

Facilitating Student Learning and Engagement in Higher Education through Assessment Rubrics

Facilitating Student Learning and Engagement in Higher Education through Assessment Rubrics

Edited by

Peter Grainger and Katie Weir

Cambridge
Scholars
Publishing



Facilitating Student Learning and Engagement in Higher Education
through Assessment Rubrics

Edited by Peter Grainger and Katie Weir

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Copyright © 2020 by Peter Grainger, Katie Weir and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-4345-5

ISBN (13): 978-1-5275-4345-4

CONTENTS

Acknowledgement.....	vii
Foreword	ix
Ernesto Panadero	
Chapter One.....	1
Introduction: The Significance of Rubrics	
Peter Grainger and Katie Weir	
Chapter Two.....	9
Understanding Rubrics	
Katie Weir	
Chapter Three	25
Rubrics as a Tool for Self-regulated Learning	
Anders Jönsson	
Chapter Four.....	41
The Efficacy of the Continua Model as an Alternative Grading Tool: Student Voices	
Peter Grainger	
Chapter Five	57
The Interplay between Exemplars and Rubrics	
Philip Smyth, Jessica To and David Carless	
Chapter Six.....	69
Creating Quality Rubrics through Conversations	
Peter Grainger and Beverly Dann	
Chapter Seven.....	87
The Rubric as a Moderating Tool in Tertiary Contexts	
Lenore Adie	

Chapter Eight.....	97
Talking about Rubrics in Higher Education: Exploring Academic Agency	
Deborah Heck	
Chapter Nine.....	111
A More than Casual Approach: An Argument for a Team Approach	
to Assessment and Rubric Design	
Gail Crimmins	
Chapter Ten	123
Critiquing and Writing Rubrics: A Law Discipline Exemplar	
Kelly Burton	
Chapter Eleven	139
Perspectives on Rubric Design and Construction	
Katie Weir and Nicholas Charlton	
Contributors.....	153
References	157

ACKNOWLEDGEMENT

The editors would like to acknowledge funding from the University of the Sunshine Coast Centre for Support and Advancement of Learning and Teaching, Learning and Teaching Grant Scheme, for funding which enabled the research for chapters six and eight to occur.

FOREWORD

ERNESTO PANADERO

Rubrics have become commonplace in educational settings across all instructional levels and throughout the world. Educators and policymakers rely on and put their trust in assessment rubrics because they believe these tools can improve the reliability and validity of assessment instruments and the judgments made by assessors. Nevertheless, there are occasions where rubrics are imposed on teachers, or where rubrics are poorly designed, or their implementation is flawed. These occasions do not only decrease the quality of teaching and student learning, they also undermine the efficacy of rubrics. To avoid this, educators and policymakers need to be better informed about assessment rubrics, especially regarding their design and effective implementation.

Importantly, rubrics have received their fair share of criticism. Some scholars and educators blame rubrics for shifting the focus from learning toward assessment, whereas others argue that rubrics short-circuit creativity. Many of these perceived flaws, however, have not been empirically tested. Actually, studies that have examined the effectiveness of rubrics consistently agree: rubrics are highly effective learning tools when they are well designed and appropriately implemented. Typically, the task of selecting rubrics or creating new ones falls on the shoulders of teachers as assessors, hence assessors need to be informed about what constitutes an exemplary rubric. Unfortunately, finding resources that present a comprehensive overview of rubric use in higher education is currently a difficult task. Learning materials for academic teaching staff, such as the book you have in your hands right now, are sorely needed.

This edited volume has a clear and unifying vision on how rubrics can be used to enhance student learning while, at the same time, distilling more specific areas of work in crucial topics revolving around this instructional tool (e.g. self-regulated learning, combined use with exemplars). The different chapters evoke a potent message: the combination of well-designed rubrics with the right instructional ingredients can be one of the most strategic practices in any educator's repertoire. The positive effects of rubrics, when properly implemented, are

brilliantly brought forward by the contributors. Because of all of the above, this volume will serve as a useful guide for researchers, students and practitioners alike, offering a comprehensive overview of current inquiry into rubrics.

Ernesto Panadero
Educational researcher at Universidad de Deusto (Bilbao, Spain)
www.ernestopanadero.es

CHAPTER ONE

INTRODUCTION: THE SIGNIFICANCE OF RUBRICS

PETER GRAINGER AND KATIE WEIR

Since the late twentieth century when there was a seismic shift in approaches to assessment in higher education, rubrics have become a ubiquitous tool for judging student learning. This is because rubrics are more effective for judging the quality of student learning against pre-set criteria. Prior to adopting criterion-referenced frameworks for assessing student learning outcomes, the approach to assessment in higher education was predominantly normative whereby student learning was quantitatively evaluated to enable the rank ordering of students against a statistical norm. Not only did this approach focus on what students didn't know, it failed to acknowledge the role that assessment can play in enhancing learning.

