What are the advantages, or disadvantages, or criterion-referenced / rubric-based assessment?

Compiled by Ashleigh Ward and Joy McEntee


SUMMARY: Letter to the editor. Andrade responds to Kohn’s “The Trouble with Rubrics,” defending rubrics against charges that they cause the valuation of conventions over content and superficial engagement on the part of student writers.


ABSTRACT. This study suggests that students use rubrics to support their own learning and academic performance. In focus groups, fourteen undergraduate students discussed the ways in which they used rubrics to plan an approach to an assignment, check their work, and guide or reflect on feedback from others. The students said that using rubrics helped them focus their efforts, produce work of higher quality, earn a better grade, and feel less anxious about an assignment. Their comments also revealed that most of the students tend not to read a rubric in its entirety, and that some may perceive of a rubric as a tool for satisfying a particular teacher’s demands rather than as a representation of the criteria and standards of a discipline.


ABSTRACT. This paper describes a new ICT assessment tool that reduces multi-handling of marks, comments and scores specifically where professional judgement is involved. Whereas previous ICT applications in most fields of education have focused on student learning, this tool focuses on the lecturer task of the assessment process. Unlike many ICT based assessment tools, it enables the assessor to make and record professional judgements. This tool moves the marking/recording sheet off the desk and onto the desktop (computer screen). The aim in using the tool is to reduce the unproductive busy work of marking such as adding up of marks, recording, and spell checking of comments by hand, also it allow for many different views of the information, and to increase the time spent on feedback, reflection and moderation. The tool combines features of the word processor, spreadsheet and database applications, and paper-based marking. These features are described and discussed as working examples are presented. The paper also describes other features or possible uses of the tool, including the development and refinement of rubric-based scales; the recording of feedback by markers about the quality of the marking key during and after marking; the ability to compare one’s marks with those of the coordinator electronically; the evaluating and analysing of results; and the printing out of many different views of the data.

**ABSTRACT.** This study used *The Competent Speaker*, a rubric developed by the National Communication Association (S. P. Morreale, M. R. Moore, K. P. Taylor, D. Surges-Tatum, & R. Hulbert-Johnson, 1993), to evaluate student performance in general education public speaking courses as a case study of student skills and programmatic assessment. Results indicate that students taking the general education public speaking course are below satisfactory standards on five of the eight competencies defined by the National Communication Association and are above satisfactory standards on two of the eight competencies. Implications for this particular program, other communication departments, and communication across the curriculum in general education are discussed. We also offer suggestions for those in other disciplines or educational settings in the use of performance evaluation rubrics for assessing other student skills/knowledge and for training new teachers.


**ABSTRACT.** Although the rubric has emerged as one of the most popular assessment tools in progressive educational programs, there is an unfortunate dearth of information in the literature quantifying the actual effectiveness of the rubric as an assessment tool in the hands of the students. This study focuses on the validity and reliability of the rubric as an assessment tool for student peer-group evaluation in an effort to further explore the use and effectiveness of the rubric. A total of 1577 peer-group ratings using a rubric for an oral presentation was used in this 3-year study involving 107 college biology students. A quantitative analysis of the rubric used in this study shows that it is used consistently by both students and the instructor across the study years. Moreover, the rubric appears to be ‘gender neutral’ and the students’ academic strength has no significant bearing on the way that they employ the rubric. A significant, one-to-one relationship (slope = 1.0) between the instructor’s assessment and the students’ rating is seen across all years using the rubric. A generalizability study yields estimates of inter-rater reliability of moderate values across all years and allows for the estimation of variance components. Taken together, these data indicate that the general form and evaluative criteria of the rubric are clear and that the rubric is a useful assessment tool for peer-group (and self-) assessment by students. To our knowledge, these data provide the first statistical documentation of the validity and reliability of the rubric for student peer-group assessment.


**SUMMARY.** Republication of the foreword of Maya Wilson’s *The Trouble with Rubrics*. Kohn critiques the use of rubrics for assessment and feedback on the grounds that their use to legitimate grades tends to lead to superficial approaches on the part of learners, and on the grounds of scepticism about the valuation of speed, efficiency and consistency. He also challenges rubrics’ advocates’ claims that rubrics yield objective results. Particularly, he opposes the distribution of rubrics to students.

**ABSTRACT.** The article discusses a rubric developed to assess writing in interdisciplinary college courses. The authors discuss increases in the popularity of interdisciplinary curricula but note the use of grades and surveys to measure curriculum effectiveness. They describe a targeted assessment rubric designed to evaluate interdisciplinary writing by measuring the purposefulness, grounding in disciplinary understanding, integration and critical awareness in college papers. The rubric was applied to essays by students enrolled in Miami University’s interdisciplinary courses. Analysis of variance (ANOVA) was used to confirm the validity of the rubric.


**ABSTRACT.** Growing concern over marking reliability and standards within UK higher education has led in recent years to increasing interest in criterion-referenced assessment. As part of an ongoing research project this paper examines student experience of criterion-referenced assessment and, in particular, a criteria assessment grid developed for the Business School at Oxford Brookes University. The main findings of this paper report the evident thirst students voiced for more reliable marking processes and clearer guidelines on assessment requirements and criteria. However, whilst students applauded the criterion-reference grid as a ‘good idea’ and one that did help clarify assessment requirements they also highlighted several shortcomings. Such shortcomings included issues of subjectivity and multiple interpretations of criteria and standards. The paper concludes that the criterion-reference grid, from the student perspective, appeared to be of limited practical benefit unless presented as part of multifaceted approach including explanation, exemplars and opportunities for discussion.


