are translated into grades. What the criteria in the cells represent in terms of knowledge and learning appears to be secondary, similar to the surface-learning strategies described by Torrance (2007) as induced by access to explicit criteria. This, however, does not mean that students' strategies are necessarily inadequate. On the contrary, most students have obviously developed very successful self-regulation strategies, optimized for the learning required at these schools, which (according to students' statements) to a large extent involve remembering and recalling factual knowledge on tests. This is interesting, because if most assessment situations require the same kind of skills and/or cognitive processes (i.e., remembering and recalling), the students may be able to use the same self-regulation strategies in different subjects. Consequently, there is no need for students to engage with the content of the grading criteria, which may help explain why students do not seem to use the grading criteria, even if they have successful (but not subject specific) strategies for self-regulation.

Similarly, as teachers at most schools continuously and frequently provide the students with feedback, the students do not really need to monitor or evaluate their performance by themselves. In principle, students could use the frequently provided feedback to calibrate their performance through trial-and-error.

Taken together, this means that if students are to use the grading criteria to support their learning and self-regulation, a demand for deep-learning strategies and for self-regulation needs to be created. As suggested above, learning portfolios (Klenowski et al., 2006) may very well provide the structural support needed for self-regulation processes. For in-depth engagement with the grading criteria, however, there is also a need for alignment between the subject-specific criteria and expectations of student performance, so that students cannot use the same general strategies for all subjects and assessment situations.

7.2 Limitations and further research

This study has several important limitations, which should be kept in mind when interpreting the findings. Most importantly, the study includes a limited sample of students who volunteered to participate, and who attend schools that are not representative for most Swedish schools. The experiences and perceptions of these students may therefore not be representative for a larger sample of schools or students. It should be noted, however, that the grading criteria apply to all schools and students in the country, and most schools also use some kind of digital platform for communicating information regarding assessment and



grading. The findings may therefore, at least in part, be relevant and have implications for most schools.

Another important limitation is that the interviewees are quite young (age 10-11 years), which may have limited their possibilities to reflect upon and articulate how they use the criteria and how they self-regulate. The rich descriptions made by the students to some extent contradict such an apprehension, but there were large differences between schools, where students at school E – with a substantial share of students with migratory backgrounds – were significantly more taciturn, as compared to the students at the high-performing schools. The perspectives of the student from low-performing schools may therefore be less well represented in this study. The lack of subject-specific strategies, or use of subject-specific criteria, in students' statements, may also be partly due to the interviews not being performed in relation to any specific subject.

That the interviews were performed with groups of students is both a strength and a limitation. Even if the group was thought to make the students feel safe, it may also have affected their willingness to share more personal thoughts. This was one of the reasons to use vignettes in the interviews, as they opened the possibility for the students to talk about a fictitious person rather than about themselves.

Since this is a relatively small-scale interview study, future research could include a larger sample and/or make use of other methods for collecting and analysing data, such as questionnaires, in order to validate the findings reported here. Based on the findings from this study, it could be considered particularly important to investigate how the grading criteria are used and to what extent these practices contribute to an increased performance-orientation, and/or an instrumentalist approach to learning, among the students.

References

- Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research and Evaluation*, 10, 1-11. <https://doi.org/10.7275/g367-ye94>
- 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. <https://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 Practice*, 27(2), 3-13. <https://doi.org/10.1111/j.1745-3992.2008.00118.x>
- Balan, A. (2012). *Assessment for Learning: A case study in mathematics education* [Doctoral dissertation, Malmö University, Sweden]. Malmö University Publications. <https://mau.diva-portal.org/smash/record.jsf?pid=diva2:1404417>
- Balan, A., & Jönsson, A. (2018). Increased explicitness of assessment criteria: Effects on student motivation and performance. *Frontiers in Education: Assessment, Testing and Applied Measurement*, 3(81). <https://doi.org/10.3389/feduc.2018.00081>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education: Assessment, Testing and Applied Measurement*, 3(22). <https://doi.org/10.3389/feduc.2018.00022>
- Brookhart, S., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, 67(3), 343-368. <http://dx.doi.org/10.1080/00131911.2014.929565>