The wholesale adoption of rubrics across university faculties has been staggered and research on how rubrics were being implemented was scarce in the early days. However, scholars in the field agree that research on rubric use on higher education is increasing (Dawson, 2017) and one aim of this book is to bring together some extant research on rubric use for specific learning purposes, and in specific higher education contexts. It synthesizes the findings from several reviews of the literature around rubrics in higher education (e.g. Brookhart, 2018; Brookhart and Chen, 2015; Panadero and Jonsson, 2013; Bell, Mladenovic and Price, 2013; Reddy 2011) and juxtaposes these against some recent ideas and research in an attempt to broaden our understanding of the nuanced ways in which these tools are being used to enhance the university student learning experience.

Assessment practices in Australia's higher education institutions have been under the spotlight in recent times because of an increased emphasis the quality of education on offer. This situation has been partially influenced by student dissatisfaction with their university learning experiences evident in student evaluations of courses and programs.

Assessment in higher education is often criticized for its lack of transparency and is characterized by academics mystifying assessment processes and presenting students with opaque and sometimes misaligned evaluation criteria. This confirms what Ecclestone (2001) once described as a negative image of subjective, anecdotal, even negligent, assessment.

We know that student engagement is driven by successful assessment experiences (Boud & Associates, 2010), and shaped by the clarity of assessment rubrics (Grainger & Weir, 2015). Rubrics typically contain three elements: criteria, standards and standards descriptors. Rubric clarity is manifest in standards descriptors that accurately capture the quality of learning expected through precise verbal descriptions that enable unambiguous evaluations of student responses to assessment. The description of quality identified for each standard is matched against evidence in student work. If the description of quality is not explicit then the matching exercise becomes problematic due to different interpretations of quality by different assessors, resulting in inconsistent judgments. Hence, rubric deficiencies represent major challenges to what Sadler (2010) refers to as grade integrity.

Rubrics can also act as a feedback mechanism because it links the evidence in student work to the standard descriptors, hence providing students with guidance as to why a standard was awarded. Feedback on student work is essential to student learning and effective comments on student work are considered an aspect of quality teaching practice (Hattie, Biggs, & Purdie 1996; Black & Wiliam 1998; Ramsden 2003). However, although there have been numerous studies carried out into the effectiveness of feedback for improving student learning, empirical evidence shows that while feedback communications may be appreciated by students, they often lead to little if any improvement in subsequent submissions (Sadler, 2009). This lament is borne out in evidence from Carless (2006) who noted student misunderstanding of the discourse of the rubric as being a barrier to student understanding of the quality of work expected.

A key quality assurance mechanism in assessment is moderation, a process that involves assessors comparing judgments about the quality of a student's assessment response. It is a form of calibration that ensures greater consistency of evaluative judgments which is dependent on rubric clarity and interpretability. Although moderation is institutionalized as accepted practice there appears to be limited understanding its contribution to teaching and learning, and significant confusion in relation to shared understandings of criteria, standards and the qualities that provide evidence of a standard (Sadler, 2010). Therefore, the rubric must be

viewed as key to quality assurance and accountability and fundamental to improving all stakeholders' assessment experiences and ultimately, student learning outcomes.

Assessors and students are impacted prior to grading as they engage with the rubric and deconstruct its meanings. Assessors are impacted at the point of grading as the rubric is used to match evidence of student work with standards and standards descriptors. Assessors are also impacted during the moderation process which uses the rubric to ensure consistency of judgments, and finally, students are impacted post grading when they receive results via the rubric. If the rubric is vague, students are reliant only on assessors' annotations for guidance and these too, are the subject of severe criticism as being too difficult to decode (Sadler, 2009). The rubric is therefore, the most significant assessment artefact and yet, until recently it has not been a focus of educational assessment research and scholarship.

The role of rubrics in making decisions about students' grades is one aspect of these assessment artefacts that Sadler (2005) views as particularly problematic. Universities in Australia have various procedures for awarding and monitoring the distribution of grades to students but criterion referenced assessment decisions are the norm for most tertiary institutions. Criterion referencing attempts to establish grading decisions by reference to some external and relatively fixed set of criteria. In principle, a student's work is judged against the relevant criteria in isolation from the work of other students. In addition to criteria, there has been a renewed focus measuring the standard of student achievement and this has given rise to what is commonly referred to as a criteria and standards-based assessment framework that is reliant on rubrics.

Despite university policies that mandate the use of criteria and standards, our experience indicates that for some students and academic teaching staff, the actual process of measuring student achievement remains a mystery. We suggest that a major reasons for the opacity of assessment practices is an inability to create a valid and reliable assessment rubrics that feature criteria aligned with the course learning outcomes and precise descriptions of the expected quality of learning being assessed. In most cases we have found that criteria are generally accurate in identifying aspects of learning behaviour being assessed, however, the standards descriptors associated with each criterion are too often subjective and open to interpretation. Often these standards descriptors provide insufficient guidance for assessors because they are differentiated by the use of subjective terminology. The key difficulty is often that the real standards are locked inside the marker's head as tacit

knowledge and have not been explicitly stated to students. This is not only a problem for students when interpreting the expectations, but a problem for assessors when moderating. Standards descriptors that fail to discriminate standards are a significant problem with many rubrics, resulting in inconsistency of assessment judgments by academics.