**ABSTRACT.** In the UK, there is growing pressure both within and across institutions to establish common assessment standards. This is an account of the introduction of a common assessment grid in a School of Business. The intention was to develop a comprehensive marking criteria grid to establish common standards of marking and grading for Advanced Level undergraduate modules (those normally taken by second- and third-year students) across the Business programmes. The aim was to provide better guidance to students (and better work as a result) as well as enabling consistency in marking, easier moderation, and easier provision of feedback. It was created using the attributes identified by the Graduate Standards Programme and South-east England consortium (SEEC) generic level descriptors as starting points. While the grid failed to provide comprehensive marking criteria, evidence is presented to suggest that it succeeded in its other intentions. The project also raised questions about the efficacy of setting standards through programme specifications and generic level descriptors.

**ABSTRACT.** This paper critically reviews the empirical research on the use of rubrics at the post-secondary level, identifies gaps in the literature and proposes areas in need of research. Studies of rubrics in higher education have been undertaken in a wide range of disciplines and for multiple purposes, including increasing student achievement, improving instruction and evaluating programmes. While, student perceptions of rubrics are generally positive and some authors report positive responses to rubric use by instructors, others noted a tendency for instructors to resist using them. Two studies suggested that rubric use was associated with improved academic performance, while one did not. The potential of rubrics to identify the need for improvements in courses and programmes has been demonstrated. Studies of the validity of rubrics have shown that clarity and appropriateness of language is a central concern. Studies of rater reliability tend to show that rubrics can lead to a relatively common interpretation of student performance. Suggestions for future research include the use of more rigorous research methods, more attention to validity and reliability, a closer focus on learning and research on rubric use in diverse educational contexts.


**ABSTRACT.** The degree of interaction among participants in distance courses is widely acknowledged to be an indicator of successful learning experiences; interaction has been found to contribute to both achievement and student satisfaction. As an increasing number of distance courses move to a completely asynchronous, online format, providing for better interaction is an important means of assuring course quality. To allow measurement and study of interaction in online courses, five elements were identified that contribute to interaction, and a rubric was designed to assess the degree of each in distance courses. A previous use of the rubric in one online class indicated the instrument had good convergent and divergent validity and reliability in terms of consistency of results across students. The current study focused on establishing the rubric’s concurrent validity and consistency of results across four distance courses. Recommendations are made to increase the usefulness and reliability of the rubric in practice.


**ABSTRACT.** This paper reports the findings of a two-year research project focused on developing students’ understanding of assessment criteria and the assessment process through a structured intervention involving both tacit and explicit knowledge transfer methods. The nature of the intervention is explained in detail, and the outcomes are analysed and discussed. The conclusions drawn from the evidence are that student learning can be improved significantly through such an intervention, and that this improvement may last over time and be transferable, at least within similar contexts. This work is a development within a longer and ongoing research project into criterion-referenced assessment tools and processes which has been undertaken in the pursuit of a conceptually sound and functional
assessment framework that would promote and encourage common standards of assessment; that project is also summarised.


ABSTRACT. When assessment tasks are set for students in universities and colleges, a common practice is to advise them of the criteria that will be used for grading their responses. Various schemes for using multiple criteria have been widely advocated in the literature. Each scheme is designed to offer clear benefits for students. Breaking down holistic judgments into more manageable parts is seen as a way to increase openness for students and achieve more objectivity in grading. However, such approaches do not adequately represent the full complexity of multi-criterion qualitative judgments, and can lead to distorted grading decisions. Six anomalies in the ways assessors approach the grading task are identified, together with several likely contributing factors. Overall, the conclusion is that explicit grading models do not have as strong a theoretical foundation as is commonly supposed, and that holistic appraisal merits further investigation.


ABSTRACT. The paper reports on the findings of a Learning and Skills Research Centre (LSRC) funded study investigating the impact of different modes and methods of assessment on achievement and progress in post-secondary education and training. Data were collected across Advanced-level academic and vocational preparation programmes in schools and colleges, work-based training, and adult education settings. The paper reports that clarity in assessment procedures, processes and criteria has underpinned widespread use of coaching, practice and provision of formative feedback to boost achievement, but that such transparency encourages instrumentalism. It concludes that the practice of assessment has moved from assessment of learning, through assessment for learning, to assessment as learning, with assessment procedures and practices coming completely to dominate the learning experience and 'criteria compliance' replacing 'learning'.


ABSTRACT. Reframing the Great Rubric Debate beyond the taking of sides, Eric D. Turley and Chris W. Gallagher propose a set of questions designed to help educators assess the value of rubrics. They contend that these four simple questions push educators to deliberately articulate "judgments about the uses of rubrics" and why they will or will not use them for assignments.

ABSTRACT. The traditional "black box" approach to evaluation of assignments in educational research courses has at least two effects: (1) products that fail to meet the expectations of the instructor; and (2) frustration on the part of students who do not know exactly what is expected, and who are consequently confused about or disappointed in the grades they receive. Solutions to these problems include developing checklists and rubrics that break complex tasks, such as writing research proposals and master's theses, into component parts, using the checklists and rubrics to guide students' work, and then grading the students on how well they complete each component. Included in the appendix are sample checklists for scoring content, quality of writing, and adherence to American Psychological Association (APA) style in research proposals, as well as rubrics for master's theses and group research project presentations.