- Clarke, V., & Braun, V. (2017). Thematic Analysis. *The Journal of Positive Psychology*, 12(3), 297-298. <https://doi:10.1080/17439760.2016.1262613>
- Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, 42(3), 347-360. <https://doi.org/10.1080/02602938.2015.1111294>
- Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18, 27-32. <https://doi.org/10.2307/1176716>
- Greenberg, K. P. (2015). Rubric use in formative assessment: A detailed behavioral rubric helps students improve their scientific writing skills. *Teaching of Psychology*, 42(3), 211-217. <https://doi.org/10.1177/0098628315587618>
- Jönsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840-852. <https://doi.org/10.1080/02602938.2013.875117>
- Jönsson, A., & Panadero, E. (2017). The use and design of rubrics to support Assessment for Learning. In D. Carless, S. M. Bridges, C. K. Y. Chan, & R. Glofcheski (Eds.), *Scaling up Assessment for Learning in Higher Education* (pp. 99-111). Springer. https://doi.org/10.1007/978-981-10-3045-1_7
- Jönsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity, and educational consequences. *Educational Research Review*, 2, 130-44. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Klenowski, V., Askew, S., & Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment & Evaluation in Higher Education*, 31, 267-286. <https://doi.org/10.1080/02602930500352816>
- Lipnevich, A. A., McCallen, L. N., Miles, K. P., & Smith, J. K. (2014). "Mind the Gap! Students' use of exemplars and detailed rubrics as formative assessment. *Instructional Science*, 42(4), 539-559. <https://doi.org/10.1007/s11251-013-9299-9>
- Löfgren, H., Alm, F., Jönsson, A., Hultén, M., Markström A-M., & Lundahl, C. (2021). *Betyg i årskurs 4: En studie om bedömningspraktikerna på skolor som deltagit i försöksverksamhet med tidigare betyg* [Grades in year 4: A study of the assessment practices in schools that have participated in the trial with earlier grades]. The Swedish National Agency for Education.
- Nussbaumer, A., Dahn, I., Kroop, S., Mikroyannidis, A., & Albert, D. (2015). Supporting self-regulated learning. In S. Kroop, A. Mikroyannidis, & M. Wolpers (Eds.), *Responsive Open Learning Environments* (pp. 17-48). Springer. https://doi.org/10.1007/978-3-319-02399-1_2
- 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 (2012). <https://doi.org/10.1016/j.lindif.2012.04.007>
- 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, 149-183. <https://doi.org/10.1080/02103702.2014.881655>
- Panadero, E., & Jönsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Panadero, E., & Jönsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, 30, 100329. <https://doi.org/10.1016/j.edurev.2020.100329>
- Panadero, E., Jönsson, A., & Strijbos, J.-W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In D. Laveault,



- & L. Allal (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation* (pp. 311-326). Springer. https://doi.org/10.1007/978-3-319-39211-0_18
- Panadero, E., Jönsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74-98. <https://doi.org/10.1016/j.edurev.2017.08.004>
- 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 & Practice*, 21(2), 133-148. <https://doi.org/10.1080/0969594X.2013.877872>
- Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment Research and Evaluation*, 15, 1-9. <https://doi.org/10.7275/psk5-mf68>
- Ritchie, S. M. (2016). Self-assessment of video-recorded presentations: Does it improve skills? *Active Learning in Higher Education*, 17, 207-221. <https://doi.org/10.1177/1469787416654807>
- Sadler, R. D. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education*, 13(2), 191-209. <https://doi.org/10.1080/0305498870130207>
- Sadler, R. D. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179. <https://doi.org/10.1080/02602930801956059>
- Sadler, R. D. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, 67, 273-288. <https://doi.org/10.1007/s10734-013-9649-1>
- Smit, R., Bachmann, P., Blum, V., Birri, T., & Hess, K. (2017). Effects of a rubric for mathematical reasoning on teaching and learning in primary school. *Instructional Science*, 45, 603-622. <https://doi.org/10.1007/s11251-017-9416-2>
- Smit, R., & Birri, T. (2014). Assuring the quality of standards-oriented classroom assessment with rubrics for complex competencies. *Studies in Educational Evaluation*, 43, 5-13. <http://dx.doi.org/10.1016/j.stueduc.2014.02.002>
- Statens Offentliga Utredningar (SOU). (2007). *SOU 2007:28. Tydliga mål och kunskapskrav i grundskolan [Clear goals and knowledge requirements in compulsory school]*. Swedish Government Official Reports. <https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2007/05/sou-200728>
- Swedish National Agency for Education (2018). *Skolverkets allmänna råd med kommentarer. Betyg och betygsättning [Common advice of the Swedish National Agency for Education with comments. Grades and grading]*. The Swedish National Agency for Education.
- Tomas, C., Whitt, E., Lavelle-Hill, R., & Severn, K. (2019). Modelling holistic marks with analytic rubrics. *Frontiers in Education: Assessment, Testing and Applied Measurement*, 4(89). <https://doi.org/10.3389/feduc.2019.00089>
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in postsecondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281-294. <https://doi.org/10.1080/09695940701591867>
- Vasileiadou, D., & Konstantinos, K. (2021). Examining the impact of self-assessment with the use of rubrics on primary school students' performance. *International Journal of Educational Research Open*, 2(2), 100031. <https://doi.org/10.1016/j.ijedro.2021.100031>
- Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, 48, 135-147. <https://doi.org/10.1080/00461520.2013.794676>



Acknowledgments

The work presented here is based on an evaluation funded by the Swedish National Agency for Education, and the interviews were performed in collaboration with Håkan Löfgren (project leader), Fredrik Alm, Ann-Marie Markström, and Magnus Hultén from Linköping University, and Christian Lundahl from Örebro university.

Anders Jönsson

Anders Jönsson is professor of Didactics at the Faculty of Teacher Education, Kristianstad University, Sweden. His main research interests are in assessment, grading, and feedback.