As a result of these deficiencies, and despite the existence of common criteria and known standards, identified in rubrics, assessors may grade the same piece of assessment differently (Grainger, Purnell, & Zipf 2008).

When judging the quality of a student's assessment response, the assessor must possess a concept of quality appropriate to the task and be able to judge the student's work in relation to that concept (Popham, 2005). In other words, a strong understanding of the knowledge, values and expectations is needed to ensure the assessor's judgment is also of quality. This knowledge can only be utilised if the assessor has had the experience of marking the task previously, in short, has the tacit knowledge.

According to Race (2006), academics who are experienced assessors do possess tacit knowledge of what quality looks like in student work. This concurs with Sadler's (2011) observation that competent appraisers can consistently identify quality when they see it. This tacit knowledge has been shown to enable assessors to make accurate interpretations of sometimes vague descriptions of student behaviour in order to discriminate between standards or levels of achievement (Grainger, Purnell, & Zipf 2008), which means there is no compromise of assessment integrity and reliability in terms of assessor judgements. However, not all academics understand or are experienced with sound assessment practices and it takes some years to get to know how to align evidence of quality with relevant achievement standards and achieve consistency of judgement over time. For assessors who are unclear about learning quality, vague assessment grading tools are not, in fact, objective arbiters of performance, nor are they defensible, nor do they encourage consistency of teacher judgements.

In this way, even an experienced assessor, who marks a task for the first time, can be referred to as a novice assessor (Grainger, Adie, & Weir 2015). Where there are multiple markers, and even multiple novice markers, the results of marking and moderation are even more problematic. Hand in hand with this, and prior to grading, is the very real issue of creating a new rubric for an assessment task, for the very first time. That is, without experience, the rubric creator can only imagine how students will respond to the assessment task which makes rubric design and construction an inexact science at best.

This introduction to our book has highlighted the many and varied problems with rubrics and their implications for teaching and learning in

higher education. The remainder of this book aims to resolve some of these problems with rubrics by presenting research that examines rubric design and use in a range of different university contexts. All the research presented here is framed by the perspective of assessment for learning purposes and the role that rubrics can play in enhancing student learning outcomes at university.

To achieve our goals in writing this book, the next chapter (Chapter Two) sets the scene for this book by outlining the multiples roles rubrics hold in the assessment process. Weir situates this information in the context of our increased understanding about how integral assessment is for effective learning. She begins by deconstructing the meaning of the term rubric, highlighting the often-simplified definitions that underlie a construct that is surprisingly complex and complicated. The chapter proceeds with details about the many roles rubrics have in outlining assessment expectations for students and assessors, in making valid judgements about students' learning, in accountability to the course curriculum and transparency of assessment practices, and for formative assessment purposes to enhance student learning. Based on extensive experience in criterion-referenced assessment, Weir suggests that more effective, targeted professional development for academic teaching staff around rubric design and implementation would result in increased student satisfaction with university teaching and learning.

In Chapter Three, Jonsson examines student self-regulated learning (or SRL) as fundamental aim of higher education and reviews the research on how rubrics can facilitate this practice. He begins by explaining Zimmerman's (2013) cyclical model of SRL and the role rubrics can play in each of its three phases. When reviewing the relevant literature Jonsson has two aims: the first is to examine the efficacy of rubrics for SRL; and the second goal is to look how rubric design and implementation can affect SRL. In the first instance, the research on rubric efficacy for SRL shows that most studies indicate rubrics do play a role in improving SRL (along with improving student performance on task), however some findings were inconclusive or contradictory. The research regarding rubric design and use also generated some mixed results but the overall conclusion is that rubric "accessibility" enhances both SRL and the learning achievements. Jonsson calls for more research in both areas to gain a better understanding of how students use rubrics and the effects of student co-construction of rubrics on their learning achievements.

In Chapter Four, Grainger's research responds in part to Jonsson's call by investigating how students interpret and use an alternative model of rubric design: the continua model. Grainger compares the features of this

model rubric to the more familiar rubric with a matrix format. He then reports on a study he conducted into its efficacy for a small group of teacher education students enrolled in his course on assessment. The findings show that the participants reacted positively to the unique design features of the continua model perceiving it to be more flexible and easier to construct than traditional rubrics.

Chapter Five also tries to fill some gaps in rubric research about how teachers communicate the success criteria and standards to students. rubrics. In this chapter Smyth, To and Carless draw on findings from a larger study about the interplay between exemplars and rubrics that is based on Sadler's (2005) argument that exemplars reveal more about quality work than rubrics can alone. The authors report on the two variations on rubric and exemplar use that highlight tensions between transparency and instrumentalism and plagiarism. These studies demonstrate some benefits from using using exemplars but suggest a need for more studies that show how students transfer the insights gained from these (and from rubrics) to their own work.

The next four chapters are all centered around professional conversations between assessors and the implications of these for the social practices associated with rubric design, construction and use. In Chapter Six, Grainger and Dan present the findings from a small action research project aimed at improving the quality of rubrics through a reflective process of problem solving, working with other academic staff in triads as small communities of practice. Analysis of the conversations held between triad members as they worked collaboratively to construct a rubric identified the need for clarity, for explicit descriptions of behavior, and for words that are devoid of subjective interpretation and that can readily differentiate standards for the benefit of both student and assessors. These robust and often challenging conversations reduced the divide between novice and expert assessors by sharing tacit knowledge which led to improved consistency in marking.

Chapter Seven shifts the focus from conversations around constructing rubrics, to dialogue centered around rubrics as tools for moderating judgments about students' assessment performance. In this Chapter Adie examines the use of rubrics for social or consensus moderation that aims to increase the consistency and comparability of judgments. She explains that during these targeted conversations the quality of a variety of students' responses to an assessment task may be interrogated. Her findings show that these conversations hinged on discourses about the different purposes of moderation including equity, justification, community building and accountability. Adie also discusses the implications of some important

factors often overlooked in rubric and assessment research in higher education contexts, such as the increased casualization of the workforce, online delivery of course curricula and student dialogue as a type of informal moderation.

In Chapter Eight Heck explores the notion of academic agency through conversations about rubrics held between a course coordinator and a casual tutor. The specific aim of their conversation was to evaluate the efficacy of the rubrics (and assessment) from a recently completed course of study. The focus of Heck's research is on how academics engage dialogically and draw on academic agency as they improve their assessment literacy and practice. The research found academic agency to be evident in two key revision commitments they made about revising the course assessment. Heck suggests that professional conversations between teaching teams immediately following the delivery of a course can provide time for academics to reflect on their professional practice and develop their assessment literacy.

Professional conversations are also at the core of Crimmins' research in Chapter Nine but here the spotlight is on dialogue with casual academics. This study responds to Adie's request for further examination of the implications of increased casualisation of the workforce in higher education by presenting a balanced argument for the collaborative co-construction of rubrics by all members of a course teaching team. Crimmins presents a model that involves a systematic approach to engaging casual academics in the design of assessment tasks and rubrics. She concludes the article by suggesting that this approach can enhance quality assurance of assessment, can provide much needed professional development for casual academics, and facilitate their social integration into a community of practice.

In the penultimate chapter of this book, Chapter 10, Burton describes her learning journey implementing criterion-based assessment in the Law discipline. She presents the program-level rubrics she designed for evaluating threshold learning outcomes in a Bachelor of Laws degree. She reflects on the problems associated with designing rubrics for uploading to learning managements systems, and also discusses the effects and implications of co-creation of rubrics with law students.

The book concludes in Chapter 11 with practical advice from Weir and Charlton on how to construct an holistic rubric for making summative judgments about the quality of student learning performance for an authentic assessment task. The authors step the reader through the design process demonstrating how to establish constructive alignment with the course curriculum. They also explain how to write explicit, task-specific

criteria using cognitive verbs that match those of the course learning outcomes. Once constructed the same rubric is then transferred to the continua model to further illustrate the design features of this alternative rubric introduced by Grainger in Chapter Four.

CHAPTER TWO

UNDERSTANDING RUBRICS

KATIE WEIR

The aim of this chapter is to expand our understanding of what rubrics are and what they do. It highlights the complexity of rubrics and explores the many roles they play in the assessment process. The chapter begins by foregrounding the assessment theories and perspectives about student learning that frame how rubrics are represented and discussed. The term “rubric” is then unpacked by reviewing its etymology and various scholarly definitions and examining how it came to be such a ubiquitous tool to evaluate learning. Deconstructing rubrics as an educational construct paves the way for a deeper understanding of how rubrics function in assessing and enhancing student learning. This chapter outlines a number of significant roles that rubrics play in the teaching – learning – assessment cycle. These rubric functions are discussed in terms of assessment accountability and transparency, validity, quality assurance and moderation of assessment results, and formative assessment. Relevant research is reviewed to examine some implications of these roles for course assessment design and implementation.

Assessment and learning

This chapter rests on the axiom that effective learning only occurs when it is assessed, and feedback is provided to link students’ learning achievements with the success criteria. In other words, learning needs an assessment “event” to establish the scope and depth of what has been learned. The notion that assessment is crucial for learning to occur signifies the necessity for educators to embrace assessment and to develop an assessment mindset (see Masters, 2014b) whereby assessment is at the forefront of all curriculum planning and implementation. The significant role of assessment in the learning process also underpins the role that rubrics can play in enhancing students’ learning outcomes.

According to Masters (2014a, p.1), assessment has one purpose, which is:

to establish and understand where learners are in an aspect of their learning at the time of assessment ... [which] usually means establishing what they know, understand and can do.

This means that assessment is primarily designed to gather and elicit evidence of student learning. It is what teachers do with that evidence that determines whether the assessment is for formative, diagnostic, or summative purposes.

Based on Masters's explanation of the fundamental purpose of assessment, an assessment "event" refers to the act of establishing what learners know and can do and evaluating their 'performance' against criteria that accurately capture the desired learning behavior(s). Effective feedback ensures that the assessment event can potentially improve student learning.

Conceptualizing assessment as necessary for effective learning paves the way for educators to acknowledge its pedagogical importance and possibly re-think their strategies for implementing summative assessment tasks in their courses of study. This is especially important when tertiary education is conceived within a standards and accountability framework that demands a more transparent approach to university assessment. This chapter situates the rubric in this educational context and highlights its integral accountability role in aligning the course learning outcomes with the assessment criteria and implementing appropriate pedagogical strategies.

Unpacking rubrics

Having established the link between assessment and learning, the focus of this chapter narrows to examine the importance of rubrics for indexing the learning that "counts"—that is, "Rubrics make the learning target more visible to students, ... [and they] guide teaching by highlighting for both students and teachers the key concepts in the assignment or course" (Wolf, Connelly & Komara 2008, p.29). Furthermore, pre-set criteria "help clarify assessment intentions and processes" (Carless, 2015, p.138). These statements suggest the import of rubrics for the effective delivery of a course of study and they presume the use of rubrics for summative assessment purposes—that is, assessment that evaluates the extent to which students can apply the key concepts, knowledge, skills, and attitudes related to the course learning outcomes for reporting and/or certification purposes.

It is now widely recognized and understood that rubrics can play an important formative assessment role in monitoring and enhancing student learning. However, in criteria and standards-based assessment frameworks, the role of rubrics for evaluating student achievement cannot be underestimated in an era of increased curriculum accountability, hence the focus here is on their utility in summative assessment.

When we talk about rubrics, the underlying assumption is that educators recognize and understand the concept and how to design the construct. Apart from the interchangeability of terms used to describe a rubric—grading tools, scoring matrices, criteria sheets, marking guides—recent research shows that the presumption of a shared understanding of what constitutes a rubric is, at best, spurious. For example, Dawson (2017, p.348) discusses the exponential increase in research literature around rubrics and explains that:

The problem with the combination of conflated meanings and rapid proliferation of “rubric” is that it has been mandated, evaluated and built into technology, often as if here was some sort of shared understanding. Some departments or institutions have adopted policies that mandate the use of a “rubric” without providing a working definition of the term, leaving it open to a very diverse array of interpretations that may not be in the spirit the policy-makers intended.

Dawson’s (2017, p.357) insights led him to propose a framework for understanding rubrics (discussed below), which, he claims, “demonstrates the breadth of assessment practice undertaken under the banner of that one word”. These insights also necessitate that the term “rubric” needs unpacking to establish a deeper understanding of its complexity and to disrupt a common perception that rubrics are simply a practical tool to evaluate student learning.

The process for deconstructing the meaning of “rubric” begins with a brief examination of the etymology and common definitions of the term, followed by a more scholarly review of research on rubrics. The aim is a shared understanding of rubrics that is underpinned by an appreciation that meanings are context-dependent and thus, whatever definition this review generates, cannot be definitive.

According to the *Oxford English Dictionary* (2018), the origins of “rubric” all associate the term with something written in red to make it distinctive. This red writing has then been translated into Modern English as a noun that refers to a title or header on a document or, alternatively, a set of instructions, guidelines, or procedural rules. These interpretations suggest why rubrics have been reconfigured for educational purposes as a

set of guidelines to assess student learning. According to Popham (1997, p. 72), this occurred around the 1970s when:

[M]easurement specialists who scored students' written compositions began to use the term to describe the rules that guided their scoring. They could have easily employed a more readily comprehensible descriptor, such as *scoring guide*, but *scoring guide* lacked adequate opacity. *Rubric* was a decisively more opaque, hence technically attractive, descriptor.

Despite this opacity, assessment scholars would then attempt to define rubrics in relatively simple ways. For instance, one of Popham's (1997) early explanations of rubrics is a purely structural perspective on their standard components: evaluative criteria, quality definitions for those criteria at particular levels, and a scoring strategy. Another version on the same theme is later provided by Reddy and Andrade (2010), who construed rubrics as documents that articulate the expectations of an assignment by listing the criteria for what is particularly important and by describing levels of quality on a scale from excellent to poor. More recently, Brookhart (2018) described rubrics as having two parts: criteria that express what to look for; and performance level descriptions that index how the criteria are instantiated in students' work at varying levels of quality. Brookhart points out that it is the performance level descriptions across a continuum of quality that distinguish rubrics from other assessment tools such as checklists or rating scales and that this feature is what enables them to be used for formative, as well as summative, purposes.

Rubrics are also often distinguished from other grading tools by their evaluative function, with different emphases on the type of assessment or the way they are used. For example, Jonsson and Svingby (2007, p. 131) characterize rubrics as:

a scoring tool for qualitative rating of authentic or complex student work. It includes criteria for rating important dimensions of performance, as well as standards of attainment for those criteria.

In contrast, Carless (2015, p. 132) describes rubrics in terms of how they are used to make judgments about the standard of work, viewing these scoring matrices as a combination of:

criteria and standards in the form of a grid to provide statements of the performance needed to achieve different standards. They can involve holistic interpretation of criteria in which overall judgement is made;

analytic schemes in which judgements are made in relation to individual or separate criteria; or some combination of the two.

Using a framework that has fourteen features or “design elements”, Dawson (2017) makes a more recent attempt to capture the breadth of assessment practices associated with rubrics. Some of the design elements are familiar and expected—for example, evaluative criteria, scoring strategy, judgement complexity, quality definitions, and specificity. The remaining design elements—including such factors as the creators of rubrics, the users and uses, secrecy, exemplars, quality processes, accompanying feedback information, presentation, and explanation—serve to demonstrate the complexity of the meaning of the term and its possible signifiers.

This scope on the different ways to describe rubrics demonstrates some interesting parallels with relatively recent changes to the way assessment is conducted in higher educational contexts. For instance, the widespread use of rubrics emerged in response to a shift from traditional testing, quantitative evaluations, and normative ranking of students’ achievements to the current emphasis on implementing a broader array of assessment instruments that employ criteria and standards to make qualitative judgments about what a student knows and can do with reference to the course learning outcomes. The problem with any confusion around the term appears to have arisen because the shift to criteria-based assessment in higher education was unaccompanied by targeted professional development for academic staff around the design and use of rubrics.

Based on the literature reviewed here, any definition of a rubric must consider the current educational context framed by standards and accountability. It also must consider the fact that a rubric is a text that has different roles and target audiences. For the purpose of this chapter, a rubric is defined as a template for presenting the assessable elements of a task (the criteria for marking) and a description of the degree of quality (or standard descriptor) expected for each criterion across a range of achievement standards. For assessors of student learning, rubrics are designed for evaluating student achievement in an assessment event and to demonstrate alignment between the course learning outcomes and what is being assessed. For students, rubrics can guide and monitor their learning progress towards the assessment event and indicate the assessment expectations—that is, what assessors are looking for and the quality or degree of the learning performance.

Having established what is meant when talking about rubrics, the discussion now centers on the major roles of rubrics in assessing student learning outcomes and enhancing learning achievements in higher education. The next section explores some significant functions of rubrics

across the teaching – learning – assessment cycle and reviews relevant research to examine related pedagogical and curricular implications.

How rubrics function in assessment and learning

As previously mentioned, at its basic level a rubric is a text that has varied audiences and purposes before, during, and after an assessment event. As texts, rubrics move through different social fields, each with its own set of social actors who employ rubrics in different ways. These social actors include designers of rubrics, assessors who employ rubrics during marking, students for whom rubrics highlight valued knowledges and skills and index the quality of work expected, and a range of other educational stakeholders such as program directors, faculty directors of teaching and learning, and even industry regulators who view rubrics for accountability purposes. This perspective on rubrics facilitates a deeper understanding of the various functions that rubrics can adopt to enhance assessment design and implementation in higher education. Four major functions of rubrics are explored here, including roles associated with assessment accountability and transparency, making valid judgments about student learning, quality assurance and moderation, and formative assessment strategies that can enhance student learning. Each role is explored in turn, with comment on the implications for course curriculum and relevant pedagogical strategies.

Rubrics as evidence of accountability and transparency

University assessment for summative purposes is perceived as “high stakes” for all higher education stakeholders. This is especially true when teaching and learning is situated within a quality standards and accountability framework approach to measuring learning outcomes. In this current climate, rubrics play an important role in ensuring that the teaching and learning outcomes of a course of study are accountable to the course/program curriculum, to faculty goals, to students’ learning expectations and entitlements, to university policies and guidelines, to professional associations, and to broader societal expectations of higher education.

Maintaining this chapter’s focus on summative assessment in a single course of study, rubrics (and their associated task guidelines) represent the quality (and quantity) of learning that is valued and important enough to be assessed. In this role, rubrics provide evidence of accountability to the course curriculum by employing the same discourse used to describe the

course learning outcomes and the teaching and learning activities. If careful consideration is given to the language used to construct rubrics, they can provide visible and valid evidence of “constructive alignment” (Biggs 2014) between the course curriculum, assessment, and pedagogy.

The rubric has another important role associated with accountability, that of enabling transparency in assessment practices. In higher education, the past few decades have witnessed considerable efforts to improve fairness and transparency in marking, securing appropriate standards and consistency in grading by demanding explicit criteria within marking rubrics (Bloxham et al. 2015). The rationale underpinning the gradual mainstreaming of rubrics into higher educational policies and practices is described by Bennett (2016, p. 67) as:

premised on the need to restore trust (e.g. accountability and transparency) in academic judgement and grading, our teachers and, by extension, higher education itself. The introduction and promotion of rubrics is part of a strategic response to the current crisis of confidence in educational standards brought about, at least in part, by increasing fiscal constraints placed on universities, attendant casualisation of staffing and quality assurance and risk auditing.

In contrast to this instrumental perspective, Sadler (2007) adopts a more philosophical approach and claims that the rise of transparent assessment practices in higher education is first and foremost about ethics and equity. That is, identifying criteria and specifying them in advance is about the right of students to know how the quality of their work is to be judged, and that all student responses to the same task will be assessed according to the same criteria.

Transparency is an essential attribute of quality assessment instruments and refers to the clarity of assessment expectations for students and the clarity of procedures for making judgments about a student’s work. Transparency has also been defined in terms of clear and explicit instructions and criteria for marking (Carless, 2015), and the need for tasks to be clear, understandable, and do-able (Tillema, Leenknecht & Segers, 2011; Readman & Allen, 2013). However, Tierney (2014) describes transparency as an aspect of “fairness”, which he sees as a multifaceted quality in classroom assessment with the primary purpose of supporting student learning through communicating criteria. According to Bamber (2015), the benefit of transparent assessment is its contribution to increased stakeholder confidence in the rigor and robustness of assessment systems and processes.

For Sadler (2008), transparency means clear and explicit criteria that provide students with the information they deserve about how their work will be judged before they commence a course of study. He also explains the ethical standpoint that students need assurance that the same criteria will be used to assess the whole cohort undertaking the same assessment. In this way rubrics make the assessment process accurate and fair because they are used to identify the most important aspects of the course and assessment task for both educators and students (Wolf, Connelly, and Komara 2008, p. 29). Worth (2014) further advocates for academics to operate with a sense of fairness to support student engagement with assessment criteria. In a keynote address, Sadler (2008) sums this notion up in his claim that “we have to be truthful” when making complex judgments to ensure that the assessment experience using rubrics is reliable and equitable, since marking is an ethical process.

This increased emphasis on assessment accountability and transparency means that the process of developing rubrics requires tremendous investments of time and effort (Reddy, 2011). However, other functions attributed to rubrics in tertiary teaching and learning contexts need to be examined and taken into consideration when planning to employ rubrics for summative assessment purposes.

Making valid judgments about student learning

For assessors, the primary role of a rubric is to facilitate an informed and valid judgment about the evidence of learning generated by the assessment event. This means that the assessable elements and their degree of quality must be observable in the students’ work and this has implications for the language choices made when constructing a rubric. The criteria and standards descriptors must be carefully crafted to narrow interpretations of the text by all end-users. This also highlights the importance of deconstructing the text with the team of assessors to ensure valid and reliable judgments are being made about the quality of work presented.

Recent research on the validity of judgments made using rubrics casts doubt on how well a rubric can fulfil this role, especially for authentic or performance-based tasks. For instance, Bennett (2016) explores an emerging discourse that contends that assessment rubrics are not capable of measuring and evaluating complex thinking skills. Bennett’s (2016, p. 50) research focused on assessing students’ writing skills and her results suggest rubrics can potentially:

limit the independent responses of students and the professional judgement of markers, encourage compliance jeopardising student commitment and

creativity, and promote a false sense of objectivity in the marking and grading of student work.

These claims concur with Humphrey and Heldsinger's (2014) research on the use of matrix-style rubrics for evaluating students' narrative writing. These authors claim this rubric format can create a threat to valid performance assessment and describe the problem as one of structural alignment, questioning whether the same numbers of gradations of quality faithfully capture that which is observed in student performances for each criterion. They suggest that it is more convenient to construct rubrics in this way for ease of marking; however, their research shows that structural alignment forces ratings to be artificially alike, thus limiting assessors from capturing variation in separate aspects of the construct. In other words, structural alignment leads to a *halo effect*—that is, the tendency for ratings on separate criteria to reflect “a general rater impression of performance” (Humphrey & Heldsinger, 2014, p. 254). They conclude that their research casts doubt on whether rubrics with structurally aligned categories can validly assess complex skills.

Other studies on the validity of rubrics explore how assessors use and interpret the language in the text and any threats to construct validity that this poses. One example is Bloxham, Boyd, and Orr's (2011) research that examined the nature of criteria used by experienced assessors and the differential meanings they assigned to each criterion to determine if this was a potential cause of variability in the awarding of a mark/grade. Their results suggest:

that additional criteria are used above those that are explicitly stated. A key difficulty with this apparent practice is the potential for assessors to vary in the additional criteria they use, and the likelihood that they may not be conscious of all the criteria they use in making judgements. (p. 474)

These authors further explain that even though assessors appear to use similar criteria, they often interpret such criteria differently, which has the potential to threaten the validity and reliability of judgments made. This research led the authors to advise universities to be more honest with themselves and with students, and to be open about the complexities involved in applying assessment criteria and that there is “rarely an incontestable interpretation of their meaning” (Bloxham, Boyd & Orr 2011, p. 479).

This advice echoes sentiments from Sadler (2008) during a keynote address when he called for greater truth in the way educators portray to their students (and, I would add, to colleagues and the broader community)

how complex judgments are made. This suggestion followed his claims that pre-set criteria have “grave limitations” and he proffered a solution to the problem, which was to induct students into how holistic judgments are made. Research published by Sadler the following year examined the claim that pre-set criteria lead to greater objectivity in evaluating student work. His findings show that “such approaches do not adequately represent the full complexity of multi-criterion qualitative judgements, and can lead to distorted grading decisions” (Sadler, 2009a, p. 159). More importantly, the outcome of Sadler’s research was that all rubrics have structural inadequacies, which he termed “indeterminacy”, whether the grading processes associated with them are holistic or analytic.

So far, this chapter has presented the potential benefits and pitfalls of rubrics. It has reported on research that demonstrates the role of rubrics in increasing the transparency of assessment practices in higher education, along with studies that question their ability to enhance the validity and reliability of the judgments they generate. This highlights the importance of understanding rubrics from both a discursive and sociocultural perspective in that they are texts that require a range of associated social practices to be implemented if they are to enhance student learning achievement. Some of these practices, such as social moderation pre- and post-assessment, and deconstructing the text with students and teaching staff, are discussed in detail in other chapters of this book.

The point here is that rubrics should always be approached as a work-in-progress: they are predictive texts in that they try to capture what the educator envisions students will submit. The effectiveness of a rubric for evaluating students’ assessment performance can only be measured when they are put to work, meaning when they are for judgment purposes. Only then can any unexpected nuances in student responses and discursive flaws in the template be identified and acknowledged in further iterations of the text.

Quality assurance and moderation

Despite the limitations of rubrics per se for making valid judgments, strategies are available to reduce inconsistencies in teacher judgment, such as the process of moderation. Rubrics have a significant role in this quality assurance process. Paramount for the success of moderation is the clarity of the task requirements, the criteria for marking, and the accuracy of the standards descriptors in capturing the expected quality of student performance. According to Rust, Price, and O’Donovan (2003), the role of rubrics in moderation is to help raise the quality of marking through

greater consistency (for a team of markers and for an individual marker) of the judgments being made—whereas Grainger, Adie, and Weir (2015, p. 4) explain the process as “a practice of engagement in which teaching team members develop a shared understanding of assessment requirements, standards and the evidence that demonstrates differing qualities of performance”. The purpose of moderation is to ensure that judgments are consistent and align with pre-set criteria and standards in order to quality assure assessment processes (Adie, Lloyd & Beutel, 2013).

For rubrics to fulfil this purpose, moderation needs to happen before, during, and after the assessment event so that the assessment intent is fully understood and the assessment instrument is reliable and produces credible results. Grainger et al. (2017) highlight that during the design phase of an assessment instrument, it is optimal for all teaching staff to be involved in the construction of the rubric, although university timelines and policies often prevent this from occurring. In this case, it is essential that rubrics are unpacked and shared between the team of assessors when the course of study commences. This strategy helps to narrow the interpretation of the text and can ensure that forthcoming assessment messages relayed to students are consistent across the teaching team.

During the assessment event rubrics are employed in their primary role to evaluate responses against pre-set criteria, and descriptors of quality and moderation ensure the rubric is applied consistently. Lenore Adie has written extensively on this topic and provides details about how moderation can occur in tertiary contexts in Chapter 6 in this book.

Another form of quality assurance in which rubrics play a role is when a course or program of study is evaluated. Although the role of rubrics in course or program evaluation is not well researched, Bharuthram’s (2015) study into the value of rubrics from a lecturer’s perspective highlights the need for a deeper understanding of how rubrics demonstrate constructive alignment and, ultimately, course and program accountability. Other findings indicate that few research respondents acknowledge the instructional value of a rubric and many view its creation as a one-off task and not as a fluid and recursive process. These findings, along with the dearth of research on this specific quality assurance role for rubrics, suggest that it is an aspect that requires greater attention because it points to the potential of rubrics as valuable tools for course curriculum review and also for whole-of-program review and development.

Rubrics and formative assessment

In addition to the role of rubrics in assessing student learning, they can also enhance student learning and assist in effective delivery of the course curriculum. In other words, rubrics open summative assessments up to formative purposes where specific pedagogical strategies are employed to improve student learning achievement. Panadero and Jonsson (2013) conducted a meta-analysis of research into the formative use of rubrics in order to investigate if, and how, rubrics have an impact on student learning. Their findings indicate that rubrics may have the potential to positively influence student learning, and that the use of rubrics can mediate improved performance in several ways, including by reducing anxiety, aiding in the feedback process, improving student self-efficacy, and supporting student self-regulation.

Rubrics can improve student learning because, along with the assessment instrument they accompany, they represent the learning goals of the course of study and what is valued in the course curriculum. As such, rubrics are imbued with the power to better inform students of the course learning intentions and, as previously discussed, render the assessment more transparent. This feature of rubrics has implications that can affect more positive relationships between teachers and students. Gallavan and Kottler (2009, p. 18) explain this in terms of the role of rubrics “in disrupting disempowering power relations between students and lecturers as they enable students to participate more actively in the assessment process.” The importance of involving the learner more intentionally in assessment practices is highlighted by Bourke’s (2016) research that focused on students using rubrics for self-assessment and demonstrated that this not only assists in their understanding of how they learn but also improves their understanding of the relationship between assessment and learning.

Recent research on the use of rubrics for formative assessment purposes in higher education highlights a range of implementation strategies for mediating delivery of the course curriculum. All the research reviewed for this chapter appears to agree that sharing the rubric with students is a fundamental requirement for transparency in the assessment process. Carless (2015, p. 141) is emphatic about this role of rubrics: “the extent to which criteria impact students’ learning is largely dependent on the quality and depth of their engagement with rubrics”. Additionally, research by Simpson and McKay (2013, p. 26) about the role of rubrics in supporting the uptake of deep learning concluded that: