Making the Grade: Theorising Musical Performance Assessment

Jack Walton

BMus(Hons), Griffith University

A thesis submitted in fulfilment of the award of Doctor of Philosophy

February 14th, 2020

Abstract

This thesis is a theoretical work that explores approaches for conceptualising and interpreting musical performance assessment practices in higher education contexts. Musical performance assessment is characteristic of higher music education, and the principal form of assessment associated with tertiary musical performance training. Although educational assessment is a ubiquitous feature of higher education around the world, it has remained a contested realm of practice for decades, reflecting competing purposes, approaches, and philosophies. Despite much theoretical advancement having been made over recent decades, assessment practices themselves have been slow to change. This is especially true in higher music education contexts where traditional approaches continue to be valorised. This situation has prompted calls for new ways of viewing assessment, which is what this research is about.

This thesis describes an explorative, qualitative research project that sought to investigate meaningful ways of conceptualising musical performance assessment practices. Three main questions guided the research, each reflecting a different facet of assessment practices.

1. How can bases for achievement in musical performance work be meaningfully conceptualised?
2. How can bases of legitimation for assessment participants be meaningfully conceptualised?
3. How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised?

The emphasis on legitimation in these questions reflects the theoretical positions that frame the research, including the adoption of Legitimation Code Theory (LCT), which is a sociological framework for studying social practices. The questions are underpinned by a basic view of assessment as a socially-situated practice, characterised by multiple overlapping systems of legitimation. In exploring these questions, the aim was to develop useful conceptual resources for viewing assessment practices by explicitly connecting theoretical concepts with data. In this, the intention was to produce

theoretical ideas that could actually relate to musical performance assessment practices.

To anchor this process, data were purposefully collected from a range of sources, including the academic literature, interviews with assessment participants (n=26) including both assessors and students involved in Australian higher music education, and the body of publicly available documentation that governs assessment practices in this sector. The data were analysed using a multi-strategy qualitative methodology, and the exploration of intersections between the data and concepts from LCT was the main approach by which the research questions were explored.

The research has several main outcomes. The conceptual framework developed in the first part of the thesis provides a package of perspectives and resources for studying and theorising musical performance assessment practices. It situates musical performance assessment in higher education as an object of study and outlines a set of theoretical ideas than can be used to make sense of assessment practices. The latter part of the thesis develops responses to the research questions and offers theoretical resources for conceptualising various aspects of musical performance assessment practices. The main focus is on translating between concepts from LCT and aspects of substantive musical performance assessment practices. As part of this process, the later chapters of the thesis describe theoretical approaches for 1) conceptualising underlying characteristics of criteria for musical performance assessment, 2) for conceptualising how the attributes of assessment participants matter in terms of their ability to successfully participate in practices, and 3) for conceptualising and interpreting the legitimation of assessment practices themselves.

Ultimately, these outcomes are starting points. Although the work described in this thesis is deliberately oriented toward the world of practice, it is nonetheless primarily theoretical, and so its main contribution is to offer points of departure for further research and/or the development of approaches for use in substantive contexts.

Table of Contents

List of Figures

List of Tables

Certification

I hereby certify this work is original and that this work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Signed: _____

Date: _____14/02/2020_____

Acknowledgements

I cannot thank enough the supportive network of friends, family, and colleagues who have given selflessly of their time, energy, insight, and patience during my doctoral journey. My discovery of the world of research is almost entirely credible to Andrea Walton, who inspired me to undertake my first proper research training in 2015, and to Louise Denson who guided me through my very first research project. It was through Louise that I met Don Lebler, who in turn introduced me to the world of assessment. Don has given selflessly of his time, wisdom, and enthusiasm, and he has been an inspiration throughout the doctorate. Likewise, Gemma Carey has given generously of her time and expertise, balancing an at times overwhelming professional workload to provide timely guidance. As supervisors, Don and Gemma have been consistently encouraging and I am extremely grateful for their mentorship. Others in the Queensland Conservatorium and wider Brisbane research community have also provided valuable professional and personal guidance and support throughout the doctorate, and I would particularly like to thank Natalie and James Lewandowski-Cox, Brydie-Leigh Bartleet, and Paul Draper. Further abroad, I have benefited greatly from the insight and infectious enthusiasm of friends in the music assessment and Legitimation Code Theory communities—their generous contributions of time and expertise have been immensely valuable.

This project involved many personal ups and downs, and I could not have asked for a more supportive network of close friends and loved ones. Special thanks go to the Nussey family, whose support has been constant and unconditional, and with whom I am always made to feel at home. Special thanks go in particular to Daisy, who has been a constant companion throughout the doctoral journey and a pillar of support, especially in the more challenging periods of the journey. Finally, I would like to thank my family, whose support and enthusiasm has been unwavering. Thanks go to my sisters, Lucy and Mia, who have kept me grounded, and who set constant examples through their decency and zest for life. Most importantly, I thank my parents Neville and Katrina, without whose continued support both in work and general life this research would not have been possible.

Chapter 1: Introduction to the Thesis

1.1 Introduction

This thesis relates to assessment practices in the context of higher music education. It is specifically concerned with practices for formally assessing musical performance work in higher education settings, and the means by which these practices are conceptualised and understood. The research was undertaken in Australia, however the concepts and ideas developed in the thesis are not locked to this context, and are framed in such a way that their applications to other contexts remain accessible. The purpose of this short chapter is to introduce the basic elements of the thesis across four parts. The first part (Section 2.2) provides context for the thesis, and describes the basic problem-situation with which the thesis is concerned. The second part (Section 2.3) outlines the main aims of the research and the basic approach taken in the work, including the questions that were formulated to focus the thesis. The third part (Section 2.4) explains the significance of the research, and the final section (Section 2.5) provides an overview of the remainder of the thesis.

1.2 The Problem-situation

This thesis is a response to a particular issue that has arisen at an intersection of fields, including music education, higher education, and educational assessment. At the broad level, the issue is that the inner workings and abstract organising principles underlying assessment practices are complex, and difficult to make visible in meaningful ways. One of the main reasons for which this issue warrants attention is that assessment continues to exert a strong influence on the landscape of higher education that grows with the sector (Torrance, 2017). Preceding the research questions introduced in the next section is a basic conundrum: How can assessment practices be meaningfully understood? In this, the notion of *meaningful* interpretation of practice underscores the idea that the array of interpretive approaches on offer are not equally equipped in terms of their conceptual power. Despite the broad literature on educational assessment, the topic remains difficult to pin down. The scholarly landscape is characterised by a multitude of concepts (formative

assessment, summative assessment, assessment for learning, assessment as learning, authentic assessment, assessment literacy, measurement, evaluation, and so forth) which are inconsistently defined and understood, and which are valued in different ways by different groups. As Brown and Knight (1994) observe, "[a]ssessment may be many things for many people…Not only are there different purposes for assessment, but these purposes are neither separate nor entirely compatible" (p. 13). The breadth of the assessment research agenda has, problematically, led to a situation where "[t]he literature on assessment often looks more like the output of a cottage industry than the product of coordinated scholarship" (Knight & Yorke, 2003, p. 209). Scholars do, however, appear to be in relative agreement on one point which is that

> [t]he dominance of assessment in the student experience and the social, economic and policy climate have led to a situation where assessment is in a state of flux, facing pressures for enhancement while simultaneously coping with demands to restrict its burden on students and staff. (Bloxham & Boyd, 2007, p. 6)

Almost ten years after Bloxham & Boyd (2007), Boud et al. (2016) concluded that "[d]espite important advances in assessment theory and practice over the past three decades, it is apparent that change does not come easily" (p. 1107). Reflecting on a national Australian project in which they studied decision-making in the design of assessment, Boud et al. (2016) pinpoint a need for new ways of viewing assessment as an object of research study:

> [D]ecision-making about assessment is far from a simple process of rational choice between alternatives… It does not appear to be a simple matter of an inability to find out what can be done or simply a lack of resources. For us it points to limitations in the ways in which research on assessment is conceived and framed and the inadequacies of conventional assessment discourses to provide a focus for effective change. (p. 1108)

Others, such as Shay (2008a) have levied similar propositions, signalling the need for a broader interest in some of the basic issues for assessment practices, including:

- What is the field of educational assessment making claims about?
- When we assess, what are we assessing?
- How do philosophical paradigms influence what we think we are assessing, and how we go about assessing?
- Are we assessing knowledge, knowers, or knowing?
- What might be the relationships between these three possible 'objects'? (list adapted from Shay, 2008a, p. 599)

At the general level, this thesis is a response to the program of theoretical inquiry set out by Boud et al. (2016) and Shay (2008a), and the approaches taken in the research explicitly foreground a conceptual view of assessment practice. This work takes as its more particular focus, however, the very specific subject of assessment practices in higher music education contexts. Even more specifically, it is about the assessment of students' musical performance work, which is "characteristic of higher music education" (Lebler, 2015, p. 2). This is a complex area of practice, in which the competing pressures of increasing regulatory accountability and adherence to tradition are at a counterpoint with contrasting ideological positions on the assessment of creative work (Lebler, Carey, & Harrison, 2015a; Wrigley, 2005). As Burnard and Fautley (2015) tell us, creative works are "arguably the most difficult to assess" (p. 255), in that even simple tasks provide considerable latitude for expression. At the same time, institutions such as conservatoires have been painted as somewhat mysterious occupants of the higher education landscape. They have been described as secret gardens (Burwell, Carey, & Bennett, 2019; Hyry-Beihammer, 2010; Perkins, 2013), and they have been seen to place a great deal of emphasis on traditional Western approaches to musical performance training (Carey & Grant, 2015; Parkes, 2010). Echoing the sentiments of Boud et al. (2016), Partti, Westerlund, and Lebler (2015) tell us that more recent theoretical development in assessment "has not yet had a strong impact on higher music education institutions where the established

practices arise from the traditional 'conservatory model' of assessing with teacher panels" (p. 477). At the same time, music performance educators describe professional frustrations with assessment practices—in the words of one research participant:

> Nobody likes it. I don't think anybody likes it. No, I don't think anybody likes the way we do it. I don't know if anybody has an answer for how it should be done. I certainly don't.

This thesis is not, however, about new or better ways of doing assessment. As Broadfoot (2012) tells us, "even a cursory glance at the literature on assessment reveals the predominance concern about techniques" (p. 17). Similarly, Knight and Yorke (2003) describe a profusion of

> reports of experiences of using this technique in that module, or of the levels of reliability that can be attained by being scrupulous in the use of a certain method for certain purposes, or of experiences which made teacher and students feel that good learning had taken place. This is important but it is not sufficient unless we are prepared to see the undergraduate years as a farrago of unrelated modules taught without reference to systematic appreciations of the research literature on human learning in general and undergraduate learning in particular. (Knight & Yorke, 2003, p. 209)

The focus in this thesis is, instead, on the meaningful interpretation of assessment practices, and in particular, reifying some of the underlying principles that govern them. The intention is to make visible some of the socially-situated *rules of the game* (Bourdieu, 1998), or what Maton (2014) refers to as *organising principles,* that influence the shapes and outcomes of practices. This thesis is about exploring means of making these rules and principles more explicit in the higher music education context. In this, it is a response to the broader theoretical directions proposed by scholars such as Boud et al. (2016) and Shay (2008a), joined with an interest in the specialised disciplinary practice of musical performance assessment.

1.3 Research Aims and Approach

The main aim of the research was to explore means of meaningfully conceptualising, interpreting, and theorising musical performance assessment practices in higher music education. A qualitative methodology was adopted in view of the explorative nature of the project (Creswell, 2007; Maxwell, 2009, 2013; Robson, 2011; Silverman, 2011). The basic approach of the research follows Maxwell's (2009, 2013) suggestion of an interactive model of research design in which emphasis is placed upon dialogue between the different parts of the design (Figure 1.1). A brief description of each component is provided below.
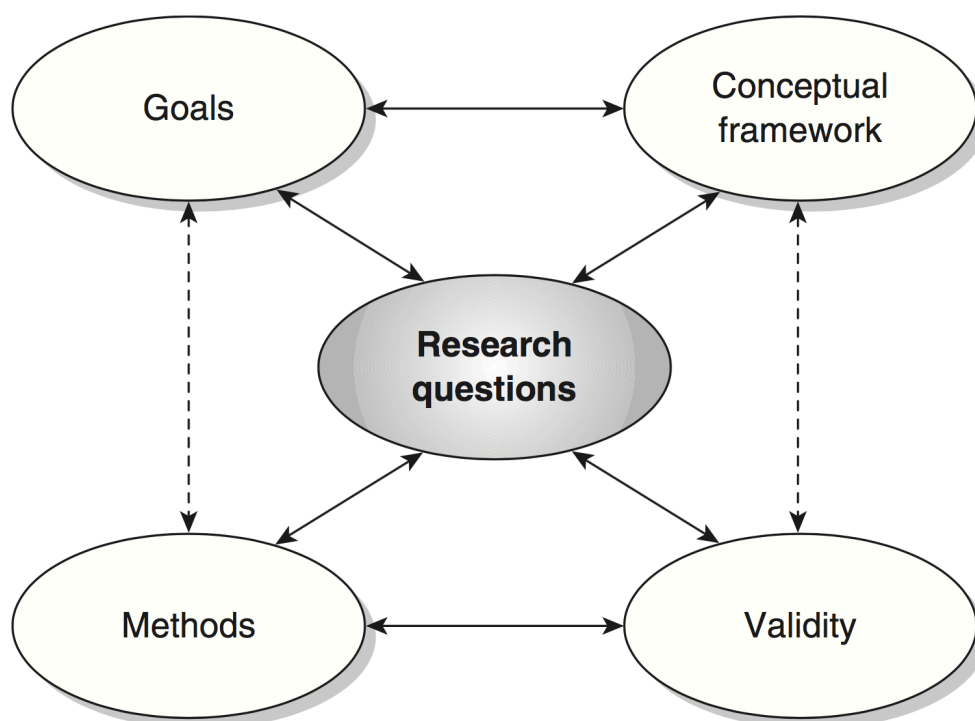
Figure 1.1 Maxwell's model of interactive research design (Maxwell, 2013, p. 5), reprinted with permission

1.3.1 Goals

The main goal of the study was to explore means of meaningfully conceptualising assessment practices in higher music education, focusing on the assessment of musical performance. Research questions were formulated to focus this goal and provide a manageable scope for the project. Given the conceptual emphasis of the project, a theoretical framework was sought that would offer useful explanatory potential, and ideally, enable insights that could be put to practical use.

1.3.2 Conceptual framework

The broader conceptual framework that informs this research includes three general components. Firstly, the field of research relating to educational assessment and music education provides the research context for the study. Secondly, this research adopts a realist qualitative research approach (Maxwell, 2012, 2013; Robson, 2011) that conceives of practices and knowledge as socially-situated and produced (Maton & Moore, 2016b). Lastly, and following the interest of this study in a social practice approach, a more specific theoretical framework was sought that would provide useful concepts for exploring the organising principles of assessment practices. To this end, Legitimation Code Theory (LCT) was adopted as the main theoretical framework for the study after a period of substantial review and scoping in the first year of the project. LCT is an interpretive theoretical framework for studying social practices (Maton, 2014), and in this thesis it provides the main set of conceptual tools that are used in theorising musical performance assessment practices. It is a generic framework, in the sense that the concepts it offers are not tethered to a disciplinary context, enabling a range of data to be conceptualised within the same theoretical space. The different aspects of the conceptual framework remained in dialogue with one another throughout the project, and the ultimate trajectory of the research reflects the combined influence of all elements of the framework.

1.3.3 Research questions

To focus the scope of the project, three questions were derived to guide the research. These questions are more targeted responses to the broader problem-situation in that they reflect more specific substantive issues for assessment in higher music education while retaining latitude for the broader theoretical motivations of the work. The three questions formulated to guide the research are:

1. How can bases for achievement in musical performance work be meaningfully conceptualised?
2. How can bases of legitimation for assessment participants be meaningfully conceptualised?

3. How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised?

The first question was formulated to place the focus specifically on the object with which assessment is concerned, which in this study is students' musical performance works themselves. It reflects an interest in the means by which achievement[1] can be attained (or limited) in musical performance assessment events. This is a central issue in musical performance assessment where appraisals are complex subjective judgements, and approaches for dealing with this issue are contested (Sadler, 2009a, 2009b; Wrigley, 2005).

The second question shifts the focus onto assessment participants, including both staff and students who are collectively viewed as practitioners of assessment. This question places the emphasis on the legitimation of those people involved in assessment, rather than the product on which assessment events are focused. In this it provides a point of departure for exploring how *who you are* is important in musical performance assessment.

The third question places the emphasis on assessment practices themselves. It provides a point of departure for exploring why some musical performance assessment practices and approaches are construed as more or less legitimate than others.

1.3.4 Methods

To explore the research questions, a research methodology was formulated with the aim of operationalising concepts from LCT to conceptualise the foci of the questions. The role of data in this research was, therefore, to provide an empirically-observable context in which to anchor conceptual theory-building. Following standard qualitative research practice, a range of data were sought to enable triangulation and enhance the interpretability of the findings (Robson, 2011). Data were sought from three main sources, including interviews with a diverse range of assessment practitioners, institutional documents such as

---

[1] In this thesis the term *achievement* is used both in specific reference to student achievement denoted by the awarding of grades, as well as in the general sense of reflecting a successful outcome in a particular situation. The concept is discussed in greater detail later in the thesis (see Section 6.2.1).

authoritative course outlines and legislative documents that govern the Australian higher education sector, and the academic literature on assessment itself, which were collected and reviewed throughout the project. The data were purposefully selected (Robson, 2011) with the aims of the research in mind. The analysis of data was conducted using an integrated approach inspired by Maxwell (2013) and Maton and Chen (2016). It involved bringing together concepts from LCT in combination with a standard thematic analysis approach (after Braun & Clark, 2006; Maxwell, 2013) to make sense of the data and to synthesise responses to the research questions.

1.3.5 Validity

Since this research is explorative, interpretive, and qualitative in nature, it does not express validity in the statistical sense. After Robson (2011) and Maxwell (2013), validity in the context of this work is conceived of as more closely related to the robustness of interpretations developed in the research. To this end, both Robson and Maxwell suggest a risk-management approach that involves dealing with threats to validity, which include threats related to description, interpretation, theory, bias, and reactivity. These threats are discussed more fully in Chapter 5.

1.4 Significance of the Research

The basic motivation for this research is the centrality of assessment in the student experience, given that students cannot—"by definition if they want to graduate" (Boud, 1995, p. 35)—escape the effects of assessment practices. At the broader level, this work is a response to the identified need to look at assessment in alternative ways (Boud et al., 2016; Shay, 2008a), and in particular, to develop some meaningful conceptual approaches that can be used to make sense of and theorise practices to productive ends. In this, the research contributes to the landscape of assessment scholarship by developing a range of theoretical ideas that provide fresh means of interpreting assessment practices. It is a contribution to the body of explanatory theory available for assessment researchers and practitioners, and to the development of a language of description for making sense of practice that places an emphasis on meaningful conceptual interpretation in which

theoretical ideas and practical concepts are explicitly linked. The tools developed within the thesis are, therefore, proposed with a view to enabling assessment practitioners to meaningfully interpret, engage with, and effect assessment practices.

The focus on musical performance assessment more specifically is a contribution to the complex and contested disciplinary discourse about assessment that is central to higher music education learning and teaching practice. In this it is a response to the pressures facing an educational sector sandwiched between tradition on the one hand (Parkes, 2010; Partti et al., 2015), and a rapidly changing, highly regulated, globalised higher education sector on the other (Allais, 2012b, Bloxham & Boyd, 2007; Torrance, 2017; Vaira, 2004). The focus on conceptual approaches means that the outcomes of the research are not locked to context, and are more accessible to a broader range of educational situations. In this, the work speaks to the broader discourse around educational assessment, and may be applicable in other educational sectors, particularly where the emphasis is on creative work.

1.5 Structure of the Thesis

Following this introductory chapter, the remainder of the thesis is developed over the course of a further eight chapters. Chapters 2–4 begin by describing the conceptual framework of the project. Chapter 2 introduces the conceptual framework and explains its basic features, including the main elements from the framework that contextualise the research at the broader level. Chapter 3 focuses on the subject of assessment in detail. Its main role is to sharpen the context of the study by discussing the main issues that give rise to the research questions. Chapter 4 completes the conceptual framework. It explains the theoretical framework adopted in the thesis and details the main suite of concepts from LCT that are enacted in the research. Chapter 5 explains the methodology of the research, including the means by which concepts from LCT were integrated into the study. Chapters 6, 7, and 8 each offers a response to the three research questions. Chapter 6 explores a means of theorising bases of achievement for the assessment of students' musical performance work. Chapter 7 extends the conceptual territory developed in Chapter 6 to the assessment participants themselves, exploring and

developing a conceptual frame for the legitimation of those involved in assessment. Chapter 8 explores the legitimacy of assessment practices, and discusses some of the bases that underlie their legitimation. Chapter 9 concludes the thesis, and offers a summary of the main points developed within the thesis, as well as a discussion of the contributions, limitations, and points of departure for further study.

Chapter 2: Conceptual Framework (Part 1)

Positioning the Research

2.1 Introduction

This thesis is about interpreting assessment practices in the context of higher music performance education. The previous chapter introduced the research, and outlined the general background to the thesis as well as the research problem, research aims, and overall approach. This is the first of three chapters devoted to the *conceptual framework* of the thesis. Briefly introduced in Chapter 1, the conceptual framework is a central aspect of the research design that sits in close dialogue with the aims and guiding questions of the research. The approach followed here has been strongly influenced by Maxwell (2006, 2009, 2013), who defines a conceptual framework as "the system of concepts, assumptions, expectations, beliefs, and theories that supports your research" (2009, p. 222). An important feature of this approach is that it positions these elements as integrated parts of the research itself. In this, the conceptual framework is explicitly distinct from a *literature review* in the traditional sense*,* and includes the full range of conceptual resources that inform the research (Maxwell, 2006, 2009). Since the components of the conceptual framework are closely intertwined, the segmentation of the framework across these three chapters is unavoidably artificial. Notwithstanding this segmentation, I have attempted to present the framework in a logical order to be as clear as possible about the relationships of each part to the project.

The focus in this chapter is on the main contextual positions that inform the thesis. Here, a narrower focus on assessment is temporarily delayed in favour of clarifying the general territory of the study, since this has a meaningful relationship to the views of assessment developed later in the thesis. This chapter includes some of the main theoretical perspectives that inform the research, as well as some explanation of the substantive context of the work. Chapter 3 focuses on the research questions that guide the study and is more specifically about assessment itself, while Chapter 4 describes the

explanatory framework used in the thesis and completes the conceptual framework.

The first part of this chapter (Section 2.2) provides a general overview of the conceptual framework, including a more detailed explanation of the broader approach to formulating the framework. Following this, the remaining sections describe the contextual positions adopted in the work. Section 2.3 outlines the main interpretation of *theory* sustained in the thesis, and describes some of the broad theoretical positions that overlay the research. Section 2.4 is concerned with the issue of conceptualising the field of assessment in higher music education. It describes a simple means of understanding the contextual relatedness of constituents from the field (including substantive research studies, artefacts, people, and so forth) to the context of musical performance assessment in higher education that was influential in this project. The final part of the chapter (Section 2.5) develops an interpretation of *higher music education*, and locates musical performance work and assessment in this setting.

2.2 Overview of the Conceptual Framework

The conceptual framework is an important aspect of this research for several main reasons. Firstly, it clarifies the positions that anchor the research. This is useful in the sense that it makes explicit the kinds of claims upon which the research is predicated, which in themselves are intended as a resource for future work. To this end, the conceptual framework can be thought of as an outcome of the research, as well as a record of the main points of departure which developed through the course of the research. Secondly, the conceptual framework supports the validity of the study by making the positions that inform the research (and to which it is accountable) explicit. Thirdly, it explicitly positions the components of the conceptual framework as integrated parts of the research. Maxwell (2009), after Becker (1986), suggests that conceptual frameworks can be thought of as comprised of a variety of *modules*, collected and constructed by the researcher. The modules that make up a conceptual framework do not come from a prescribed list, being responsive to the requirements of the research: they may include, for example, insights from the research literature in a particular area, philosophical paradigms

(constructivism, positivism, realism, etc.), theories and theoretical frameworks, and personal experiential knowledge. These modules serve a range of functions over the course of a research project, including the marking out of particular contexts and standpoints, as well as providing tools and data for the research itself (Maxwell, 2009). The focus in this suite of chapters (Chapters 2–4) is on the specific modules that make up the framework. Underpinning the selection and synthesis of these modules are four general perspectives which are briefly described here.

2.2.1 Research questions

The research questions developed to guide the study serve as anchor points for the conceptual framework in the sense that the various aspects of the framework are closely connected to their interpretation and exploration. The strategy followed in this project reflects the proposition that explorative qualitative research should be guided by a *reflexive* approach to research questions (Maxwell, 2009, 2013; Robson, 2011). In this study the research questions were not viewed as absolute starting points, but rather as integrated, interactive components that both directed the research and were at the same time responsive to insights that emerged through the research process. To quote Maxwell (2009),

> [t]his is different from seeing research questions as the *starting point* or primary determinant of the design. Models of design that place the formulation of research questions at the beginning of the design process, and that see these questions as determining the other aspects of the design, don't do justice to the interactive and inductive nature of qualitative research. The research questions in a qualitative study should not be formulated in detail until the goals and conceptual framework (and sometimes general aspects of the sampling and data collection) of the design are clarified, and should remain sensitive and adaptable to the implications of other parts of the design…. However, these specific questions are generally the *result* of an interactive

design process, rather than the starting point for that process. (p. 229, italics original)

2.2.2 Literature review

The role of literature review in this work is shaped by two important distinctions. Firstly, the approach taken here reflects the view that conceptual frameworks are distinct from literature reviews in that they can include a broader range of resources, including the experiences of the researcher, unpublished work, insights produced in earlier stages of a project, and theoretical tools (Maxwell, 2009). Secondly, this study explicitly recognises the difference between "literature review articles for publication (reviews *of* research)... [and] dissertation literature reviews, which are primarily reviews *for,* rather than of, research" (Maxwell, 2006, p. 28, italics original). Maxwell (2006, 2009) and others (Becker, 2007; Locke, Spirduso & Silverman, 1999) highlight several key issues with traditional notions of dissertation literature reviews. Maxwell (2009) tells us that it can be a "dangerously misleading term, for three reasons" (p. 223), which are that:

- It can lead to a narrow focus on literature at the expense of other useful conceptual resources that "may be of equal or greater importance... including unpublished work, communication with other researchers, and your own experience and pilot studies" (p. 223)
- It can promote strategies of "'covering the field' rather than focusing specifically on those studies and theories that are particularly *relevant*" (p. 223, italics original) to the research
- It can convey an emphasis on description, which Maxwell (2009) asserts is only part of the purpose of the conceptual framework, and that the literature should be treated "not as an *authority* to be deferred to, but as a useful but fallible source of *ideas* about what's going on, and to see alternative ways of framing the issues" (p. 223, italics original)

Accordingly, this thesis does not include *a* (singular) literature review in the sense of a literature review for independent publication. Rather, the emphasis here is on literature review as an ongoing process integrated into the research,

which is a truer representation of how the research actually developed over the course of the project. This is important for the reason that it directly responds to the issues highlighted by Maxwell (2009) and is an approach which helped to prevent the ossification of the critical research process.

2.2.3 Research paradigm

Maxwell (2009, 2013) suggests that one of the most important aspects in the design of research studies is connection with a *research paradigm* (or paradigms, see also Robson, 2011). In this setting, the term refers to "a set of very general philosophical assumptions about the nature of the world (ontology) and how we can understand it (epistemology)" (Maxwell, 2009, p. 224). Research paradigms of the most general kind include philosophical positions such as constructivism, positivism, and realism, while more specific examples include phenomenology, post-modernism, critical theory, feminism, and interpretivism (Maxwell, 2009, Robson, 2011). The choice of paradigm is important for the reason that it articulates high-level principles according to which claims are made about the validity of a study, and the general choice of approaches used within it (Maxwell; 2009, Robson, 2011). Research paradigms are intimately connected with the world of *theory*, and in this framework they are understood as *meta-theories* (Maton, 2014) which are distinct from other kinds of theories that are of interest in the thesis (see Section 2.3).

2.2.4 Theoretical modules

Related to the previous two aspects of the conceptual framework are the more specific theoretical modules used in the study. The kinds of theory encapsulated in these modules include theoretical ideas generated through substantive studies, as well as explanatory theoretical concepts. In this thesis, these are modular in the sense that placing them in different combinations can produce different kinds of outcomes. The kinds of theoretical distinctions drawn in the thesis are described in greater detail in Section 2.3.

2.2.5 Experiential knowledge

Maxwell (2009) notes that "[t]raditionally, what you bring from your research background and identity has been treated as 'bias,' [*sic*] something whose influence needs to be *eliminated* from the design, rather than a valuable component of it" (pp. 224–225). More recently, however, *experiential data* (Strauss, 1987) has been recognised as a "major source of insights, hypotheses, and validity checks" (Maxwell, 2009, p. 225, after Berg & Smith, 1988; Denzin & Lincoln, 2000; Jansen & Peshkin, 1992; Strauss, 1987). As Maxwell (2009) emphasises, this perspective is not "a license to impose your assumptions and values uncritically on the research" (p. 225) but rather reflects Reason's (1988) *critical subjectivity*, which refers to

> a quality of awareness in which we do not suppress our primary experience; nor do we allow ourselves to be swept away and overwhelmed by it; rather we raise it to consciousness and use it as part of the inquiry process. (p. 12, as cited in Maxwell, 2009, p. 225)

To this end, a researcher identity memo (Maxwell, 2013) was developed to describe and clarify my own personal position in relation to the project, including the experiences and insights that have informed the work, as well as those that have been derived from it over time. This memo is included in Appendix A.

2.3 Theoretical Position of the Thesis

The purpose of this section is to outline, in broad strokes, the general theoretical ideas that guide the research. The reason for introducing these concepts at this early stage is that they explain the basic set of theoretical propositions according to which the various constituents of the research are framed. Theory is central to this thesis in the sense that it is about making sense of—that is, theorising—assessment issues.

Although a good deal of the scholarship on educational assessment deals with concepts that are theoretical in nature, conceptions of theory and/or the theoretical qualities expressed by the ideas proposed and developed within

the field are often unclearly demarcated. This is important, for the reason that concepts are not made equal—there is variation in what they do and how well they are equipped to work in given circumstances (Maton, 2014). Part of the complexity around this is that the term theory can have a range of interpretations—a "polysemy" (Maton, 2014, p. 14) that, Maton (2014) suggests "results partly from failing to distinguish theories from paradigms" (p. 14, after Boudon, 1980). Two sets of distinctions are useful in alleviating this issue, and help to clarify what comes later in the conceptual framework. The first set is offered by Maton (2014), who develops ideas from Archer (1995) to distinguish three broad kinds of theory, which are

- *meta-theories* that reflect ontology and epistemology
- *theories* comprising explanatory frameworks
- *substantive theories* of research studies

These forms of theory can be simply explained as follows. Beginning at the finer-grained level, substantive theories offered by research studies provide particular accounts, explanations, or conjectures on the basis of research. At the broader level, meta-theories develop and promote ideas about the natures of knowledge and being—some examples of meta-theories are positivism, constructivism, and realism. In between these two kinds of theory, explanatory theoretical frameworks offer means of generating explanations. Although they may be enacted in substantive research studies, and although they are informed by (and inform) meta-theoretical ideas, explanatory frameworks are neither specific accounts of research nor are they positions on knowledge and being. Rather, they sit at an interface with these other kinds of theory (Maton, 2014).

The second useful distinction is one proposed by Bernstein (2000), who conceives of a "discursive gap" (p. 445) between the worlds of theory and empirical data (Maton & Chen, 2016; McLean, Abbas, & Ashwin, 2013). Maton and Chen (2016), after Bernstein (2000), tell us that "a key source of this problem [the discursive gap] is the form taken by theories themselves" (Maton & Chen, 2016, p. 27). To this end, Bernstein (2000) proposes a distinction between *internal* and *external* languages of description to differentiate

respectively between "the languages of theory or concepts... [and those] which are rooted in the empirical world" (McLean, Abbas, & Ashwin, 2013, p. 265). Citing Bernstein (2000), Maton and Chen (2016) describe the distinction as "between the 'internal language of description' of a theory or how the constitutive concepts are interrelated, and its 'external language of description' or how those concepts relate to referents beyond the theory" (p. 28, see Bernstein, 2000, p. 132). An important limitation of internal languages of description is that they exhibit "powerful and persuasive internal conceptual language but reduced powers to provide externally unambiguous descriptions of the phenomena of concern" (Bernstein, 2000, p. 168). The potency of an internal language of description is "greater where concepts are tightly interrelated and weaker where they are less related" (Maton & Chen, 2016, p. 27), and languages of this kind "should be capable of going beyond the descriptions created by members [of a given context]" (Bernstein, 2000, p. 135). External languages of description offer a means for connecting the world of theory with the world of data, and the strength of external languages reflects their capability to relate "concepts and referents… in relatively unambiguous ways" (Maton & Chen, 2016, p. 27). The business of *theory-building*, to follow Bernstein's (2000) thinking, involves traversing the discursive gap between theory and the world of data (Maton & Chen, 2016; McLean, Abbas, & Ashwin, 2013), and the development of external languages of description is an approach for making this traversal more explicit. The connection of theory with the world of musical performance assessment is a broad aim of this thesis, and some more specific means of translating between the internal languages of theory and referents from this world (data) are explained in Chapter 5.

These sets of distinctions help to explain the modules of the conceptual framework that are described in this and the chapters that follow. The sections that immediately follow are concerned with the meta-theoretical position of the thesis and complete the basic theoretical overlay. Chapter 3 is more directly concerned with substantive theoretical ideas and the internal language of assessment research, being focused on the landscape of research about assessment in higher education and the more specific body of work related to musical performance assessment. The focus in Chapter 4 is on Legitimation Code Theory, which is a specific explanatory theoretical framework for

studying social practices, and which is one of the main bases for the discussions developed in the second half of the thesis.

2.3.1 A realist philosophical perspective

This section describes the basic philosophical position of the thesis, which is most closely aligned with the realist paradigm. A range of texts from the literature on research methodology offer useful accounts of the various philosophical paradigms that have been influential in social research, which include constructivism, positivism, and realism amongst others (see for example, Maxwell, 2013; Merriam, 2009; Robson, 2011; Silverman, 2011). The view taken in this research aligns most closely with positions from critical realism (Maxwell, 2012, 2013; Robson, 2011), and social realism (Maton, 2014; Maton & Moore, 2010b). It is a perspective that reflects the combination of

> a realist ontology (the belief that there is a real world that exists independently of our beliefs and constructions) with a constructivist epistemology (the belief that our knowledge of this world is inevitably our own construction, created from a specific vantage point, and that there is no possibility of achieving a purely 'objective' account that is independent of all particular perspectives). (Maxwell, 2012, p. vii)

The basic premise, as such, is that "[a]ll knowledge is… 'theory-laden,' [*sic*] but this does not contradict the experience of a real world to which this knowledge refers" (Maxwell, 2012, p. vii). This view emphasises the rejection of a dichotomous view of positivist and social constructionist perspectives (Maxwell, 2012, 2013; see also Maton, 2014) on the basis that it constitutes

> a false dichotomy between, on the one hand, the belief that knowledge must be decontextualized, value-free, detached and 'objective' and, on the other hand, the idea that knowledge is socially constructed within particular cultural and historical conditions (and necessarily entwined with issues of interest and power). (Maton & Moore, 2010a, pp. 1–2)

Rather, the realist philosophy upon which this work is predicated asserts that both perspectives are compatible (Maxwell, 2012, 2013) and involves a view of knowledge as "both constructed and based on the reality of the world we experience and live in" (Robson, 2011, p. 28). Sayer (2000) explains that it is

> the evident *fallibility* of our knowledge—the experience of getting things wrong, of having our expectations confounded, and of crashing into things—that justifies us in believing that the world exists regardless of what we happen to think about it. (p. 2, italics original)

The social realism paradigm adopts general principles reflecting all of these ideas, but places additional emphasis on the social circumstances of knowledge while retaining the idea that it can possess qualities of its own—it is a view of knowledge as "not only social but also real (hence its name) in the sense of possessing properties, powers and tendencies that have effects" (Maton, 2014, pp. 9–10). This philosophical view underlies the emphasis in this thesis on assessment as a *social* practice, and speaks to the positioning of this work as a response to calls for such approaches (see for example Boud et al., 2016; Shay, 2008a, 2008b). The next section describes a social practice perspective on assessment in greater detail.

2.3.2 Seeing assessment as a social practice

Conceptualising assessment in any context is complicated for the reason that assessment is, broadly speaking, inherent in daily life: As Broadfoot (2012) observes, "[i]t is very hard to set limits to assessment. Passing judgment is a central part of social behaviour" (p. 12). This is especially true in musical fields, where "[q]uality judgements form a routine part of listening" (Thompson & William, 2003, p. 21). An interest in assessment is therefore, in the general sense, an interest in a field of practice that operates at many levels in day-to-day life and society. Speaking to the broader field of education research, Schatzki (2012) informs us that *practice theory* is an approach that has "gained currency over recent decades" (p. 13). The conceptual landscape associated with practice theory is diverse, however at the broad level, the basic idea is that practices are inherently social: they are "an organised constellation of

different people's activities" (p. 13). A helpful pair of articles by Shay (2008b) and Boud et al. (2016) summarise what taking a social practice perspective means in educational assessment research contexts. Shay (2008b) tells us that it is a view which places an emphasis on peoples' activities as "located in space and time, and [these activities] cannot therefore be understood outside of the social, cultural, economic and political contexts in which they occur" (p. 160). Practices are also seen as purposeful, in that: "[T]he notion of 'practice'… suggests that these doings are habitual; they have hardened into a kind of relative permanency… This relative permanency means that people engage in practices purposefully, but not always consciously so" (Shay, 2008b, p. 160). Building on these basic ideas, Shay (2008b) outlines a research position where

> the analytical focus is not on individual agents, though the role of agents within practice is crucial. Nor is it on structures, though the macro social conditions which enable practice are equally crucial. A social practice perspective is focused on the relationships between the acts of agents within practices (including their understandings of these acts) and the broader social conditions which constitute these practices, with a particular interest in their co-determination. Bourdieu refers to this as 'thinking relationally' (Bourdieu & Wacquant, 1992, p. 228). Ultimately the interest lies in the *underlying principles* which explain or make sense of practice in a particular social, historical, institutional context. The emphasis on the particularity of practice is not antithetical to a search for the more general or the universal. The goal of such research is 'to grasp the particularity within the generality and the generality within the particularity' (Bourdieu & Wacquant, 1992, p. 75) and to uncover 'the universal buried deep within the most particular' (Bourdieu & Wacquant, 1992, p. 44). (Shay, 2008b, p. 160, italics added)

A summary of this thinking is that the field of assessment can be construed as one of *discursive* social practice, in the sense that it is a field in which a range of people engage, in different ways, and to different ends. At the substantive

level it is a view of assessment practices as complex social activities, which involve a range of purposes, locations in place and time, actors, material artefacts, and an associated "discourse of sayings and doings" (Boud et al., 2016, p. 1110), where "[d]ecisions about even apparently technical matters are influenced by political and ideological considerations" (Gipps, 1999, p. 355). Although recent decades have seen an increase in the adoption of social practice perspectives for assessment research (Boud et al., 2016; Bearman & Ajjawi, 2018; Broadfoot, 1996, 2012; Filer, 2000; Gipps, 1999; Shay, 2004, 2005, 2008a, 2008b), this approach appears not to have been deeply explored in studies of assessment practices in higher music education contexts. Within this broader conception of assessment practices are several more specific ideas that further sharpen the basic philosophical position of the research. These ideas directly anticipate the inclusion of the explanatory theoretical framework (Legitimation Code Theory) in the research (detailed in Chapter 4).

2.3.3 A view of social practices as sites of struggle over legitimacy

The landscape of literature on social theory and practice includes contributions from a range of scholars who have proposed principles, theories, and devices according to which the social world can be understood. This work is influenced most strongly by the view that social practices are characterised by discourses of legitimation which influence the ways in which individuals or groups are able to successfully participate in social fields of practice. This perspective is argued for by Maton (2014) who develops contributions from Bourdieu and Bernstein (amongst others). Both Bourdieu and Bernstein perceived social practices as sites of cooperation and struggle in which people—sometimes referred to as *actors*—seek to maximise outcomes that benefit them, and within which particular rules influence the ways in which different groups are able to successfully participate (or not) in practices (see for example, Bernstein, 2000; Bourdieu, 1998). Navigating fields of social practice is often complex in that *rules of the game*[2] are not by definition explicit, and challenging

---

[2] This phrase is used in a general sense here which reflects a common paraphrasing of Bourdieu who offered the analogy of a *game* as a tool for of making sense of social practices (see for example, Bourdieu, 1998; Maton, 2019; Grenfell, 2008). This is distinct from other usages of the term, for example, in the context of game theory or gamification.

these rules can be difficult (Maton, 2014, 2019). Building on Bourdieu and Bernstein, Maton (2014) conceptualises a *legitimation device* that influences the ways in which success is distributed in practices. For Maton, the legitimation device is a conceptual device that positions legitimacy as a defining feature of what enables success in practices. Maton (2014) exaplains:

> Underlying the structuring of fields, and acting as a kind of exchange rate mechanism among currencies, is the *Legitimation Device…* Whoever controls this 'device' establishes specific legitimation codes as dominant and so defines what is legitimate, shaping the social field of practice as a dynamic field of possibilities. To analyse legitimation codes is thus to explore what is possible for whom, when, where and how, and who is able to define these possibilities, when, where and how. (p. 18, italics original)

In this view, the actions of people and—the artefacts that they produce— "embody messages concerning what should be the dominant measurement of achievement within a field" which are "*languages of legitimation*" (Maton & Chen, in press, p. 2, italics original). For assessment, this view is a fairly natural fit—assessment is, after all, directly concerned with legitimation. At the broader level, Broadfoot (1996) argues that assessment "provide[s] one of the principal mechanisms by which the changing bases for social control within the broader society which gave birth to mass schooling in the first place are translated into the educational process" (p. 9). As Burke (2012) emphasises, "education is never neutral but is always a site of struggle over meaning-making and knowledge… Universities are significant institutional sites of the legitimization of certain forms of knowledge and identity" (pp. 35–36). For the purposes of this project, this perspective on practice signifies an interest in seeking to better understand what the rules of the game look like for assessment in higher music education contexts, and how they can be productively framed to make visible the kinds of cooperations and struggles that delimit successful participation in assessment practices.

2.3.4 An additional note on knowledge

Knowledge, in the general sense, is a recurring theme in this thesis. In addition to what is described above, the main theoretical ideas that frame knowledge are described in Chapter 4. One point to make clear here however, is that this thesis devotes little attention to distinctions between kinds of knowledge, in the sense of concepts like declarative knowledge, procedural knowledge, embodied knowledge, and so forth. These are important ideas, but not the focus of this work. Rather, the focus in later chapters is on theoretical relationships that frame and characterise knowledge in its various forms, as well as its counterpart *knowing,* which invokes an emphasis on the attributes of people—or *knowers*—including what knowledge and skills they possess (Maton, 2007, 2014).

2.4 Conceptualising the Territory

Having defined the basic theoretical positions of the research, this section concerns the problem of making sense of the conceptual field within which the research is situated. The issue here is that the research overlaps several regions of educational scholarship that are simultaneously distinguishable yet interrelated. Some of these fields include:

- The field of higher education
- The field of research about educational assessment
- The field of music education—in particular, musical performance training
- The field of research about assessment in higher music education

Although the focus of this research on musical performance assessment in higher education is fairly specific, it nonetheless exists at a particular intersection of these regions. This has implications for the ways in which assessment is viewed and contextualised within the conceptual framework. Highlighting the overlap is important because it makes explicit the combined influence of these regions. Many parts of assessment have been identified, explored, and theorised, and there is thus a good deal of research that frames the subject of musical performance assessment but that does not deal with it

explicitly. At the same time, prominent assessment concepts that have been taken up in research about musical performance assessment (for example, peer assessment) have a broader life of their own beyond the musical performance assessment literature. Further, these concepts may be interpreted in different ways in different parts of the world, and so their definitions cannot be taken for granted.

To provide a means of conceptualising the relatedness of research constituents to the musical performance assessment context, I pre-empt some of the theoretical tools described in Chapter 4. After Maton (2014, 2020), I propose that the constituents of the field in which this research is situated (including substantive and theoretical studies, educational material, social groups, discourses, and so forth) can be thought of in terms of *semantic gravity*, which "refers to the degree to which meaning relates to its context" (Maton, 2020, p. 62). Where concepts are characterised by stronger links to a particular context, they express stronger semantic gravity (SG+), while those that are less bound to context express weaker semantic gravity (SG-). This range is realised visually on a continuum of strength (Table 2.1).

Table 2.1 Using semantic gravity (Maton, 2014) to conceptualise contextual specificity

Continuum of relative strength of semantic gravity	Relative strengths and weaknesses of semantic gravity	Examples
SG- ↑ ↓ SG+	Weakest semantic gravity (SG--)	Assessment in higher education
	Weaker semantic gravity (SG-)	Assessment in higher education disciplines (e.g. music)
	Stronger semantic gravity (SG+)	Assessment of disciplinary activities in higher education (e.g. musical performance)
	Strongest semantic gravity (SG++)	Disciplinary assessment activities in particular settings

These ideas are expanded on further in Chapter 4, however for the purposes of this chapter they enable a simple distinction between more generic concepts and more discipline-specific concepts. Expressing the least gravity are the most context-independent ideas. A general discussion of the purposes of assessment, for example, expresses weaker semantic gravity than one that is linked to the specific context higher music education. Expressing the most gravity is the most context-dependent material: for example, studies of assessment practices and/or concepts in specific contexts, such as units of coursework. To give a more nuanced example, consider a group of three papers about the subject of holistic assessment (Iusca, 2014; Sadler, 2009a, 2015). In one of these papers, Sadler (2009a) discusses the concept of holistic assessment relatively independent of its application in a specific disciplinary context. Another paper by Sadler (2015) covers similar conceptual subject matter, but was published in a book about assessment in music education. Iusca (2014) compared segmented and holistic strategies in the assessment of recorded musical performances. Of these three, Sadler (2009) expresses the weakest semantic gravity, Iusca (2014) expresses the strongest semantic gravity, and Sadler (2015) is in between the two. What is important here is that semantic gravity is *not*, by definition, a barometer of the relevance of a particular topic to the research. It is useful precisely because it enables us to conceptualise one of the ways in which less context-specific subjects relate to the more specific disciplinary context that is the substantive focus of the thesis, retaining the emphasis on the relevance of the former. In this, it helps to rationalise why some parts of the conceptual framework are devoted to concepts independent of their immediate relevance to musical performance contexts. Several educational studies suggest that movement across the contextual continuum of semantic gravity is productively involved in the construction of meaning (see for example, Blackie, 2014; Clarence, 2016; Maton, 2020; Matruglio, Maton, & Martin, 2013). Drawing links between more general ideas and the specific context of musical performance assessment (as well as those from other contexts) is therefore a broader contribution of this study.

2.5 Defining the Higher Music Education Context

Having described the broader conceptual points of departure for the research, the final part of this chapter describes in detail the features of the higher music education context to which the thesis is oriented. Part of the complexity of the music education landscape is that the terminology *music education* can refer to a range of kinds of education (Fautley & Murphy, 2015), just as assessment can also have a number of meanings depending on the context (Sadler, 2005). This part of the conceptual framework develops the contextual framing for the focus of the thesis on assessment in musical performance programs in higher education. Section 2.5.1 considers the meaning of *higher education* and explains what this refers to in the Australian context where this research was conducted. Section 2.5.2 examines the concept of *higher music education* more closely to distinguish the kind of music education that this thesis most closely relates to. Section 2.5.3 contextualises musical performance as a central feature of higher music education curriculum, and Section 2.5.4 describes the basic kinds musical performance assessment found in higher education settings.

2.5.1 Higher education in Australia

In this thesis, higher education refers to the educational sector that offers legally accredited, non-compulsory programs of study, upon completion of which graduates are awarded one of a range of possible qualifications. These programs of study are mainly available to applicants who have completed compulsory schooling. Depending on the context, the group of programs considered to be constitutive of higher education may differ. In Australia, higher education is distinct from *tertiary education*—although the two are often conflated—which is broader in scope, and includes *all* legally accredited post-school programs of study. The kinds of programs included in the local tertiary education sector are specified according to a *qualifications framework*, which has become the standard means by which to specify non-compulsory formal qualifications (Allais, 2012b). Tuck (2007) offers a helpful summary of their usual features:

A Qualifications Framework is an instrument for the development, classification and recognition of skills, knowledge and competencies along a continuum of agreed levels. It is a way of structuring existing and new qualifications, which are defined by learning outcomes, i.e. clear statements of what the learner must know or be able to do whether learned in a classroom, on-the-job, or less formally. The Qualifications Framework indicates the comparability of different qualifications and how one can progress from one level to another, within and across occupations or industrial sectors (and even across vocational and academic fields if the NQF [national qualifications framework] is designed to include both vocational and academic qualifications in a single framework). (p. V)

In Australia, the range of formally recognised non-compulsory qualifications available is specified according to a national qualifications framework called the Australian Qualifications Framework (AQF, see Australian Qualifications Framework Council, 2013). In the AQF, qualifications are organised on a scale of levels from 1–10, where level 1 indicates basic skills-based training certificates, level 7 corresponds with bachelor degree studies, and level 10 is associated with doctoral qualifications. The feature that distinguishes an institution as legally capable of awarding qualifications defined by the AQF is registration with the Tertiary Education and Quality Standards Agency (TEQSA). TEQSA is the sole national regulatory body responsible for oversight of institutions' accreditation to award qualifications specified by the AQF. The standards according to which higher education providers are regulated are defined in the Higher Education Standards Framework (HESF, see Australian Government, 2015) which includes standards for assessment. Currently, qualifications considered to be constitutive of higher education in Australia are those specified between levels 5 (a diploma) and 10 (a doctorate) of the AQF (TEQSA, 2017). In other words, "[t]he defining feature of a higher education provider is the legal power to issue qualifications, ranging from a diploma to a PhD" (Norton & Cherastidham, 2018, p. 9).

Higher education in Australia is provided by a range of different kinds of providers. The broadest distinction between registered providers lies between universities and non-university institutions. Although most higher education students in Australia attend universities (Norton & Cherastidtham, 2018), some 131 non-university institutions are registered to award qualifications. In this thesis, the focus is on the higher music education provided by universities, although the discussion developed throughout the thesis is applicable to the broader range of institutions. The reason for this narrower focus is that the inner workings of non-university institutions are considerably less available to public scrutiny. While most universities make publicly available a substantial amount of detail relating to institutional policies and curriculum, access to similar documentation is far more limited within the non-university higher music education sector.

2.5.2 Higher music education

At the international level, higher *music* education includes a diverse span of curricula offered by various kinds of institutions. Fautley and Murphy (2015) distinguish between three general forms of music education:

- *Music education*: In this usage... the principal focus is with a general education concerning music. It encompasses music generative activities, such as composing and improvising; performing opportunities, and learning about and listening to music.

- *Musical education*: This involves a very specific focus on a small subset of the above, often performing technique, and takes this to be the principal reason for the activity taking place. It can encompass learning at school, in communities or outreach centres; in some jurisdictions this becomes what is meant in totality by 'music education' (thus adding to the confusion). This is certainly what takes place at the conservatoire, and is in many cases delineated by a concentrated focus on developing expertise.

- *Music in education*: Unlike the first two, this takes as its centre of concern the role that music plays in the general education of learners. It might involve either (or both) of the above, but is conceptually

separated in that it pays singular attention to the way music functions as a timetabled subject on the school curriculum. (List adapted from Fautley & Murphy, 2015, pp. 2–3, italics added)

While Fautley and Murphy (2015) emphasise that these definitions are simplistic and have "a degree of porousness between them" (p. 3), they provide a useful point of departure for distinguishing curricular emphases in higher education. In this thesis, higher music education generally includes the first two kinds proposed by Fautley and Murphy, both of which can include a substantive focus on the performance of music. Contrasts in curricular emphases have been associated with different kinds of educational institutions. Jørgensen (2010) offers a distinction between

two broad groups of institutions concerned with higher music education. The first one is the European conservatoire tradition, where the institutions are called Conservatoire of Music, Academy of Music, College of Music, Musikhochschule, and Musikkhogskole. Most of them have programs for a wide range of professional activities, with the education of professional performers as a central mission…. In the U.S. there are also several conservatoires, and there are conservatoires in Canada, China, Korea, Japan, Australia, and many other countries. The other type of higher music education institution belongs to the university tradition. These institutions are primarily found in the U.S. and they are mostly called School of Music. These institutions may have a performance program, but the education of music teachers is generally their major mission.[3] (pp. 67–68)

While Jørgensen (2010) correctly observes that a number of Australian higher music education institutions are formally named as conservatoires, this group of institutions has been absorbed into the local university sector, blurring the distinction between the two historically distinct types of institutions. An

---

[3] In this thesis, higher music education is considered distinct from the education of music teachers, usually referred to as *music teacher education*, which is generally not the focus of this research.

important historical background to this situation is the program of reform that followed the release of the Dawkins White Paper[4] in 1988 (Department of Employment, Education and Training). The reforms that took place in the following decade have had a lasting impact on the landscape of higher music education in Australia (Forbes, 2016), and perhaps the most significant development was the merging of then-independent conservatoires and academies with universities. Harrison (2014) describes this as the bringing together of "two very different cultures" (p. 5): where historically "[c]onservatoires had an intense focus on performance at the undergraduate level; universities focused equally intensely on research at the honours, masters and doctoral levels" (Harrison, 2014, p. 5). According to Forbes (2016), Australian higher music education is "still grappling" (p. 53) with the effects of this reform. She notes: "Some 30 years later the aftershocks of these changes still reverberate throughout the academy, and the pace of reform during the 2000s has not abated" (2016, pp. 53–54). Perhaps more optimistically, Harrison qualifies that "[s]ignificant progress has been made in merging these two disparate cultures since that time, though music education still tends to privilege practice, performance and virtuosity, sometimes at the expense of scholarly-based approaches to learning and teaching" (2014, p. 5). Although the Dawkins reforms have attracted critique (Wilson, 2018), one of the by-products of this reform is that it has created a situation where the secret gardens (Chapter 1) of higher music education in Australia are perhaps both more available and more able to participate in research studies such as this one.

2.5.3 Musical performance

Although higher music education institutions such as conservatoires have traditionally been associated with classical performance training and jazz education more recently (Tregear et al., 2016), the scope of the sector has

---

[4] The White Paper, also referred to as the Dawkins Review, laid out a plan for long-term development and restructuring of Australia's higher education system concerned with determining "reforms that would expand the capacity and effectiveness of the higher education sector" based upon the perception of a need to "develop our advanced manufacturing and service industries, and increase reliance on high level skills, adaptability and entrepreneurship, to compete in a globalised economy" (Department of Education and Training, 2015, p. 11).

diversified considerably in recent decades. As Monkhouse (2010) notes, it is not unusual "for the performance stream of a Bachelor of Music to offer instruction in a range of musical idioms such as classical, contemporary, jazz and rock or to have more than one staff member teaching a particular instrument" (p. 3). Although many non-performative program offerings are now available, performance studies remain characteristic of higher music education (Lebler, 2015; Monkhouse, 2010; Wrigley, 2005), and today a Bachelor of Music is the "degree of choice" (Monkhouse, 2015, p. 62) for those seeking professional musical performance training. For students, their performance abilities are an important form of social capital (Kingsbury, 2001; Perkins, 2013) within the social dynamics of higher music education institutions: As Vella and English (2015) assert, many students' identities are "defined by their relationship to their instrument" (p. 114).

Despite the rapidity of change, higher music education is steeped in the traditional practices of Western conservatoire education which has "been in the business of 'conserving tradition' for hundreds of years" (Parkes, 2010, p. 101, see also Daniel & Parkes, 2015). While the uptake of new assessment approaches has thus been slow (Partti, Westerlund, & Lebler, 2015), the introduction of relatively new disciplines of musical study into higher education—such as popular music—have offered fertile ground for the development of alternative assessment methods (see for example Lebler, 2008). Scholars agree that the *master-apprentice* model of learning and teaching has continued to dominate in higher music education (Bergee & Cecconi-Roberts, 2002; Daniel & Parkes, 2015; Monkhouse, 2010; Parkes, 2010; Partti et al., 2015). This approach characteristically positions the teacher as "the initiator and verifier of the activity" (Westerlund, 2006, p. 120) and is underpinned by "the notion of the expert performer or teacher as assessor and provider of feedback" (Daniel & Parkes, 2015, p. 91). This model is not without strengths, as Monkhouse (2010) points out: "[T]he strength of this modality is that tuition is individually tailored for each student and adjusted to accommodate the speed of accomplishment of any specific task" (p. 3). Nonetheless, "what works well in an individual studio may not necessarily transfer easily to an institutional environment" (Monkhouse, 2010, p. 3). While scholars generally agree about the dominance of the master-apprentice

tradition, its influence is not universal. Lebler (2008), for example, tells us that "[p]eer learning is common in popular music, where knowledge acquired alone is shared, while the traditional master apprentice and formal tuition models found in the study of classical and jazz music are relatively uncommon" (p. 195).

2.5.4 Musical performance assessment tasks

The kinds of assessment practices with which this work is concerned are those associated with the formal assessment of students' musical performance work, which are *formal* in the sense that they are fixed in the design of curriculum. The kinds of assessment tasks used in higher music education are helpfully reviewed in a chapter by Zhukov (2015) who tells us that the dominant approach to assessment reflects a reliance "on practical examinations as means of assessing instrumental learning" (p. 54). Zhukov contends that "[t]his ingrained tradition of assessment tends to influence music administrators and heads of departments to continue in the same way" (p. 54).

In Australia, the assessment of musical performance at the university level typically takes place either in classroom settings—performances in workshops, for example—or in examination settings characterised by recitals, technical juries, public performances, and so forth. Although differences do exist between institutions, the general approach is that performance assessment takes place in a live setting, in the presence of either a single assessor or in front of a panel who develop a written report on the student's performance during the examination event (Zhukov, 2015). While some institutions do record students' performance assessments, such practices tend to be exceptions to the norm (Daniel, 2001; Zhukov, 2015). Some musical performance assessments, such as recitals, are typically for a substantial proportion of students' total grades. Others, such as scheduled workshop performances, may be non-graded assessments in which students are simply required to take part.

The following chapter develops a more substantial discussion of assessment itself, however two other categorical distinctions are worth noting here. Firstly, a pair of general categories of performance assessment that are common in higher music education are *solo* performance assessments—

where the focus is on the work of an individual student—and *ensemble* performances, where the focus is on multiple students' work at one time[5]. Although this thesis is not directly concerned with this distinction, it is notable for its conceptual influence on the landscape of research about musical performance assessment where studies are often organised according to these categories or otherwise make explicit mention of the distinction (see for example, Barratt & Moore, 2005; Ginsborg & Wistreich, 2010; Harrison, Lebler, Carey, Hitchcock, O'Bryan, 2013; Russell, 2015). Secondly, Blom and Poole (2004) highlight that meaningful categorical distinctions between kinds of assessment include those where the emphasis is on who is involved as the assessor(s). Traditional models of musical performance assessment emphasise members of institutional staff or external assessors (usually professional performers) as the legitimate assessors (Blom & Poole, 2004; Partti et al., 2015). Although it is well-established, this approach has attracted some criticism—after Heron (1988), Blom and Poole (2004) refer to it as "authoritarian" (p. 112), for example. Concurrently, increasing attention is being devoted to what Partti et al. (2015) refer to as *participatory assessment*, which are those forms of assessment where students are actively involved in the assessment of their work, as in peer and self-assessment (see for example, Blom & Poole, 2004; Blom & Encarnacao, 2012; Blom, Stevenson, & Encarnacao, 2015; Daniel, 2001, 2004; Lebler, 2006, 2008; 2015).

2.6 Chapter Summary

This chapter was the first of three chapters to develop the conceptual framework of the thesis. The main focus here was on describing the contextual location of the work by clarifying some of the basic positions by which the research is oriented. The first part of the chapter (Sections 2.1 & 2.2) introduced the conceptual framework of the research and explained the role and characteristics of the framework. The second part of the chapter (Section 2.3) described the basic theoretical positions that overlay the research. In

---

[5] The boundaries between solo and ensemble assessment are not always clear—many forms of jazz performance assessment, for example, involve the assessment of individual students performing in ensemble contexts (see for instance, Barratt & Moore, 2005).

particular, this section articulated a realist research perspective and described a view of assessment practices as social practices characterised and regulated by discourses of legitimation. The third part of the chapter (Section 2.4) described a view of the research territory and introduced the idea of semantic gravity as a device for locating the relationship of different modules within the conceptual framework to the substantive context of musical performance assessment in higher education. The final part of the chapter described the main features of this context and their relationships to the study. The next chapter continues to describe the conceptual framework of the thesis and focuses more specifically on the subject of assessment itself to situate the research questions that guide the study.

Chapter 3: Conceptual Framework (Part 2)
Framing the Research

3.1 Introduction

This is the second chapter of the thesis that explains the conceptual framework of the research. The previous chapter introduced the framework and explained the main contextual positions of the research. This chapter focuses more specifically on situating the research questions introduced in Chapter 1. These questions are:

1. How can bases for achievement in musical performance work be meaningfully conceptualised?
2. How can bases of legitimation for assessment participants be meaningfully conceptualised?
3. How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised?

At the broader level these questions are unified by an interest in the meaningful conceptualisation of assessment practices (Chapter 1). Each section of the chapter develops the conceptual framework by focusing on one research question to explain how the question is framed by extant research. In this, the intention is not only to explain important framing concepts from the assessment literature, but to situate these in relation to other parts of the conceptual framework to make their relationships to the research as transparent as possible.

3.2 Conceptualising Bases of Achievement in Musical Performance

This section frames the first guiding question of the thesis: How can bases for achievement in musical performance work be meaningfully conceptualised? This question reflects the evolution of one of the original inspirations for this thesis which was a concern with the robustness of assessment process in general, but in particular, with how the subjectivity inherent in musical performance assessment could be productively managed (see Appendix A).

This question is closely related to the third research question, and the meaningful distinction between the two is that the emphasis here is on the issue of conceptualising musical work itself, where the emphasis in the third question is on possible approaches for designing and enacting assessment. This section begins by describing the background to the question and locates a focus on what is actually being (or to be) assessed. The second part of this section describes some approaches for conceptualising criteria for assessment in musical performance. The final part of this section considers some limitations of extant approaches and situates the project in relation to this gap.

3.2.1 Background to the question

Some basic concerns for assessment in any field are the questions of *what to assess* and *what is assessed* (Broadfoot, 2012; Lebler & Harrison, 2017; Shay, 2008a). Related to this is the "cliché" (Thompson & Williamon, 2003, p. 25) problem that appraisal of creative works such as musical performance is unavoidably subjective "and so the implication goes, unreliable" (Thompson & Williamon, 2003, p. 25, see for instance Bergee, 2003; Parkes, 2010; Russell, 2015). This in turn has been attributed to the complexity of musical performance, and of the process of appraising complex work. Musical performance is complex in production in the sense that it is "among the most complex forms of skilled serial action produced by human beings" (Palmer, 1997, p. 117). At the same time, it is complex from a creative standpoint in that there is often considerable latitude for students to respond to assessment tasks in a variety of ways (Blom, Stevenson, & Encarnacao, 2012). Daniel and Parkes (2015) develop a useful summary of the situation:

> The performance of music is a very personal act, with interpretation and individuality increasingly distinguishing musicians from their peers as they develop their own style and approach (Wesolowski, 2012). Two performers are unlikely to play a work the same way, indeed assessors who review and critique performers also rarely have exactly the same view or assessment of the performance. While in some disciplines there

are clear expectations in terms of solutions or outcomes, such as in mathematics or accounting, in music there is significant potential for interpretation of performance and hence subjectivity in how it is assessed (Monkhouse, 2010; Wesolowski, 2012). Indeed, assessment of performance has previously been described as hazardous (Stowasser, 1996), while Wesolowski (2012) recently argued that many music teachers even "rely heavily on nonmusic criteria, such as behavior, attitude, attendance, and participation to determine their grades" (p. 37). (Daniel & Parkes, 2015, p. 91)

Of course, not all musical performances are equally complicated to assess, however even simple musical performance tasks involve complex interactions between different kinds of knowledge and skill (Fautley, 2010). Many musical performance works are what Sadler (2009a) refers to as *divergent*, in that "there is no single correct or best answer, result or solution" (Sadler, 2009a, p. 160). The assessment of such works typically involves engaging students' "abilities in both design and production, the objective being to allow considerable latitude for creative solution, analysis or expression" (2009a, p. 160). Fautley and Burnard (2015) summarise some of the main reasons for which creative works are "arguably the most difficult to assess" (p. 225):

We can determine what someone knows simply by asking for their recollection or application of knowledge (declarative knowledge). We can assess understanding by asking for a response which explains or rewords, or which shows how the knowledge can be applied (procedural knowledge). It is, however, a different matter to capture, document, show, amplify, represent or even visualise routes to creativity and analyse creative practices….Of crucial importance to assessment is the myriad of dimensions expressed in the characterisations of musical creativities. (Burnard & Fautley, 2015, pp. 255–256)

The issue of what *to* assess, Lebler and Harrison (2017) argue, is given "in times of increasing regulation" by regulatory structures, where "the 'what' that

should be assessed has been defined by the regulatory authorities" (p. 96), and the bases of students' achievement are given by mandated learning outcomes encapsulated in qualifications frameworks (in Australia, the AQF; see Section 2.5). *Learning outcomes* therefore occupy a critical juncture between disciplinary curricula and the broader sectors in which curricula are located (Allais, 2012a, 2012b). The central notion is that "learning outcomes somehow capture a 'sameness'" (Allais, 2012a, p. 333) which enables comparison and compatibility "between nation states, different parts of education and training systems, or between education programmes and life (especially work) experiences" (p. 333). The focus on outcomes within qualifications frameworks—in particular, the emphasis that has been placed on generic skills outcomes (Bridgstock, 2009)—is of notable significance for higher music education, where incongruities have been observed between the kinds of *knowledge workers* that universities are seeking to develop and the kinds of graduates that musical performance institutions have historically sought to produce (Carey, Grant, McWilliam, & Taylor, 2013). It is a situation characterised by the "requirement of students to be analytical and critical when their aspirations are most often associated with playing their instrument" (Vella & English, 2015, p. 113). Related to the high degree of transferability associated with generic learning outcomes, Allais (2012a) tells us, is the situation that qualifications frameworks constitute an "integral component" (p. 333) of quality assurance practices:

> The idea here is that national regulatory bodies will be able to measure programmes against the outcomes, and employers and educational institutions, whether at home or in other countries, will then have a good sense of what it is that the bearer of a qualification is competent to do. (Allais, 2012, p. 333a)

All of this places a good deal of cumulative pressure on decisions about what to assess at the disciplinary level. This is complex for the reason that the relationship between learning outcomes at the course level and those produced at regulatory level is unavoidably abstract and interpretive (Sadler, 2014), which speaks to the value placed on curricular mapping and alignment

activities (see for example, Lebler, Harrison, Carey, & Cain, 2015). Published statements about learning outcomes are high-stakes, for the reason that they reflect claims about what graduates can expect to have acquired through the acquisition of a given qualification. This means that *what is actually assessed* has strong implications for the *fidelity* of an assessment task, which refers to the idea that "assessment evidence [should] be of a logically legitimate type" (Sadler, 2010, p. 728). In turn, this places emphasis on what it is that connects these outcomes and how these connections are, perhaps more importantly, meaningfully accessible and interpretable (or not) for those involved in assessment practices. Central to the fidelity of assessment are *criteria*, the singular *criterion* referring to "[a] distinguishing property or characteristic of any thing, by which its quality can be judged or estimated, or by which a decision or classification may be made" (Sadler, 1987, p. 194).

3.2.2 Situating criteria

The positioning of criteria as the notional basis for determining students' achievement reflects a departure from *norm-referenced* approaches where a student's achievement is determined by the relationship of the quality of their work to that of their cohort (Bloxham & Boyd, 2007). Criteria are, therefore, a focal point of research about educational assessment. The explication of criteria has been linked to the transparency of assessment and is often considered "*sine qua non* for good assessment practice" (Shay, 2008a, p. 596, italics original), although this has attracted a robust critique from some prominent scholars (see for example, Sadler, 2014, 2015). In the higher education context, part of the importance of criteria is that they represent a connection to the learning outcomes associated with a program of study.

The discussion that follows is clarified by explanation of one important conceptual point of distinction for criteria, which is that some criteria are made explicit in formal course documentation, and others are not. Sadler (2009a) refers to the former kind as *preset* criteria which are determined in advance of an assessment event and are generally provided to students beforehand. This is an important distinction and is discussed further in Section 3.4, however it is not the focus of the immediate discussion which concerns both implicit and explicit criteria for musical performance. Accurately determining criteria for

creative works such as musical performances is difficult for the reason that this kind of work involves a complex synthesis of multiple skills and knowledges relating to both production *and* evaluation:

> Appraising the quality of such responses typically involves multiple criteria, often interlocking or overlapping. The assessor's brain is the main 'instrument' for what is clearly a 'qualitative' judgment, meaning that it involves no quantification—aspects that are counted or measured. (Sadler, 2015, p. 11)

Further, impressions of the quality of students' work may be influenced by factors that are not readily apparent to assessment participants as criteria (McPherson & Thompson, 1998; McPherson & Schubert, 2004; Wesolowski, Wind, & Engelhard, 2015). This complexity sits at the centre of a debate about assessment in music education where the main point of contention relates to the appropriateness of conceptually segmenting and quantifying musical performances for the purposes of assessment (Wrigley, 2005). This subject is related to the discussion about the explication of criteria and is dealt with in further detail later in the chapter (Section 3.3.4). The focus here, rather, is on the conceptualising of criteria themselves.

3.2.3 Conceptualising criteria

The previous sections outlined some of the reasons for which the task of conceptualising criteria actually matters. The concern here is with conceptual approaches themselves which are subject to issues of categorical tension. At the broad level, one example of this is "the tension between assessing craftsmanship and artistry" (Lebler & Harrison, 2017, p. 93). Broad categories such as craftsmanship and artistry go by a range of common names—for example, technical and expressive skills (Russell, 2015)—and can act as containers for finer-grained concepts (Blom & Encarnacao, 2012; for a substantive example see the criteria developed by Wrigley, 2005). Different assessment tasks may place varying emphases on different criteria, and some scholars suggest that certain skills may be prerequisites for students to achieve in other areas (Russell, 2015). The granularities of categories

(reflecting which categories are broader than others) for criteria are not consistently defined. Reviewing the literature on musical performance skills, for example, Blom and Encarnacao (2012) distinguished eight general categories, including technical skills, analytical skills, appreciative skills, personal skills, interpersonal skills, organisational skills, on-stage criteria, and criteria particular to the rehearsal process. Further, criteria for musical performances may include constituents further removed from the musical performance itself. Citing several prior studies (Davidson & Da Costa Coimbra, 2001; Elliot, 1995/1996; Thompson, Diamond, & Balkwill, 1998; Williamon, 1999) Thompson and Wiliamon (2003) warn that assessment may be influenced by "extramusical factors" (p. 27) including appearances, "adherence to accepted protocols" (p. 27), and that assessors may employ diverse internal criteria in formulating judgements (see also McPherson & Thompson, 1998; McPherson & Schubert, 2004; Wesolowski et al., 2015).

In an effort to productively conceptualise criteria for musical performance assessment, scholars have adopted and developed frameworks that speak to more theoretical levels of abstraction. After Davidson and King (2004), Blom and Encarnacao (2012) distinguish between "two levels of knowledge" (p. 26). One level encompasses general knowledge that "provides performance rules and regulations based on historical, social, and cultural factors" (Davidson & King, 2004, p. 105), while the other relates to specific musical knowledge—"information that must be processed and responded to in an ongoing manner" (p. 105). Others have proffered similar distinctions between musical knowledge—in the sense of tightly specified concepts—and broader, more generalised bodies of knowledge relating to musical performance and its assessment. McPherson and his colleagues (McPherson & Schubert, 2004; McPherson & Thompson, 1998), for example, offer an influential framework, distinguishing between musical, extra-musical, and non-musical factors that may influence the appraisal of musical performances. Within the latter pair of categories, Wesolowski, Wind, and Engelhard (2015) further distinguish performers' characteristics, contextual characteristics, and assessors' characteristics. Goodman (2002), in the context of ensemble performance, distinguishes between musical and social aspects of interaction; the former including factors such as timekeeping and aural skills, and the latter

factors such as communication and leadership (see also Davidson & King, 2004).

In extending this group of frameworks—broadly conceptualised as "groupings of musical and non-musical/social factors" (Blom & Encarnacao, 2012)—Blom and Encarnacao (2012) propose a distinction between "'hard' and 'soft' skills… to describe technical/cognitive (hard) and behavioural (soft) skills" (p. 26). After Coll and Zegwaard (2006), Blom and Encarnacao (2012) summarise that

> 'hard' skills comprise technical skills which 'represent the ability to apply technical knowledge with some expertise' (p. 31), analytical skills 'and constructive skills [which] are concerned with problem identification and the development of solutions' (p. 31), and appreciative skills which 'refer to the ability to evaluate complicated situations, and make creative and complex judgements' (p. 31). Soft skills 'comprise personal skills, how one responds to, and handles, various situations; interpersonal skills, securing outcomes through interpersonal relationships; and organisational skills, securing outcomes through organisational networks' (p. 31). (Blom & Encarnacao, 2012, p. 26)

3.2.4 Visualising the problem

The approaches described above are related to a broader family of descriptive typological tools—*knowledge typologies* (Maton, 2014)—which includes notable contributions from Biglan (typologies of hard/soft, pure/applied, life/non-life, see 1973a, 1973b), Kolb (typologies of abstract/concrete, active/reflective, see 1981), and Becher and Trowler (taxonomy of tribes, see 2001), whose works have been influential in the context of educational research. The traction that typological approaches have had in the music education context speak in part to the broader orientation of education towards the "knowledge economy" (Carey & Lebler, 2012), where knowledge is the principal resource to be cultivated and accumulated (Allais, 2012b; Maton, 2014; Torrance, 2017). Making knowledge visible has, as such, been an important item on the educational research agenda (Maton, 2014; Shay,

2008a). Although knowledge typologies "valuably bring knowledge into view as an object of study... they also possess an inbuilt limit to their explanatory power thanks to their segmental form" (Maton, 2014, p. 127). On the one hand, typological approaches have helped to make visible a broader spectrum of substantive influences on student achievement that would otherwise lay "hidden beneath the general guise of assessment" (McPherson & Schubert, 2004, p. 78). On the other hand, these models are limited in terms of their capacity to actually represent practice. This limitation helps to explain conclusions from scholars such as Blom and Encarnacao (2012), who found a strict delineation between categories (hard and soft) to be "neither desirable nor possible" (p. 26). Likewise, McPherson and Schubert (2004), reflecting on a distillation of four main categories (technique, interpretation, expression, and communication), describe an "inevitable overlap" (pp. 64–65) between categories. The common issue is that "[w]hether expanding or contracting, overlapping or integrating the forms of knowledge delineated... [typological models] nonetheless offer a series of static types into which few empirical practices and processes fit" (Maton, 2014, p. 127). Maton (2014) argues that deliberation over the qualities or shortcomings of typologies highlights yet misunderstands the problem itself:

> The problem is not whether typologies offer sufficient categories to capture the manifold diversity of knowledge practices but rather that *this kind of theorizing* cannot *by itself* embrace such diversity. This is not to argue that typologies are a misstep but rather that they are a valuable first step. The next step is to additionally conceptualize the organizing principles that generate these diverse types (and other types yet to be delineated). (Maton, 2014, p. 127, italics original)

Oriented to the subject of criteria, Maton's (2014) argument helpfully places the focus on the interpretation of the problem. What alternatives might be available if we suspend a typological view of criteria, and what insights might such approaches afford? The explorations conducted in the later chapters of this thesis are, in part, a response to this line of questioning.

3.2.5 Summary

This section contextualised the first research question that guides the study. It described a gap in the landscape of conceptual thinking about criteria for musical performance assessment by highlighting the limitations of typological approaches. The rationale for exploring further possibilities for conceptualising the forms of knowledge inherent in criteria is the dominant role of knowledge in broader discourse of higher education, including as a currency for regulatory discourse. The theoretical framework used in this study was selected for the reason that it offers a set of conceptual devices for exploring assessment criteria on different terms—the main features of this framework are outlined in the following chapter. The next section in this chapter contextualises the second research question that guides the study.

3.3 Conceptualising the Legitimation of Assessment Participants

This section frames the second guiding question of the thesis: How can bases of legitimation for assessment participants be meaningfully conceptualised? *Assessment participants* is used as a broader term here and includes both students and assessors (who might be staff members, external assessors, other students, and so forth). At the general level, this question is inspired by the relevance of *personhood*[6] to assessment practices—it reflects a basic concern with how *who you are* matters in assessment practices. This question is distinct from the first and third research questions in that it evolved later in the project following a period of engagement with the theoretical framework described in the following chapter (the rationale for this was provided in Section 2.2).

3.3.1 Background to the question

The main argument that frames this question is that studies of musical performance assessment have provided very little explanatory theory about assessment participants themselves. This is not an argument that the people involved are obscured in musical performance assessment research studies—

---

[6] The quality of being an individual person.

indeed, most studies on musical performance assessment highlight the importance of assessment participants' characteristics. Building meaningful theory about the characteristics of assessment participants is rarely the object of these studies however, and the primary focus is usually on issues of design or approach. This is the more specific gap to which this research question is oriented, and its existence became apparent later in the project following engagement with the theoretical ideas described in other parts of the conceptual framework (see Section 2.3 and Chapter 4). An important influence in this respect was the work of Maton (2014) who highlights the significance of *social relations* in practices. In Maton's (2014) terms, social relations exist "between practices and their subject, author or actor (who is enacting the practices)" (Maton, 2014, p. 29). For assessment practices, they signify an interest in how the characteristics of people and their relationships to one another and to practices can influence the capacity of those people to successfully engage in those practice. At a very general level this is related to the idea that people bring to the world of practice various resources, perspectives, and positions (including formal positions as well as identities)—different forms of what Bourdieu referred to as *capital*[7] (Moore, 2008). Maton (2014) highlights that these attributes of people include both knowledge and skills (epistemic attributes), and values, identities, and taste (social attributes). Thinking in these terms during early interviews with musical performance staff (assessors) and students at one institution emphasised the relevance of a focus on the people involved in assessment as distinct from products (students' works) or the actual assessment practices. Comments from these research participants included, for example:

> I have a problem with a numerical result [for] performance. Because, *being a performer myself* I think it works against the concept of performance. (Assessor B, a classical department leader, italics added)

---

[7] Bourdieu used this term to describe different kinds of assets, including but not limited to monetary forms. Some well-known forms include economic, cultural, and social capital—a detailed but accessible description is provided by Moore (2008). In this thesis capital is used in the Bourdieusian sense, but only at the general level.

> *I feel as the examiner* aware of that and aware of *wanting to make people feel as comfortable as possible* because I want to see them do their best, I don't want to see them fail... So yeah I think there is *an emotional component* that is I think a little bit tricky. (Assessor D, a classical performance teacher, italics added)

> It was an American-written piece and *she [a teacher] had a very strong basis of American study* and was drawing upon the different inspirations that contributed to this work's composition... When I performed that she was on my panel, as well as *another teacher whose training background was from Versailles in France.* And so he had actually a lot to say about that piece... what more I could make of it as a musical demonstration in terms of, often, a lot of contrast and colour and character. Very different interpretive contexts. (Student A, a classical percussion student, italics added)

In the context of research about musical performance assessment, the relationships between people and practices have been studied less clearly than have issues related to the conceptualisation of musical performances (the previous section) or strategies for assessing performance (the following section). Of course, extant research about musical performance assessment usefully highlights the importance of social relations in assessment processes. Indeed, this is perhaps true of most research about musical performance assessment which highlights myriad ways in which assessment participants' characteristics matter. The studies cited in Section 3.2, for example, highlight the influence of social factors on the grading of students' work.

Several other studies make a particularly compelling case for this research direction. Barratt and Moore (2005) for example studied the ways in which jazz students at one British institution undertook recital assessment tasks. The recital task in question involved students forming small ensembles with other students to provide a performing context for the assessed student. Barratt and Moore discovered that students adopted particular strategies in

assessed recitals which could inadvertently lead to a limiting of students' achievement—these strategies included, for example, attempts on the part of the supporting students not to over-interact with the assessed students, and thereby divert attention away from the assessee. Barratt and Moore noted that in making such allowances students were inadvertently inhibiting their ability to demonstrate the levels of interaction and risk-taking typically associated with high level jazz combo performance. Barratt and Moore make the following comment in discussing their findings:

> Traditional assessment approaches, when applied to jazz, were encouraging students to make allowances for the assessed musician, by moving away from the natural group interactions, the improvisatory flair, and the democratic contrapuntalism necessary for most fine jazz performance. Decisions concerning the production of ideas, as well as direction and form of the group performance, seemed to be left to the member that was being assessed at the time, which in turn appeared to generate a false emphasis on one individual within a medium that is predominantly a group activity. (Barratt & Moore, 2005, p. 303)

Barratt and Moore continue to develop a useful discussion of the effects of assessment design on students' approaches, and their article provides a helpful discussion of the influence of assessment approaches inherited from classical performance training traditions in non-classical disciplines such as jazz. In directing their attention to assessment design, however, the authors shift attention from the people who are enacting the assessment practice. This is not a critique of the contributions of Barratt and Moore (2005) but rather an illustration of a type of blind-spot that appears to be relatively common in research studies about musical performance assessment—that is, the people involved in assessment as a central focus of theoretical inquiry and theory-building. It is perhaps an elusive blind-spot for the reason that its effects may be less clearly apparent than those of the more tangible features of assessment encapsulated within designs. It is difficult to locate an explicit discussion or recognition of this gap in the literature on musical performance

assessment, however contributions from the broader landscape of research about assessment in higher education provide some useful examples of what a more explicit focus on people might look like in assessment research. Boud et al. (2016) identify several studies of assessment in higher education that adopt socio-cultural perspectives, including James and Diment (2003), Shay (2004), Pryor and Crossouard (2008) and Rust, O'Donovan, and Price (2005). Boud et al. (2016) tell us that these studies place

> assessment as part of a social activity in which students acquire understanding of assessment processes, criteria and standards through active engagement and participation. These accounts have gained little traction amid the dominant everyday discourse of testing and marks, leading us to consider whether they might be extended to create more productive perspectives on assessment. (p. 1109)

Within the broader landscape of research on music education socio-cultural perspectives appear much more common, and a range of examples are available in edited publications by Wright (2010) and Burnard, Trulsson, and Söderman (2015). One possible explanation for the differences between kinds of research approaches that have seen uptake in the fields of music education research and musical performance assessment research is that the fields have yet to meaningfully converge as regions of scholarship. Another explanation might be differences in regional perspectives on assessment around the world. Neither of these possibilities are explored in any detail here, however they are worth noting as possible points of further research in the future.

The study conducted by Barratt and Moore is a useful exemplar here for the reason that the effects of students' thinking about assessment are rendered particularly visible by the implications of approaches they took. Further indicators of the importance of the attributes of people in assessment can be found littered throughout the literature on musical performance assessment. Some contrasting examples (amongst very many possibilities) include:

- Pulman (2009), who describes a process model for engaging students in peer assessment, in which explicit focus is placed upon the attributes of students
- Vella and English (2015) discuss the provision of feedback to students—a situation where care is required in light of "their aspirations and expectations" (p. 113) and "the collision of students' assessment of themselves… and the formalised assessments conducted by our staff [at the Conservatorium of Music University of Newcastle]" (p. 113)
- In a well known study, Stanley, Brooker, and Gilbert (2002) observed differences in assessors' perceptions of assessment criteria and recommended that examiner training should include a focus on several aspects of assessment related to criteria

In most cases, theoretical discussions of the attributes of assessment participants are not developed to explanatory theoretical ends—while rich description is often provided, the explanatory power of these descriptions is limited by the need to decode and decontextualise each individual study in relation to each other study within a given area of focus. Generally, it is the assessment model or design that receives the most attention, and the attributes of those involved are positioned in relation to these models and designs. Although the attributes of those people involved in musical performance assessment are clearly central to practice, a robust theoretical discourse in which the main focus is on explaining the underlying principles by which people relate to musical performance assessment practices (and to one another within this) does not appear to have developed within the field.

3.3.2 Assessment literacy

One of the areas in which assessment participants are more closely construed as part of assessment practices is *assessment literacy*, a brief summary of which is provided here. Assessment literacy has emerged as a concept for describing assessment participants' facility with "the language of assessment" (Deeley & Bovill, 2017, p. 464). Looney, Cumming, van Der Kleij, and Harris (2018) provide a useful summary of the field and have synthesised a practical definition:

> Assessment literacy is usually broadly defined, encompassing both assessment knowledge and skills related to teacher practice (Popham, 2009; Stiggins, 1995) as well as use and interpretation of evidence to inform instruction, generate feedback, guide student learning, and report student achievement (Stiggins & Duke, 2008; Webb, 2002). (Looney et al., 2018, p. 443)

According to Looney et al., the concept was introduced by Stiggins (1991), and it has since been accorded considerable attention in the broader literature on educational assessment. In view of the focus of this thesis on the conceptualisation of assessment, there are several key points to be made about assessment literacy as a scholarly concept. The first is that it appears to have been directed largely at teachers rather than students. Part of the reason for this is likely that the focus in assessment literacy research appears to have been on the pre-tertiary sector, rather than the higher education learning and teaching space. In this view, a broader conception of assessment literacy potentially has much to offer both staff and students in the latter context—particularly considering the increasing interest in participatory forms of assessment such as self- and peer-assessment. Secondly, assessment literacy is a useful conceptual pole within the broader field of assessment theory for the reason that it complements a focus on the theorising of concepts with actors' understandings of concepts. This is related to the third observation, which is that assessment literacy appears often to be concerned with actors' understandings of assessment from a technical standpoint—the emphasis, as observed above by Looney et al. (2018), is frequently on knowledge and skills. As Looney et al. acknowledge, however, the concept has been extended by some scholars to include the dispositions of assessment practitioners as well. The focus of Brown (2011; Brown, Lake, & Matters, 2011) and Thompson (1992) on teachers' conceptions of assessment, for instance, reflects a broader view that highlights a "more general mental structure encompassing beliefs, meanings, concepts, propositions, rules, mental images, preferences and the like" (Looney et al., 2018, p. 444). Looney et al. (2018) summarise:

> *Conceptions* of assessment presume a more complex and iterative relationship between knowledge and practice, than the relatively straightforward and apparently technical and instrumental relationship enshrined in assessment literacy. Further, *conceptions* foreground teachers' beliefs about assessment (even those that may well be non- or irrational) as significant in shaping classroom practice (p. 444)

Helpfully, a number of scholars have investigated conceptions of assessment in the higher education context (for example, Fletcher, Meyer, Anderson, Johnston, & Rees, 2011), and some work has already been done in this space in the context of higher music education (for example, Cain, 2015). Related to the broader concept of assessment literacy, other kinds of literacy have been proposed that are attracting the interest of assessment scholars: for example, *feedback literacy* (Carless & Boud, 2018). It appears that the concept of assessment literacy has not yet achieved strong traction within the more specific field of music education, although this has been advocated for by scholars such as Brophy (2019).

The body of work on feedback literacy helpfully brings into focus assessment participants as central to quality assessment practices. It provides useful data about the positioning of assessment participants in the sense of exemplifying claims about desired attributes for assessment practitioners—these are discussed further in Chapter 7.

3.3.3 Visualising the problem

Section 3.3 began by explaining that the second research question developed through dialogue with the concept of social relations (Maton, 2014). In this the core problem with which the second research question is associated differs to that of the first. As explained in Section 3.2.5, the problem associated with the first question relates mainly to the limitations of extant typological approaches for conceptualising bases of achievement in musical performance assessment. It foregrounds *knowledge* as an object of study in the sense of identifying criteria as the substantive focus of inquiry. The issue here, rather, is that the object of study itself—musical performance assessment

participants—has been less explicitly positioned as a focus for research: the question is about *knowers* (as distinct from knowledge) and the problem can be described as what Maton (2014) refers to as *knower-blindness*. Of course, research about musical performance assessment frequently emphasises knowers (assessment participants) and the importance of their characteristics—these are far less frequently the main object of focus however, and are often locked in states of rich description. Often, although the importance of the characteristics of people in musical performance assessment is clear, the discussion is turned to questions of design or of optimal knowledge amid the "predominance of concern about techniques" (Broadfoot, 2012, p. 17). In the current knowledge economy climate of higher education (Allais, 2012b; Maton, 2014; Torrance, 2017; Vaira, 2004) the premiums placed on knowledge explain the comparative lack of attention accorded knowers in research studies about musical performance assessment. Maton (2014) highlights the significance of this issue for arts education:

> [W]hen arguing for knowledge to be seen it is easy to valorize the kinds of knowledge most easily seen: explicit, abstract, condensed, hierarchical forms that visibly announce themselves. This tendency can drift towards offering a deficit model of the arts, crafts, humanities and many social sciences, as well as everyday understandings, where knowledge may be less explicit and more concrete, context-dependent, embodied, and axiological. At this point, knowledge-blindness gives way to seeing nothing but knowledge and obscuring practices for socializing or cultivating knowers. (p. 14)

The second research question reflects a direct attempt to acknowledge and engage with the issue of knower-blindness in musical performance assessment. It is in some ways a counterpart to the first research question which reflects a stronger emphasis on knowledge. In exploring this question through the thesis, considerations are made of the descriptive limitations identified in Section 3.2.5. The focus, therefore, is on exploring the meaningful

conceptualisation of musical performance assessment participants without resorting to a typological approach.

3.3.4 Summary

This section contextualised the second research question that guides the study. It described a gap in the landscape of research about musical performance assessment in higher education in the sense of a lack of research explicitly focused on the assessment participants themselves (without presuming recourse to assessment designs). It is a subtle gap in that assessment participants and their attributes are far from absent in research about musical performance assessment. The gap pertains, rather, to the ways in which they are meaningfully located in research studies such that cumulative theory-building about them can occur. Some aspects of the research landscape, such as assessment literacy, do place a closer emphasis on assessment participants themselves in that they recognise people as central to practice—the limitation of these studies in relation to this gap is that they are generally focused on knowledge (for example, of assessment literacy standards) rather than those for whom the knowledge is intended. In the broader field of assessment research in higher education some scholars have employed socio-cultural approaches that place a stronger emphasis on the people involved in practices, however these studies are far fewer in number than those oriented towards design aspects of assessment (Boud et al., 2016). This study responds to this gap by exploring the second research question.

3.4 Framing the Legitimacy of Assessment Practices

This section frames the third guiding question of the thesis: How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised? The focus of the first question was on kinds of criteria for musical performances, while the focus of the second question was on the legitimation of participants in assessment practices. This question is distinct from the previous two research questions, in that here the focus is turned to assessment practices themselves. More specifically, the interest is in conceptualising the valorisation and/or critique of approaches to assessing musical performance.

3.4.1 Background to the question

A defining characteristic of the scholarly literature on musical performance assessment is the abundance of concern about the trustworthiness of assessment practices. Section 3.2 outlined the basic scenario leading to this situation—it explained some of the complexities associated with assessing student works in musical performance contexts, and highlighted the issue of conceptualising the works themselves. Related to this situation is the question of *how* to assess such work. At the macro level, decisions about how to assess students' work are influenced by the regulatory structures in place (Lebler & Harrison, 2017; see Section 3.2). In Australia, this is related to the requirement for institutions to adhere to formal standards specified for various disciplinary areas (Lebler & Harrison, 2017). Universities retain autonomy, however, for their individual curricular practices. At the disciplinary level, the question of how to assess musical performances remains a contentious issue, and one that speaks to the basic philosophies and ideologies of those involved (Gynnild, 2016; Wrigley, 2005). The common practice of representing students' achievement numerically, for example, contrasts with "a widely held belief that it is artificial and inappropriate, if not impossible to objectively measure and quantify music performance" (Wrigley, 2005, p. 3).

A number of scholars including Mills (1991), Stanley, Brooker, and Gilbert (2002), Swanwick (1998), and Johnson (1997) have made the case that the conceptual segmentation of musical works for purposes of assessment are problematic. Wrigley (2005) refers to these positions as "ideological resistances" (p. 3) to the "quantification of quality" (p. 3). Others have argued the opposite, proposing that valid and reliable criteria are definable on the basis of experts' "joint authoritative knowledge" (Wrigley, 2005, p. 178, see for instance Annett, 2002; Manns, 1998; Muckler & Seven, 1992). An interest in more "objective" (Russell, 2015, p. 360) approaches for musical performance assessment is closely linked to the aforementioned concern with the trustworthiness of assessing music, and in particular, the issue of reliability. This, in turn, is connected to a broader discourse about the reliability of assessment which has been going on for some time. Elton and Johnson (2002), who reviewed the literature on assessment, assert that reliability has historically been the dominant focus in research about

assessment, defined by "an assumed basis of assessment in psychological measurement" (p. 9). Positioned at odds with this is an interpretive paradigm which gained momentum through the late 90s and 2000s (see for example Moss, 1992, 1996; Shay, 2004, 2008a) in which assessment is viewed as socially constructed and contextualised (Elton & Johnson, 2002). Although some elements of the assessment literature can give the impression of a relatively agreed-upon departure from approaches emphasising psychological measurement as the basis of reliability (see the following section), the situation remains contested in the field of higher music education, where quality assurance processes stress the importance of reliability (Bloxham & Boyd, 2007), and it thus continues to occupy a central role (see for example, Bergee, 2003; Russell, 2010, 2015; Wesolowski et al., 2015).

3.4.2 Purposes and paradigms

Related to the debate about reliable approaches for assessment are contrasting positions on what assessment is actually for and what it should aim to achieve. An ongoing issue for assessment is that it "aggregates a multiplicity of purposes" (Boud et al., 2016, p. 1109), which are frequently "complex and interdependent" (Price, Carroll, O'Donovan, & Rust, 2011, p. 481), and prone to conflict with one another (Bloxham & Boyd, 2007; Brown & Knight, 1994; Boud, 2000; Rowntree, 1987; Rust, 2002). This situation has led, Bloxham and Boyd (2007) argue, "to a situation where assessment is in a state of flux, facing pressures for enhancement while simultaneously coping with demands to restrict its burden on students and staff" (p. 6). Although many purposes of assessment have been identified (Newton, 2007), Boud and Falchikov (2006) assert that assessment has two main assumed purposes, the first of which is about certifying achievement and the second of which is to "facilitate learning" (p. 401). While "different purposes emphasise different principles of assessment" (Bloxham & Boyd, 2007, p. 32), it has been noted that "[t]here are varying views on what the principles of assessment should be" (p. 34).

Two of the main conceptual touchstones for research about assessment are *measurement* and *learning.* The concepts can be likened to poles, around which contrasting "philosophical paradigms" (Shay, 2008a, p. 599) have developed (Boud et al., 2016; Carless, 2015; Elton & Johnson,

2002; Price, Carroll, O'Donovan, & Rust, 2011). In situations where measurement is the dominant paradigm, the emphasis is generally on the robustness of procedures for measuring students' achievement. Assessment is seen to be about "tests and examinations" (Boud, et al., 2016, p. 3), wherein the underlying assumption is that "various kinds of assessment activities can be designed to measure particular learning outcomes or characteristics of students and that the purpose of research is to improve their efficacy in doing so" (p. 3). The measurement paradigm has been closely associated with concepts and practices such as *summative assessment* and assessment *of* learning (Baird, Andrich, Hopfenbeck, & Stobart, 2017; Bennett, 2011; Wanner & Palmer, 2018). In the literature on musical performance assessment, measurement is positively construed by some scholars as a means of enhancing the objectivity of assessments. The basic rationale is that

> [e]mploying a more objective approach in the design of performance assessment allows for more focused and reliable evaluations that reduce the amount of confounding influence that so often plagues solo performance assessment situations (i.e., environmental conditions, adjudicator experience, etc.). (Russell, 2015, p. 360)

The emphasis on the objective distinction of qualities reflects the close link between the measurement paradigm and the field of psychometrics, which has had a significant influence on the field of educational assessment over the latter half of the 20th century (Elton & Johnston, 2002; Moss, 1994; Shay, 2008a). The prevalence of this influence is such that in their review of research about educational assessment in universities Elton and Johnston (2002) refer to an "assumed basis of assessment in psychological measurement" (p. 11). This is a key point of distinction between the measurement and learning paradigms. The former tends to focus on the objects of assessment (in the sense of what is assessed) as *non-social*, in the sense that "causal entities exist outside the theorist's mind" (Messick, 1989, as cited in Shay, 2008a, p. 599), and therefore more objectively observable.

Where learning is the dominant paradigm, the emphasis is most strongly on the relationship between assessment and students' learning, and on deeper student learning as a desired outcome of assessment. Here, assessment is positioned as "not just a measurement that leaves the student untouched" (Boud, 2006, p. x). Rather, it is seen to be "a strong intervention into their world of studying and it points students to what is important and what they should be doing" (Boud, 2006, p. x). The learning paradigm is closely affiliated with the concept of formative assessment (Bennett, 2011; Wanner & Palmer, 2018) and the valorisation of social constructivist approaches to assessment (Rust, O'Donovan, & Price, 2005; Shay, 2008a). Social constructivist approaches reflect a view of learning as a process in which key components are interactions between people between (it is social) and a view of learning as a process of construction by the minds of those involved (it is constructivist). This is distinct from the idea that knowledge exists independent of the learner and can therefore simply be acquired—rather knowledge is seen as embodied by people who themselves also embody many relationships to other things, including other people and other knowledge (Shay, 2008a).

The influence of the learning paradigm is clearly evident in the centrality of *learning* in more recent conceptual terminology: for example, assessment *of learning*, assessment *for learning*, and assessment *as learning* (Lebler, 2008), and Carless' (2015) *learning-oriented* assessment which is a broader concept "defined as assessment where a primary focus is on the potential to develop productive student learning processes" (p. 964).

There is a consensus amongst some scholars that the field of educational assessment is currently experiencing a paradigm shift, where the emphasis is shifting from measurement to learning, and formative assessment is "in vogue" (Bennett, 2011, p. 5; see also Boud et al., 2016; Shay, 2008a). Boud et al., (2016) note that "[t]he measurement tradition served many ends well—it desirably focused on making decisions based not on privilege, patronage or social acceptability but on desirably unbiased judgements of individual performance" (p. 1108). Problematically, however, "in doing so it prioritised certain kinds of purposes, effects and outcomes that represent only part of what we now see as the wider enterprise of assessment" (Boud et al., 2016, p. 1108). Similarly, Bloxham and Boyd (2007), tell us that "researchers

are now stressing the importance of balancing concerns about assessment of learning (certification and quality assurance) with assessment for and as learning (student learning, lifelong learning)" (p. 45). Shay (2008a) describes the assertion of a "new assessment paradigm" (p. 600), wherein "[t]he centring of learners and learning has foregrounded the formative purposes of assessment, reasserting assessment as part of the teaching and learning process" (Shay, 2008a, p. 600). Nonetheless, Boud (2010) asserts that emphases on measurement and certification remain "well entrenched in all education systems" (p. 5), and this is certainly true for the music education context (see Section 3.4.1).

Of course, measurement and learning are not mutually exclusive concepts (Boud, 2000). Boud (2006), for example, observes that an emphasis on learning "does not mean ignoring the role of assessment in certification, but ensuring that learning is considered every time assessment is mentioned" (p. x). As has already been mentioned, the measurement paradigm continues to be influential in the field of musical performance assessment. According to Bloxham and Boyd (2007), "[m]ost institutions still stress the measurement aspects of assessment with little concern for other purposes" (p. 44). Similarly, Gibbs and Simpson (2005) tell us that "[w]here institutional learning and teaching strategies focus on assessment they are nearly always about aligning learning outcomes with assessment and about specifying assessment criteria" (p. 3).

3.4.3 Approaches

Related to the various purposes of assessment (and the paradigms upon which they are predicated) are a range of conceptual approaches that frame practices and which have received much attention in the literature. This part of the chapter describes some of the main approaches that have been discussed. Many of the possibilities listed on this menu are not mutually exclusive and they frequently overlap. Their significance to this research is given not only by their functionalities for assessment design purposes, but also by their availability as ideological positions on desirable assessment practice.

3.4.3.1 Reference points for grading student achievement

At the broader level are distinctions between assessment systems on the basis of general referents for achievement. Broadly, three main approaches for determining the quality of students' work have been defined: norm-referenced assessment, criterion-referenced assessment, and standards-referenced assessment. Of the three, the first two are the most common and were introduced earlier in the chapter. Bloxham and Boyd (2007) provide a simple pair of definitions: Norm-referenced assessment "measures a student's performance in relation to his/her cohort. The grade awarded depends not only on the quality of a student's work but also on the quality of others' performance" (Bloxham & Boyd, 2007, p. 235). Criterion-referenced assessment, by comparison, "means that student achievement is tested against a set of criteria such as those linked to the learning outcomes for the assignment" (Bloxham & Boyd, 2007, p. 233). Standards-referenced assessment involves "strategies which make judgements about the level of an individual's learning. This may include shared benchmarks of expected performance or be supported by exemplars" (Lebler, Carey, & Harrison, 2015b, p. 296). Differentiating between criterion-referenced and standards-referenced assessment is slightly more complicated for the reason that criteria and standards are easily conflated (Sadler, 2005). Sadler (2005) observes that while the two terms overlap in meaning,

> the situation is complicated by the fact that the overlap is not symmetrical. 'Criterion' can cover for 'standard' in more contexts and sentence constructions than can 'standard' cover for 'criterion'…. the dual meaning use of 'criterion' clouds the discussion. Within the context of assessment and grading in higher education, both criteria and standards as distinct but related concepts have a crucial role to play, and it is necessary to be clear on which is intended at each point in a dialogue. Otherwise, meanings can slide almost imperceptibly from one underlying concept to the other even within a single discourse. (2005, p. 188)

Over the last few decades, criterion-referenced assessment has increasingly replaced norm-referenced approaches. As Shay (2008a) tells us, "[t]he articulation of assessment criteria has become a *sine qua non* for good assessment practice (Shay, p. 596, italics original), and in Australian higher education, norm-referenced approaches are "no longer acceptable practice" (Lebler & Harrison, 2017, p. 97).

3.4.3.2 Formative and summative assessment

The second broad group of approaches that have become a mainstay of the education literature are closely interwoven with the purposes and paradigms described previously. Of these, *formative* and *summative* assessment are among the most entrenched concepts in the educational assessment literature. There is a growing consensus, however, that the distinction between the two is problematic (Boud, et al., 2016; Lau, 2016; Newton, 2007). While this distinction has been "extraordinarily effective in carving an identity for assessment undertaken by students and teachers in the service of learning" (Newton, 2007, p. 151), several scholars agree that it has been ultimately unhelpful in advancing assessment theory and "might have even set it back somewhat" (2007, p. 151, see also Boud, 2000). Part of the problem is that fine-grained comparisons between the two can risk becoming "fairly semantic" (Biggs, 1998, p. 107), and this is an issue that perhaps stems from the original appropriation of the terminology from Scriven (1967) who used the concepts in the context of curricular evaluation. Interestingly, Wiliam (2017) tells us that Scriven "was quite opposed to the idea of applying this distinction in the kinds of roles that evaluation might play to the assessment of individual students" (p. 313). According to Newton (2007), the introduction of the two concepts to a broader audience in the context of the assessment of students' work can be attributed to the *Handbook of Formative and Summative Evaluation of Student Learning* (Bloom, Hastings, & Madaus, 1971).

Increasingly, scholars have turned to alternative terminology that emphasises the relationship between the given assessment approach and learning. Three terminologies have been proposed in particular: Assessment *of* learning, assessment *for* learning, and assessment *as* learning. The overall impression given by the literature is that the first two terms frequently become

respective surrogates for summative and formative assessment. Wanner and Palmer (2018), for instance, simply locate the older terminology beside the new: "assessment of learning (summative assessment)… assessment for learning (formative assessment)" (p. 1033). The similarity between summative assessment and assessment of learning in particular is relatively consistent. Bloxham and Boyd (2007), for example, write that assessment of learning is "involves making judgements about students' summative achievement for purposes of selection and certification" (p. 15). Lebler (2008) observes that it "occurs when a student's understanding of curriculum content is measured and this is the traditional role of assessment" (p. 28). Formative assessment and assessment for learning are more contested. While many regard them "as synonymous, or at least interchangeable, others draw more or less sharp distinctions between the two terms" (Wiliam, 2017, p. 313). For Wiliam, the more serious problem is that "there is a profound lack of agreement about what either of these terms mean, which makes both theoretical and empirical progress difficult" (Wiliam, 2017, p. 313).

Assessment *as* learning is also contested. In one definition it "involves students in the act of assessment as active participants and this involvement is intended to produce learning in itself" (Lebler, 2008, p. 194). Put another way, "the assessment process becomes part of learning of skills and students reflect on and assess their own learning" (Wanner & Palmer, 2018, p. 1033). In this definition, it is closely related to *authentic assessment*—an influential contemporary type of assessment which is discussed further below. A second notable definition, however, has been offered by Torrance (2007), wherein "assessment procedures and practices come completely to dominate the learning experience, and 'criteria compliance' comes to replace 'learning'" (p. 282). Ultimately, it seems likely that this new term remains somewhat in flux: According to Dann (2014), assessment as learning is already being "reinterpreted to explain dominant discourses… It is not overtly clear in current literature where AaL [assessment as learning] sits in our understandings of either assessment or learning" (p. 150).

3.4.3.3 Authentic assessment

The concept of *authentic assessment* has become a firm fixture in the broader field of education. According to a systematic review conducted by Villarroel, Bloxham, Bruna, Bruna, and Herrera-Seda (2017), it was first introduced into the scholarly literature by Archbald and Newmann (1988). It has since been discussed extensively and a variety of interpretations of the concept have been proposed. Vu and Dall'Alba (2014) provide a succinct summary, telling us that "[c]onventionally, assessment is seen to be authentic when the tasks are real-to-life or have real-life value" (p. 778). Through their analysis of authentic assessment literature published between 1988 (beginning with Archbald and Newmann) and 2015, Villarroel et al. (2017) distilled a more complex interpretation:

> *Authenticity* is understood as realism, contextualisation and problematisation when teaching and assessing curricular content (Benner et al., 2009; Raymond et al., 2013). *Realism* involves linking knowledge with everyday life and work, *contextualisation* characterises situations where knowledge can be applied in an analytical and thoughtful way, and *problematisation* invokes a sense that what is learned can be used to solve a problem or meet a need. Thereby authentic assessment aims to integrate what happens in the classroom with employment, replicating the tasks and performance standards typically faced by professionals in the world of work (Wiggins 1990). (p. 841, italics original)

One of the complexities that have resulted from the uptake of authentic assessment is the obfuscation of other discussions of authenticity that are relevant. As Wald and Harland (2017) observe, authenticity has been "mentioned in passing, claimed or discussed by scholars in relation to different aspects of higher education, including teaching, learning, assessment and achievement" (p. 751): They propose that "in spite of the growing appeal of authenticity, the use of the term is often vague and uncritical. The notion of authenticity is complex, has a range of meanings and is sometimes contested"

(p. 751). Vu and Dall'Alba, for instance, contend that the view of authentic assessment as involving tasks which are 'real-to-life' is "narrow and limited" (p. 778). Rather, Vu and Dall'Alba "argue that authenticity need not be an attribute of tasks but, rather, is a quality of educational processes that engage students in becoming more fully human" (p. 778). In a similar vein, Wald and Harland (2017) propose three "ways of understanding authenticity" (p. 752):

1. Authenticity as relating to the 'real-world'
2. The existential authentic self
3. A degree of meaning (Wald & Harland, 2017, p. 752)

Perspectives such as these point to a more flexible interpretation of the terminology wherein more specific authenticities are inductively construed rather than extruded onto educational systems—Parkinson and Smith (2015), for example, propose multiple dimensions of authenticity in the context of higher popular music education.

3.4.3.4 Analytic and holistic assessment strategies

Overlaying the previous groups of concepts are a set of approaches that have been seen to fall into two broad categories: *analytic* assessment—sometimes referred to as *segmented* assessment—and *holistic* assessment—sometimes referred to as *global* assessment. Sadler has written extensively on these approaches which have received a good deal of attention both within the scholarship on musical performance assessment and in the broader landscape of assessment research (see for example, Sadler, 2009a, 2009b, 2015). Sadler (2009b) provides a useful set of distinctions:

> In *analytic* grading, the teacher makes separate qualitative judgments on a limited number of properties or *criteria*. These are usually *preset*, that is, they are nominated in advance. Each criterion is used for appraising each student's work. The teacher may prescribe the criteria, or students and teachers may negotiate them. Alternatively, the teacher may require that students develop their own criteria as a means of deepening their involvement in the assessment process…. After the separate

judgments on the criteria are made, they are combined using a rule or formula, and converted to a grade. Analytic grading is overtly systematic. By identifying the specific elements that contribute to the final grade, analytic grading provides the student with explicit feedback. The template used in implementing the process may be called a *rubric*, or any one of *scoring*, *marking* or *grading* paired with *scheme*, *guide*, *matrix* or *grid*. As a group, these models are sometimes referred to as *criterion-based* assessment or *primary trait analysis*. (p. 45, italics original)

In *holistic* or global grading, the teacher responds to a student's work as a whole, then directly maps its quality to a notional point on the grade scale. Although the teacher may note specific features that stand out while appraising, arriving directly at a global judgment is foremost. Reflection on that judgment gives rise to an explanation, which necessarily refers to criteria. Holistic grading is sometimes characterised as impressionistic or intuitive. (pp. 45–46, italics original)

This particular patch of conceptual territory is quite contested for musical performance assessment, and some explanation for this has already been developed earlier in the chapter. A range of studies are notable for their attention to analytic and holistic approaches to assessing musical performance in relation to higher education contexts. This group of studies includes Newsome (2015), Wrigley (2005) Iusca (2014), Gynnild (2016), Cain (2015), Stanley, Brooker, and Gilbert (2002), Mills (1991), Lebler (2015), Sadler (2015), Blom, Stevenson, and Encarnacao (2015), Polifonia Working Group (2014), Thompson and WIlliamon (2003), Blom and Poole (2004), Zhukov (2015), Bergee (2003), Parkes (2010), Wesolowski, Wind, and Engelhard (2015), Wesolowski (2012), Swanwick (1998), and Russell (2010), amongst others. Both approaches have been comprehensively debated, as Sadler (2009b) writes: "The relative merits of analytic and holistic grading have been debated for many years, at all levels of education" (pp. 46). The basic

argument for holistic approaches relates to a philosophy of authenticity—that they are "fuller and more authentic" than their analytic counterparts (Sadler, 2015, p. 15). Wrigley (2005) provides a summary of these arguments, positioning them as ideological resistances to the quantification of quality. Mills (1991), for example, argues that holistic approaches are, from a conceptual standpoint, more "musically credible" (p. 179), while Swanwick (1998) asserts that

> it seems to defy the holistic nature of an art to identify several different dimensions and give a separate mark for each—say melody, harmony, texture—adding them up to get a single figure. When we conflate several observations we lose a lot of information along the way (p. 2)

Similarly, Johnson (1997) argues that criteria

> are problematic as applied to musical performance…. forms of words are not readily available to distinguish objectively between the good and the mediocre performance. Criteria need therefore to be designed to complement the examiner's subjective response rather than substitute for it, and analytical concepts prove to be less reliable than 'aesthetic terms'. (p. 271)

At the same time, arguments for analytic approaches are grounded in the pursuit of approaches that provide more objective evaluations of students' work. Russell (2015), for example, tells us that the high-stakes nature of formal educational assessment calls "for as much objectivity as possible" (p. 359) and that "[i]t is imperative to continually explore the nature of music performance assessment to attain higher levels of objectivity and accuracy in measuring performance achievement" (pp. 359–360). Using different terminology, Wesolowski, Wind, and Engelhard (2015) link the applicability of analytic approaches to requirements for validity and reliability imposed by external educational structures:

> General impression marking and related holistic scoring schemes are conducive as one method to diagnose musical

performances, as qualitative feedback can offer valuable and nuanced feedback tailored to individual performances. However, the utilization of such methods for summative, product-based evaluation purposes poses many concerns related to validity, reliability, and fairness… As music organizations such as the National Association for Music Education (NAfME) and the Associated Board of the Royal Schools of Music (ABRSM) move towards assessment models that demand a need to standardize and benchmark music performance assessments, holistic scoring procedures are not suitable. Therefore, the demand for reliable, valid, and equitable trait-specific scoring mechanisms is increasing for performance-related assessments in music[8]. (pp. 148–149)

3.4.3.5 Fairness

Linked to questions of assessment strategy of the kinds described above is the concept of *fairness*. According to a review of the concept by Nisbet and Shaw (2019), the emphasis placed upon fairness in "[i]nfluential statements about educational assessment" (p. 612) has led some to see the "gods of assessment theory as moving from the 'big two' (Validity and Reliability) to the 'big three'" (p. 612). As Nisbet and Shaw (2019) emphasise, however, the concept is contentious—the authors offer two main reasons for this:

> The first is that many writers are not clear what they mean by 'fair' (or 'unfair'). The second is that the fairness of assessments has been viewed through a range of different lenses—not only those of assessment theorists and practitioners, but also the viewpoints of the courts in different legal traditions, philosophers, social and political theorists and others. (p. 612)

---

[8] This may not reflect the practices of all organisations that offer music examinations. Lebler (2019), for example, observes that examiners for both the Australian Music Examinations Board [AMEB] and the ABRSM use holistic assessment and "undertake regular high-quality training that ensures their results are reliable and consistent between specializations and in all locations, and their training processes are remarkably similar to the consensus moderation approach" (p. 203).

One main reason for debate around approaches to assessment relates to the contention that some approaches are more valid and reliable than others, and therefore fairer (Moss, 1994; Wrigley, 2005). This is a complex assertion however, for the reason that fairness is intertwined with the broader notion of *justice* (Nisbet & Shaw, 2019), and the trustworthiness of the selected strategy may reflect only one of many possible contentions about the fairness of an assessment design or event. Further, some scholars identify a tension between the ideas of fairness and authenticity (see Section 3.4.3.3); as Elton and Johnston (2002) contend:

> Perhaps the biggest difference between the educational and the so-called real world is this insistence on fairness in the former as a criterion above all others in assessment. Once it is appreciated that life is not fair and that education is a preparation for life, it will perhaps become more acceptable to take fairness off its high pedestal in the interest of better and more relevant education. (p. 31)

Notwithstanding critiques of positivistic approaches to validity and reliability in the context of educational assessment (see for example, Elton & Johnston, 2002; Moss, 1994), a key distinction between these concepts and that of fairness is that fairness seems to be more considerably more open to interpretation (Nisbet & Shaw, 2019). In the context of this study fairness is viewed as a concept that is firmly socially-situated, and positions on fairness are construed as indicators of social legitimation (see Sections 2.3.2 and 2.3.3). This is of particular pertinence to the third research question since assertions about the fairness—whatever their philosophical basis—are distinct acts by which assessment practices can be construed as more or less legitimate.

3.4.4 Visualising the problem

As explained in Chapter 1, this thesis is not about ways of doing assessment in the sense of specific designs or techniques for assessing students' work. Accompanying the range of approaches that have been conceptualised for assessment are a similar range of issues related to selection and

interpretation. To quote Russell (2015): "Unfortunately, there seem to be as many approaches to performance assessment as there are performances to assess, and, so far, no model exists to help guide music educators in the quest for objectivity in performance" (p. 360). The interest in this research is not in the proposition of better types, nor is it about advocacy for one approach over another. This project is also not about redescribing why we might adopt one position over another—indeed, redescription of this nature, I contend, is one of the very issues we are dealing with. The interest in exploring this third research question, rather, is in possibilities for overcoming dichotomous (or trichotomous, tetrachotomous, etc.) positioning to enable a stronger dialogue *between* positions. Put another way, is it possible to overcome a typological view of positions to make more visible some of the principles that *organise* assessment practices? To reference the position introduced in Chapter 1, the intention of the research is ultimately to make the rules of the game more visible, and in so doing, to promote a more accessible discourse for all involved.

3.4.5 Summary

This section contextualised the third research question that guides the study. It discussed a range of purposes, paradigms, and approaches that frame musical performance assessment practices, including orientations towards measurement, certification, and learning, as well as conceptual devices including reference points for determining achievement, formative and summative assessment, authentic assessment, and analytic and holistic approaches to assessing students' work. These aspects of assessment overlap and are not generally mutually exclusive, however they also reflect positions about what is valued in assessment, and aligning with some positions may mean the relative downplaying of other perspectives. This study responds to this situation by exploring the third research question, and in so doing places explicit emphasis on conceptualising principles underlying the legitimation of assessment practices themselves.

3.5 Chapter Summary

This chapter was the second of three chapters to develop the conceptual framework of the thesis. The main focus here was on describing the more specific framing of the three research questions that guided the study. Each of these questions reflects a more specific trajectory for responding to the broader problem-situation outlined in Chapter 1. The first part of the chapter (Section 3.2) situated the first research question: How can bases for achievement in musical performance work be meaningfully conceptualised? It located a gap in the form of a limitation of typological conceptual approaches that have been applied in conceptualising criteria for musical performance assessment. While extant approaches emphasise the importance of interpreting knowledge (as criteria) in musical performance assessment, their explanatory power is limited by their typological structures—the first research question provides an avenue through which to respond to this limitation and the theoretical tools outlined in the following chapter offer some points of departure for overcoming this.

The second part of the chapter (Section 3.3) situated the second research question: How can bases of legitimation for assessment participants be meaningfully conceptualised? This question developed through the course of the research, emerging as a result of interaction with the work of Maton (2014) and the group of conceptual tools outlined in the following chapter. The gap toward which this question is oriented is the apparent deficit of explicit explanatory emphasis on knowers (here, assessment participants) in musical performance assessment practices. While extant research usefully highlights the significance of the attributes possessed by those involved in practice, the people themselves are rarely positioned as subjects of theory-building in studies of musical performance assessment: thus, the field expresses a degree of knower-blindness. This research responds to this gap by placing explicit emphasis on conceptualising how who you are matters in musical performance assessment practices.

The third part of the chapter (Section 3.4) situated the final research question: How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised? This question recognises

the multiple purposes, paradigms, and approaches that characterise the field of assessment. The conceptual landscape associated with these aspects of assessment is a contested space, and different positions may overlap and interact in myriad ways—usually multiple interests are present at any given time. The gap identified here is a need for conceptual approaches that enable dialogue between positions, as distinct from the marking out of additional purportedly better positions. The third research question offers a means for exploring a response to this gap.

Together these three research questions provide more specific trajectories for responding to the problem-outlined in Chapter 1. As part of the conceptual framework of the research they are closely linked with—and developed alongside—the contextual positions described in the previous chapter as well as the theoretical concepts discussed in the next chapter. Chapter 4 concludes the discussion of the conceptual framework by situating and providing a detailed account of the more specific explanatory theoretical framework that was enacted in the research.

Chapter 4: Conceptual Framework (Part 3)

Legitimation Code Theory

4.1 Introduction

This is the third and final chapter to describe the conceptual framework of the thesis. The previous two chapters have mainly been about establishing the context and framing of the research. Chapter 2 introduced the conceptual framework and discussed its broader contextual aspects, including the basic theoretical points of departure for the study as well as an overview of the research setting. Chapter 3 focused more specifically on framing the research questions that guide the study and articulated the main gaps toward which this work is oriented. The role of this chapter is to discuss the more specific theoretical ideas with which this thesis engages, and to this end it describes the outcome of a search for some theoretical tools that could be productively put to work in response to the research aims (Chapter 1). These theoretical ideas are central to the study and are one of the main bases for discussion developed in the remainder of the thesis.

This chapter has three main parts. The first part of this chapter (Section 4.2) introduces the theoretical framework (Legitimation Code Theory) and explains its relevance to the work by establishing a frame of reference to some of the ideas developed in previous chapters. The second part of the chapter (Section 4.3) describes the specific aspects of this framework that meaningfully influenced the study. In discussing the more specific theoretical devices used in the research this chapter has a degree of overlap with the following chapter (Chapter 5), which explains the means by which the concepts discussed here were applied in practice. The final part of the chapter (Section 4.4) concludes the discussion of the conceptual framework and summarises the main points from Chapters 2–4.

4.2 Legitimation Code Theory

This thesis is a theoretical work focused on conceptual questions (Chapter 1), and in exploring these questions it is broadly about theorising aspects of musical performance assessment. Chapter 2 discussed the differences

between three kinds of theories, including substantive theories, explanatory theories, and meta-theories (Section 2.3). For the purposes of this research, an explanatory theoretical framework was sought that would provide theory-building resources for exploring the conceptual questions guiding the research. The menu of possible theoretical frameworks and devices for social practice research is fairly diverse (Schatzki, 2012). The direction taken in this research is in part the result of theoretical exploration conducted earlier in the project (see Appendix A). Early encounters with sociological theoretical frameworks described in a broader range of music education studies (see for example Wright, 2010; Burnard, Trulsson, & Söderman; 2016) suggested a range of possibilities, the most influential of which were the theoretical contributions of Bourdieu (1930–2002, see for example, Bourdieu, 1998) and Bernstein (1924–2000, see for example, Bernstein, 2000). Reviewing scholarly discussion of the concepts developed by Bourdieu and Bernstein led to the discovery of Legitimation Code Theory (LCT), which was ultimately selected for inclusion in this work. The adoption of LCT into the conceptual framework over the first year of the study meant that it became an important influence in defining the research as it unfolded—as with all components of the conceptual framework, it remained in dialogue with other aspects of the framework throughout the study. Throughout the research a sustained engagement with the landscape of practice theories (including an ongoing dialogue with scholarly peers interested in social theories) has highlighted a range of other possible frameworks and concepts which are beyond the scope of this study to explore. Some examples include Actor Network Theory (see for example, Bearman & Ajjawi, 2018) and Systemic Functional Linguistics (see for example, Martin, 2013, 2016). Some possibilities for engagement with these frameworks in future studies are noted in Chapter 9, and further personal detail explaining the discovery and adoption of LCT into the study is provided in Appendix A.

4.2.1 Introduction to LCT

LCT is a realist sociological theoretical framework that builds primarily on the theoretical contributions of Bernstein (for example, Bernstein, 1975, 2000) and Bourdieu (for example, 1998, see also Grenfell, 2008). The primary architect

of LCT is Maton, who proposed an early version of the framework in 2004 (see Maton, 2004). The set of major publications that detail current forms of the framework's constituents includes Maton (2014), Maton, Hood, and Shay (2016), and Maton and Howard (2018). To introduce the framework, a useful place to start is by explaining the terms that make up its title. Beginning from the end, *theory* refers to the explanatory nature of the concepts on offer—LCT is an explanatory theoretical framework in the sense of the description provided earlier in this chapter. *Legitimation* refers to the idea that the ways in which people, or *actors*, think and act—broadly, their *practices*—embody positions or claims about what is legitimate in a given context. In this sense, the basis or bases of legitimacy in a given context serve to regulate the ways in which desirable outcomes may be achieved. In LCT, the practices of actors are conceptualised as languages of legitimation (Chapter 2, see Section 2.3.3) in the sense that they reflect "messages concerning the nature of achievement" (Maton, 2016b, p. 240) in social fields. *Code* refers to the idea that there are particular organising principles underlying "practices, dispositions and contexts" (Maton, 2016b, p. 240). It can be likened to a corporate code of practice in the sense that a given code describes legitimate ways of acting and/or being. Its origins include the *code theory* of Bernstein (see for example, Bernstein, 2000), many ideas from which are subsumed into LCT. The kinds of codes theorised by LCT are *legitimation codes*, which "conceptualize organizing principles of practices, dispositions and contexts" (Maton, 2016b, p. 240). The main premise of legitimation codes is that they are concepts that "explore the basis of legitimacy and thus what practices, dispositions or contexts are attempting to establish as possible and valued" (Maton & Howard, 2018, p. 7). In summary, LCT provides general concepts— in the sense that they are not-discipline specific—that can be used to make explicit and/or theorise the bases of successful participation in social practices: that is, these concepts can be used to help clarify the rules of the game.

So far, LCT has been used in various ways to explore both assessment and music in a variety of educational contexts, although the two have not yet been brought together explicitly. On the assessment side, it has been used to explore practices in educational contexts that include chemistry (Rootman-le Grange & Blackie, 2018), biology (Antia & Kamai, 2017), work-integrated

learning (Garraway & Reddy, 2017), feedback on students' writing (van Heerden, Clarence & Bharuthram, 2016), engineering (Wolff & Hoffman, 2014), and effects of social constructivism on assessment in higher education (Shay, 2008b). On the music side, LCT has been used in studies of school music programs in England (Lamont & Maton, 2008, 2010), higher education students' writing about jazz and the integration of notation into jazz students' written texts (Martin, 2013, 2016, 2018), and student popular musicians' informal learning (Carroll, 2017, 2019).

4.2.2 The broad structure of LCT

LCT is comprised of a range of parts which serve different functions within the framework. A distinctive feature of LCT is that the various concepts on offer can be utilised in a modular fashion as appropriate to the given research context or other problem-situation. It is not necessary to use all aspects of the framework—indeed, most studies using LCT adopt only a few aspects of the framework at a time. The organisation of LCT is such that the framework is divided into multiple *dimensions* which bring together resources for exploring different kinds of legitimation codes. An important feature of the dimensions is the narrowness of their focus on specific aspects of social practices. LCT does not aim to conceptualise the abstract totality of social practices, but rather, focuses on individual elements through the dimensional organisational approach. A useful conceptual analogy for enacting the dimensions likens the framework to a suite of medical imaging machines, such as may be found in a hospital theatre (K. Maton, personal communication, February 14, 2020). The different machines each provide some information about the patient. The nature of the information required, or the problem, will dictate which machine or group of machines is required. Some machines may provide multiple kinds of information about the patient—depending on the situation it may be necessary to utilise only some of this information. The kind of information produced by the machines usually requires some form of interpretation (even if this is only basic): Since LCT is an interpretive framework, translation between the internal language of the theory and data is necessary to make use of the concepts (Chapter 2). Further, the conceptual tools on offer are not necessarily equally applicable to a given problem or data, and the use of one

concept does not presuppose the application of another. Rather, analysis and interpretation are guided by what the problem-situation and data available require or enable (Maton, 2016a).

There are currently five dimensions to LCT. Since the publication of Maton (2004), the dimensions of Specialization, Semantics, and Autonomy have been substantially developed and enacted in a range of research studies. Temporality and Density have been explored less, and will likely undergo revision in the coming years.

Table 4.1 Overview of LCT dimensions, adapted from Maton (2016b)

Dimension	Description
Specialization[9]	explores practices in terms of *knowledge–knower structures* whose organizing principles are given by *specialization codes* that comprise strengths of *epistemic relations* and *social relations*. These are mapped on the *specialization plane* and traced over time on *specialization profiles*
Semantics[10]	explores practices in terms of their *semantic structures* whose organizing principles are given by *semantic codes* that comprise strengths of *semantic gravity* and *semantic density*. These are mapped on the *semantic plane* and traced over time on semantic profiles
Autonomy	explores practice in terms of *relatively autonomous social universes* whose organizing principles are given by *autonomy codes* that comprise relative strengths of *positional autonomy* (PA) and *relational autonomy* (RA). These are mapped on the *autonomy plane* and traced over time on *autonomy profiles*
Temporality	a dimension of LCT that explores practices in terms of their temporal features—currently this dimension has received little attention, though it has been developed further than Density (below)
Density	A dimension of LCT that has received relatively little attention as yet. Likely to be renamed when developed further within a major research study to avoid confusion with semantic density

---

[9] The spelling of Specialization with a "z" is a convention of LCT.

[10] The term Semantics (in its capitalised form) takes on particular meaning in the context of LCT, and its usage is distinct from linguistic interpretations where the concern is with meaning itself. Wilmot (2018) summarises: "Semantics in the LCT sense is specifically concerned with the context-dependence (semantic gravity) and complexity (semantic density) of meaning, not meaning itself" (p. 39).

4.2.3 General characteristics of LCT dimensions

There are some characteristics common to all LCT dimensions. Firstly, all dimensions comprise "a series of concepts centred on capturing a set of organizing principles underlying dispositions, practices and contexts" (Maton, 2016a, p. 11). The different sets of organising principles represent different kinds of *legitimation codes*, which are named for their respective dimension: Specialization generates specialization codes, Semantics generates semantic codes, and so forth. Each dimension includes a set of constitutive relations, portrayed together on Cartesian planes, each of which combines the typology of codes generated by the dimension with the topological[11] space generated by the two continua (see Figure 4.1).
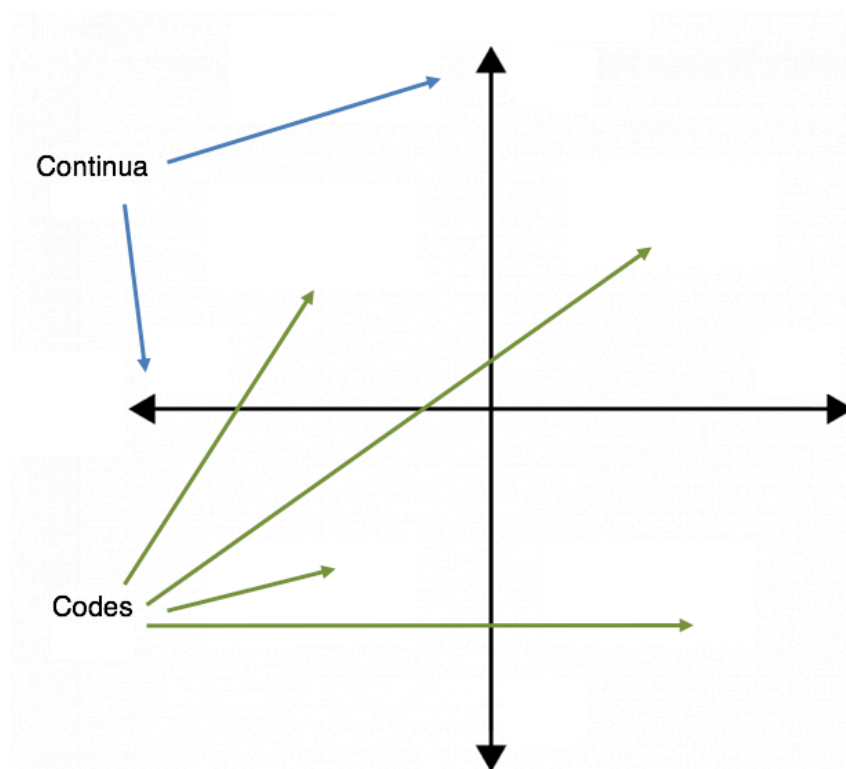
Figure 4.1 An illustration of how Cartesian planes are interpreted in LCT

---

[11] The concept of *topology* as it is used in this thesis reflects the general usage of the term to refer to the arrangement of parts of a space or object. This is distinct from other technical uses of the term—for example, the sub-discipline of mathematics that goes by the same name.

Importantly, codes are not mutually exclusive: They may *clash*, *match*, or *shift* to various degrees (Maton, 2014, 2016a). Maton and Chen (in press) explain:

> A specific code may dominate as the basis of achievement, but may not be transparent, universal or uncontested. Not everyone may recognize and/or be able to realize what is required, there may be more than one code present, and there are likely to be struggles among actors over which code is dominant. (in press, p. 4)

Maton and Chen (in press) further explain that shifts in which code is dominant "effectively change the 'rules of the game'" (p. 4), and can render "previously successful actors unable to continue to succeed" (in press, p. 5).

4.3 Concepts From LCT Used in This Thesis

This section explains the characteristics of the main parts of LCT that are employed in this project. There are four main aspects of the framework that are drawn upon, including three of the five dimensions (Specialization, Semantics, and Autonomy), and one set of sub-dimensions (the 4-K model, within Specialization). Some other auxiliary concepts will also be used throughout the thesis, however these are less substantial in scope than those discussed here and will be introduced as they arise. An important point to re-emphasise here is that while this section provides an account of the various components of LCT enacted in this thesis, the manner and degree of their enactment across the project differs substantially—in one sense, this section describes the theoretical menu from which approaches were selected at various points throughout the project. Each of the concepts discussed in this section is accompanied by some short examples to help clarify the mechanics of the concepts in the context of this study. This approach somewhat reflects the descriptive style of the main publications from which this section is synthesised (Maton, 2014, 2020; Maton, Hood & Shay, 2016; Maton & Howard, 2018), and the examples given are at times fairly direct re-formulations of those given in these publications.

4.3.1 Specialization

The first of the three LCT dimensions drawn on in this work is Specialization (denoted by the capitalised form of the word). The conceptual point of departure for the Specialization dimension is that "practices are about or oriented towards something and by someone" (Maton, 2016a, p. 11). It can be used to explore the bases on which particular knowledge and ways of knowing are valorised or downplayed within given domains of practice. Specialization develops Bernstein's concepts of horizontal and vertical discourse[12] (explained in Bernstein, 2000) into two theoretical relationships:

- *epistemic relations* [ER] between practices and their object or focus (that part of the world towards which they are oriented); and
- *social relations* [SR] between practices and their subject, author or actor (who is enacting the practices) (List adapted from Maton, 2014, p. 29, italics original).

In other words, "[t]hese relations highlight questions of: what can be legitimately described as knowledge (epistemic relations); and who can claim to be a legitimate knower (social relations)" (Maton, 2014, p. 29). Both epistemic relations and social relations are conceptualised on respective continua of strength—that is, the degree to which they are strongly or weakly emphasised, which taken together are visually depicted by the Specialization plane (Figure 4.2). The use of continua to depict sets of relations in this way is a feature of all dimensions of LCT, and degrees of emphasis, or strength, are broadly signified by plus (+) and minus (-) symbols: for example, stronger epistemic relations would be abbreviated as ER+.

Although subsequent chapters provide more detailed explanations of how these concepts are realised empirically, some simple examples help to clarify the mechanics of these concepts. For epistemic relations, a useful example anticipates Chapter 7: In simple terms, the interest is in whether successful participation in assessment situations is more (ER+) or less (ER-) dependent upon the possession of particular knowledge and skills in a given

---

[12] Specialization also subsumes the Bernsteinian concepts of *classification* and *framing* (see Bernstein 1975 and 2000 for an overview of these concepts)

context, which could include both musical kinds—for example, abilities in technical execution, or idiomatic interpretation—and assessment knowledge of the kind that is associated with assessment literacy (Chapter 2). For social relations, the interest is in whether successful participation in assessment situations is more (SR+) or less (SR-) dependent upon the possession of particular traits, dispositions, or relationships: for example, to what degree gender influences participants' abilities to achieve successful outcomes.

This is a useful point at which to highlight that in working with concepts from LCT the main interest is in theorising the *bases* of achievement in social practice. In Specialization these are given by the relative emphases accorded epistemic relations and social relations in given contexts. In LCT, *basis* is distinct from *focus* in the sense that the latter determines the object of study (Maton, 2014). This is an important distinction for the reason that objects of study can be positioned using the same theoretical terminology. For example, Chapter 6 of this thesis is about assessment participants, which can be analytically construed as a focus on social relations. The interest in the chapter, however, is in the bases by which assessment participants are able to succeed in assessment practices, including whether this is by possession of knowledge and skills (a basis in epistemic relations) or particular social characteristics (a basis in social relations).

Together, the continua of epistemic relations and social relations generate four general modalities, or *Specialization codes*, visually depicted by the Specialization plane (Figure 4.2). Inherent in each code are particular characteristics produced by the relative strengths and weaknesses of the epistemic and social relations. These can be summarised as:

- *knowledge codes* (ER+, SR-), where possession of specialized knowledge, principles or procedures concerning specific objects of study is emphasized and the attributes of actors are downplayed;
- *knower codes* (ER-, SR+), where specialized knowledge and objects are downplayed and the attributes of actors are emphasized as measures of achievement, whether viewed as born (e.g. 'natural talent'), cultivated (e.g. 'taste') or social (e.g. feminist standpoint theory);

- *élite codes* (ER+, SR+), where legitimacy is based on both possessing specialist knowledge and being the right kind of knower; and
- *relativist codes* (ER-, SR-), where legitimacy is determined by neither specialist knowledge nor knower attributes—'anything goes'. (Maton, 2016a, p. 13, italics original)
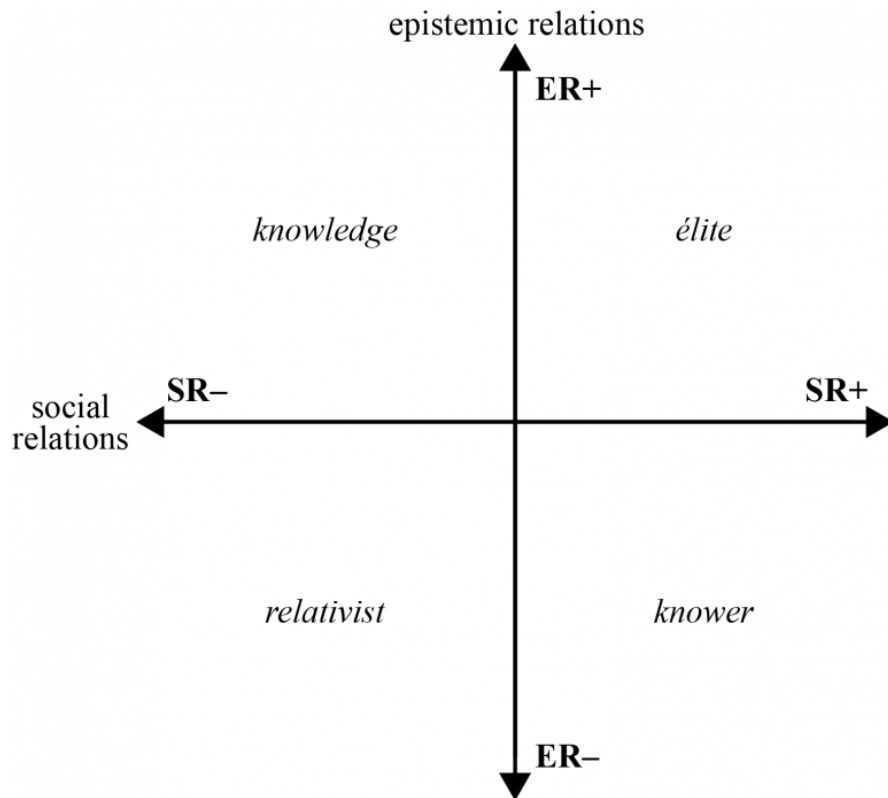
Figure 4.2 The Specialization plane (Maton, 2014, p. 30), reprinted with permission

4.3.2 Semantics

The second of the three LCT dimensions drawn on in this work is Semantics (denoted by the capitalised form of the word). Within this dimension is *semantic gravity*, which was briefly introduced in Chapter 2 (see Section 2.4). The conceptual point of departure for Semantics is a view that

> conceives of social fields of practice as *semantic structures* whose organizing principles are conceptualized as *semantic codes* that comprise *semantic gravity* and *semantic density*. Put simply, semantic gravity conceptualizes context-dependence and semantic density conceptualizes complexity. (Maton, 2020, p. 62, italics original)

A more specific definition for semantic gravity is that:

> *Semantic gravity* (SG) refers to the degree to which meaning relates to its context and may be stronger (+) or weaker (–) along a continuum of strengths. The stronger the semantic gravity (SG+), the more meaning is dependent on its context; the weaker the semantic gravity (SG–), the less dependent meaning is on its context. (Maton, 2020, p. 62, italics original)

Music education provides a useful example for clarifying the mechanics of this concept. Chapter 2 explored some of the meanings associated with this term, and some space was devoted to explaining the landscape of Australian higher music education. Using semantic gravity, the broader definitions of music education discussed in the previous chapter can be interpreted as expressing weaker semantic gravity, while the description of Australian higher music education expresses relatively stronger semantic gravity (it is more context dependent). This example highlights the relational nature of LCT as an explanatory framework: general conceptions of music education are taken as expressing weaker semantic gravity *in relation to* more context-specific versions, such as Australian higher music education. Moving to semantic density, a more specific definition for the concept is that

> *[s]emantic density* (SD) refers to the degree of condensation of meaning within sociocultural practices, whether these comprise symbols, terms, concepts, phrases, expressions, gestures, clothing, etc. (sic) Semantic density may be relatively stronger (+) or weaker (–) along a continuum of strengths. The stronger the semantic density (SD+), the more meanings are condensed within practices; the weaker the semantic density (SD–), the fewer meanings are condensed. Put another way, semantic density conceptualizes complexity: the stronger the semantic density, the more complex the practices. (Maton, 2020, p. 63, italics original)

A useful example here is the concept of assessment. Firstly, the contextual specificity of the concept could be described using semantic gravity: Formal musical performance assessment in higher education contexts is a more strongly contextualised concept than assessment in higher education. Semantic density, by comparison, reflects the degree of condensation of concepts within the notion of assessment. For example, assessment is often thought of by staff and students in educational communities as simply an unavoidable aspect of education that is sometimes stressful and sometimes useful. In the field of scholarship about assessment, however, the term is more explicitly related to a myriad of other concepts—standards, criteria, authenticity, reliability, validity, subjectivity, objectivity, and so forth. The aim in this example is to illustrate that the range of meanings with which a term is imbued is linked to the complexity of the semantic structure in which it is embedded (Maton, 2020). In this way, semantic density can be construed in terms of relations between meanings, where "the more relations established with other meanings, the stronger the semantic density" (Maton, 2020, p. 63).
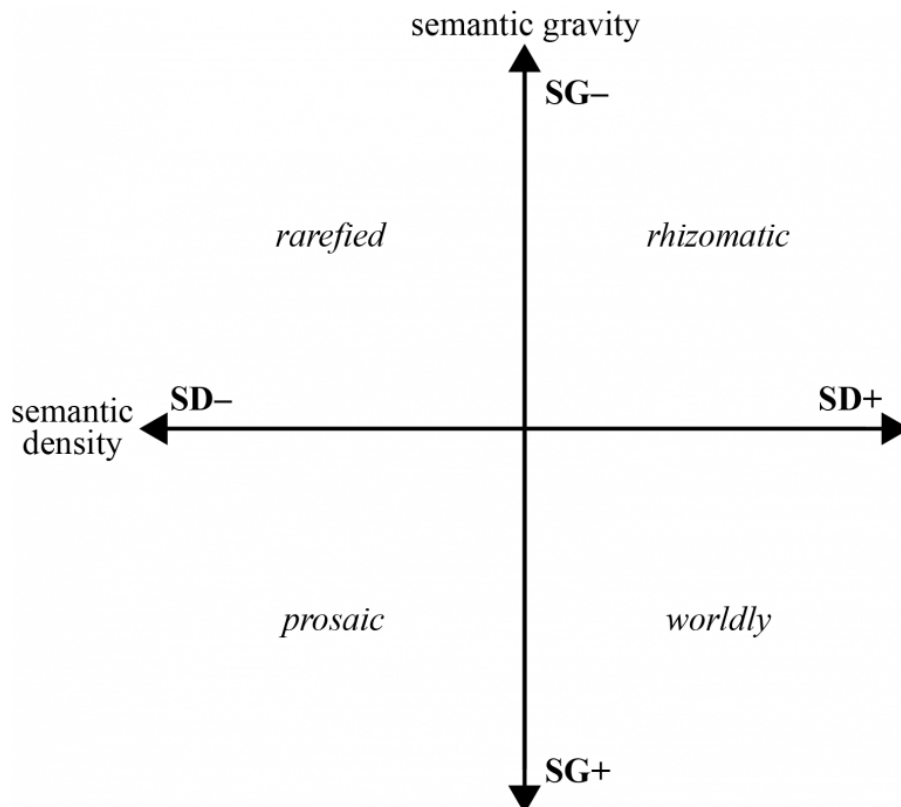
Figure 4.3 The semantic plane (Maton, 2014, p. 131), reprinted with permission

Taken together, the continua of semantic gravity and semantic density are visually depicted by the semantic plane (Figure 4.3). Like Specialization, the semantic plane generates four principal modalities, or semantic codes (Maton, 2016a, pp. 15–16):

- *rhizomatic* codes (SG–, SD+), where the basis of achievement comprises relatively context-independent and complex stances;
- *prosaic* codes (SG+, SD–), where legitimacy accrues to relatively context-dependent and simpler stances;
- *rarefied* codes (SG–, SD–), where legitimacy is based on relatively context-independent stances that condense fewer meanings; and
- *worldly* codes (SG+, SD+), where legitimacy is accorded to relatively context-dependent stances that condense manifold meanings.

4.3.3 Autonomy

The last of the three LCT dimensions drawn on in this work is Autonomy (denoted by the capitalised form of the word). Autonomy is the dimension of LCT that has received the most recent developmental attention. As with the other dimensions of LCT, the original ideas upon which it is premised were proposed over a decade ago (see Maton, 2004), and the most recent revision is detailed in Maton and Howard (2018). The conceptual point of departure for Autonomy is that

> any set of practices comprises constituents that are related together in particular ways. The constituents and the basis of how they are related together may take many forms. Constituents may be actors, ideas, artefacts, institutions, machine elements, body movements, sounds, etc.; how such constituents are related together may be based on explicit procedures, tacit conventions, mechanisms, explicitly stated aims, unstated orthodoxies, formal rules, etc. (Maton & Howard, 2018, p. 6)

Autonomy is a useful foil to Specialization and Semantics for the reason that it is focused on the integration of practices. Where the former dimensions

highlight structural features of practices "neither dimension conceptualizes whether and how those knowledge practices are being integrated" (Maton & Howard, 2018, p. 5). The function of Autonomy codes is, therefore, to "explore the boundaries that practices establish around their constituents and the boundaries they establish around how those constituents are related together" (p. 6). The degree to which constituents of practices and relations between them are bounded can be conceived of in terms of *insulation* (Maton & Howard, 2018). This premise is realised analytically in the distinction between

- *positional autonomy* (PA) between constituents positioned within a context or category and those positioned in other contexts or categories; and
- *relational autonomy* (RA) between relations among constituents of a context or category and relations among constituents of other contexts or categories. (Maton & Howard, 2018, p. 6, italics original)

As in Specialization and Semantics, positional autonomy and relational autonomy are construed on continua that realise stronger and weaker positions "where stronger represents greater insulation and weaker represents lesser insulation" (Maton & Howard, 2018, p. 6). The main positions of strength and weakness can be summarised as follows:

> Stronger positional autonomy (PA+) indicates where constituents positioned in a context or category are relatively strongly delimited from constituents attributed to other contexts or categories, and weaker positional autonomy (PA–) indicates where such distinctions are drawn relatively weakly. Stronger relational autonomy (RA+) indicates where the principles governing how constituents are related together are relatively specific to that set of practices, i.e. purposes, aims, ways of working, etc. are autonomous; and weaker relational autonomy (RA–) indicates where the principles governing how constituents are related together may be drawn from or shared with other sets of practices, i.e. purposes, aims, ways of working, etc. are heteronomous. (Maton & Howard, 2018, p. 6)

An example interpretation of these concepts anticipates Chapter 7 of this thesis, wherein it is suggested that positional autonomy can be productively used to frame the degree of insulation of constituents (concepts, ideas, people, and artefacts) considered to be *musical* from those considered to be *non-musical*, where stronger positional autonomy reflects greater insulation. Relational autonomy in this example reflects ways of working, where musical ways of working are construed as more or less insulated from other ways of working. For instance, many Australian universities advocate that the general rule of thumb for how many hours a week students should devote to higher education coursework amounts to between 10 and 15 hours per standard unit. Relational autonomy highlights that this means of allocating workload may be more or less insulated from the world of study experienced by musical performance students, where conceptions of workload may reference other cultural structures: The cultures of conservatoires, for example, may encourage students to practice for many hours a day (see Jørgensen, 2002; Perkins, 2013).

Taken together, these concepts are visually depicted by the autonomy plane (Figure 4.4), where independent variations of positional autonomy and relational autonomy generate four main autonomy codes. The four main autonomy codes can be summarised as follows:

- For *sovereign codes* (PA+, RA+) status is accorded to strongly insulated positions and autonomous principles. What is valued emanates from within the context or category and acts according to its specific ways of working: internal constituents for internal purposes.
- For *exotic codes* (PA−, RA−) legitimacy accrues to weakly insulated positions and heteronomous principles. What is valued are constituents associated with other contexts or categories and ways of working from other contexts or categories: external constituents for external purposes.
- For *introjected codes* (PA−, RA+) legitimacy resides with weakly insulated positions and autonomous principles. What is valued are constituents associated with other contexts or categories but oriented

towards ways of working emanating from within: external constituents turned to internal purposes.

- For *projected codes* (PA+, RA−) status resides with strongly insulated positions and heteronomous principles. What is valued are constituents from within that are oriented towards ways of working from elsewhere: internal constituents turned to external purposes. (Maton & Howard, 2018, p. 7, italics original)

Figure 4.4 The Autonomy plane (Maton & Howard, 2018, p. 6), reprinted with permission

4.3.4 The 4-K Model

The final theoretical device discussed in this chapter, *the 4-K Model*, is not a primary dimension of LCT, but rather, sits within the Specialization dimension. It is a set of concepts which have their basis in epistemic relations and social relations (Section 4.3.1) and which provides some additional theoretical tools. Functionally, the 4-K Model is structured in a similar manner to the primary dimensions of LCT. The main points of departure for these sets of relations are summarised as follows (illustrated in Figure 4.5):

[T]he concept of *epistemic relations* (ER) highlights that practices may be specialized by both *what* they relate to and *how* they so relate, or by relations to the objects of their focus and to other possible practices. One can thereby analytically distinguish *ontic relations* (OR) between practices and that part of the world towards which they are oriented, and *discursive relations* (DR) between practices and other practices. (Maton, 2014, p. 175, italics original)

[T]he concept of *social relations* highlights that practices may be specialized by knowers in terms of both who they are (such as social categories) and how they know (such as cultivation), or kinds of knowers and ways of knowing. One can thereby analytically distinguish: *subjective relations* (SubR) between practices and the kinds of actors engaged in them; and *interactional relations* (IR) between practices and the ways of acting involved. (Maton, 2014, p. 184, italics original)
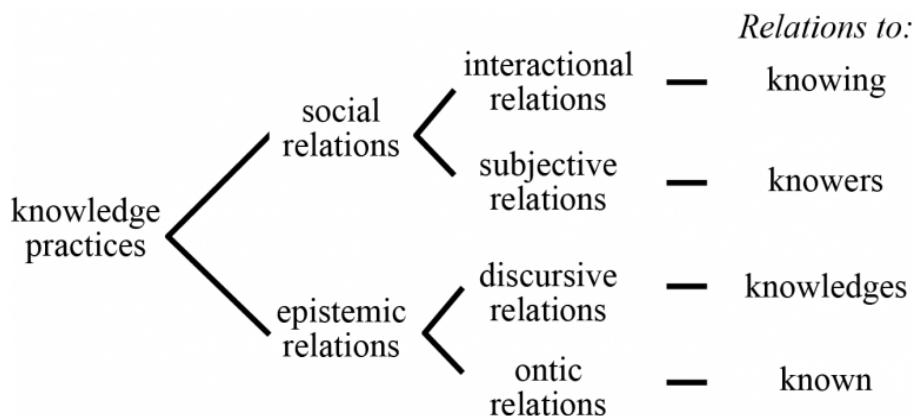
Figure 4.5 Relations comprising the 4–K Model (Maton, 2014, p. 193), reprinted with permission

It is worth clarifying before proceeding further that these subsets of relations do not necessarily reflect the totality of their parent concepts: that is, social relations are not, by definition, the conceptual sum of interactional relations and subjective relations. As with the main dimensions of LCT, these sets of relations are taken together to generate topologies, visualised on Cartesian planes (illustrated below). These spaces are, similarly, not necessarily to be

taken as geographical subdivisions of the Specialization plane: rather, they simply offer some finer-grained means of unpacking organising principles within epistemic relations and social relations.

Beginning with the derivations of epistemic relations, a simple example helps to highlight the distinction between ontic relations and discursive relations. In the context of assessment, these concepts can be used to highlight the distinction between what is assessed (ontic relations) and how assessment is conducted (discursive relations). Stronger ontic relations (OR+) are given by an emphasis on questions of what: for example, we may be interested in the importance accorded the question of what to assess in a given context, which is distinct from the question of how to assess (discursive relations). Taken together, ontic relations and discursive relations generate four modalities of epistemic relations, referred to as *insights* and visually depicted by the *epistemic plane* (Figure 4.6).
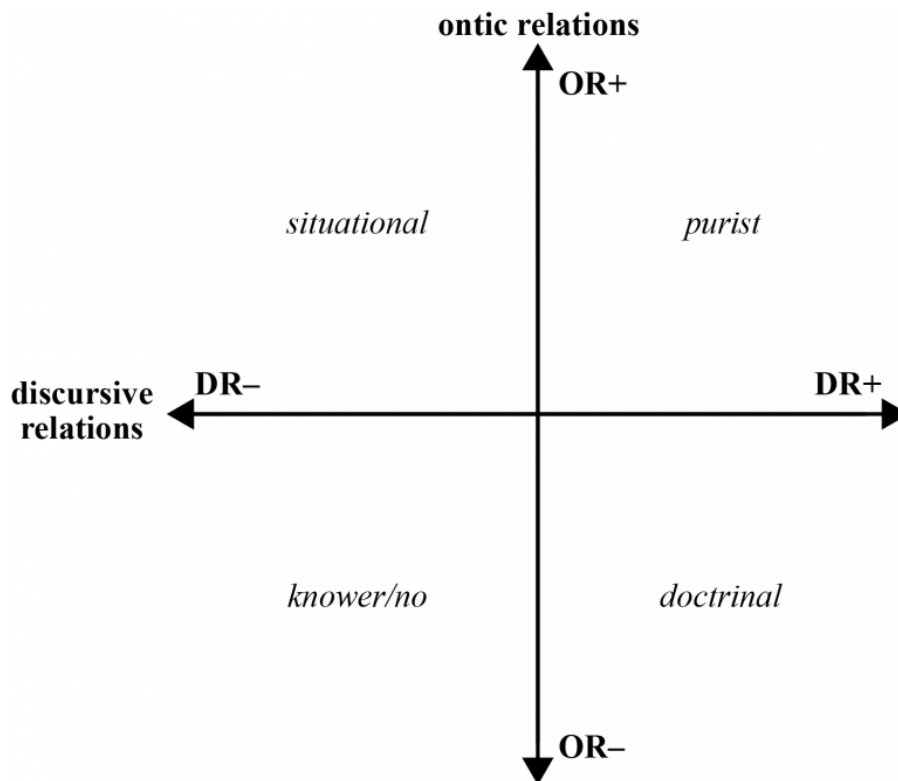
Figure 4.6 The epistemic plane (Maton, 2014, p. 177), reprinted with permission

Maton (2014) summarises these insights as follows:

- Practices characterized by *situational insight* relatively strongly bound and control their legitimate objects of study but relatively weakly bound and control legitimate approaches for constructing those problem-situations (OR+, DR−). Simply put, *what* one is studying matters but not *how*. Knowledge practices are thus specialized by their problem-situations, which may be addressed through a range of approaches: procedural pluralism or, at its weakest possible strength of DR, procedural relativism.

- Where practices emphasize *doctrinal insight*, legitimate problem-situations are not restrictively defined but relations between the legitimate approach and other possible approaches are relatively strongly bounded and controlled (OR−, DR+). Legitimacy flows from using the specializing approach: *what* is studied is less significant, *how* it is studied matters. This combines theoretical or methodological dogmatism with ontic promiscuity or, at its weakest strength of OR, ontic relativism.

- Practices based on *purist insight* relatively strongly bound and control both legitimate objects of study and legitimate approaches (OR+, DR+). Legitimacy is thus conferred by *both* 'what' and 'how' – one must use a specific approach to study a specific phenomenon. Using the legitimate approach to analyse other phenomena or using other approaches to study the legitimate phenomenon are both devalorized.

- Practices with *knower* or *no insight* relatively weakly bound and control both legitimate objects of study and legitimate approaches (OR−, DR−). With different strengths of social relations, these weaker epistemic relations may form part of either a knower code (ER−, SR+), where legitimacy flows from attributes of the subject, or a relativist code (ER−, SR−), where 'anything goes', depending on the strength of social relations. It could thus be described as *k(no)wer insight*. (Maton, 2014, p. 176, italics original)

For social relations (comprised of subjective relations and interactional relations), a helpful example is illustrated by considering the question of whether musical performance teachers should be allowed to assess their own students in higher music education settings. In some Australian institutions local policies prohibit this arrangement. We may construe from this example that the basis of legitimation is not in subjective relations, which would reference the social categories to which teachers and students belong (for example, categories such as ethnicity or gender), but in interactional relations which reflect interactive relations between people and significant others (for example, through apprenticeship), given here by the teacher-student dynamic. Taken together, subjective relations and interactional relations generate four modalities of social relations, referred to as *gazes,* and visually depicted by the *social plane* (Figure 4.7).

Maton (2014) summarises these gazes as follows:

- Where legitimacy is based on knowers possessing a *social gaze*, practices relatively strongly bound and control the kinds of knowers who can claim legitimacy but relatively weakly limit their ways of knowing (SubR+, IR–). For example, standpoint theories base legitimacy on membership of a specific social category (social class, gender, ethnicity, etc.), regardless of knowers' past or present interactions.

- Practices that base legitimacy on the possession of a *cultivated gaze* weakly bound and control legitimate categories of knower but strongly bound and control legitimate interactions with significant others (SubR–, IR+). These often involve acquiring a 'feel' for practices through, for example: extended participation in 'communities of practice' (Lave & Wenger 1991); sustained exposure to exemplary models, such as great works of art; and prolonged apprenticeship under an acknowledged master.

- Practices that define legitimacy in terms of possessing a *born gaze* relatively strongly bound and control *both* legitimate kinds of knowers and legitimate ways of knowing (SubR+, IR+), such as religious beliefs of an act of God towards a chosen person or people, and claims to

legitimacy based on both membership of a social category and experiences with significant others (e.g. standpoint theory that additionally requires mentoring by already-liberated knowers in consciousness-raising groups).

- Practices that relatively weakly bound and control both legitimate kinds of knowers and legitimate ways of knowing (SubR–, IR–) are characterized by weaker social relations that, alongside different strengths of epistemic relations, may form part of either a knowledge code (ER+, SR–) underpinned by a *trained gaze* that emphasizes the possession of specialist knowledge and skills, or a relativist code (ER–, SR–) that offers a *blank gaze.* (Maton, 2014, pp. 185–186, italics original)
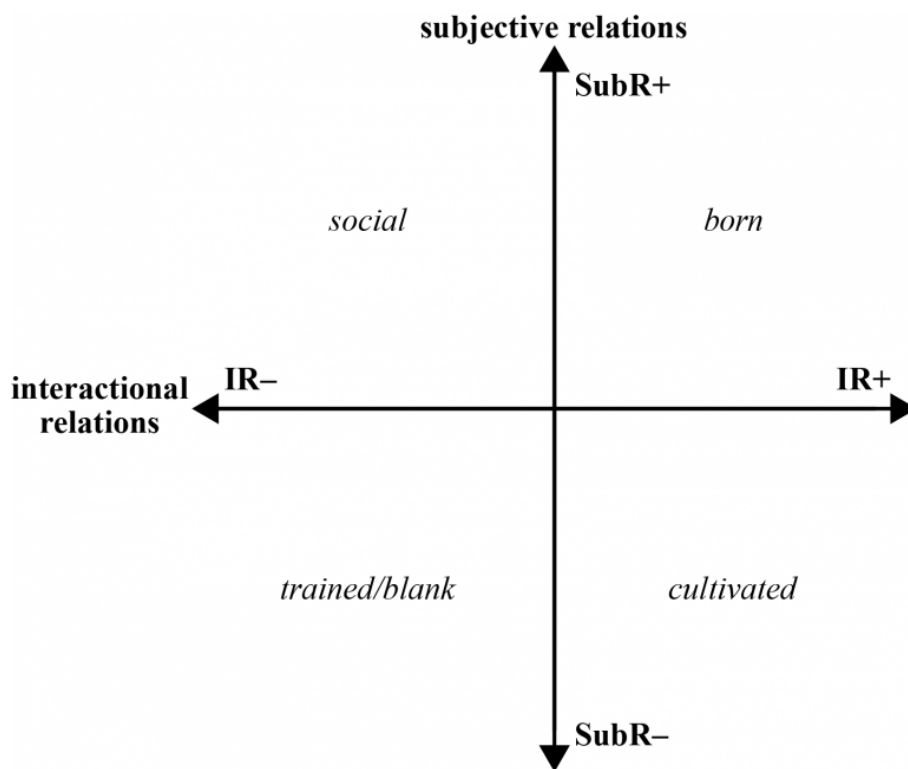
Figure 4.7 The social plane (Maton, 2014, p. 186), reprinted with permission

4.3.5 Additional concepts: Cosmologies and constellations

A general feature of LCT is that it highlights theoretical structures underlying social practices. This is particularly explicit in the Semantics dimension where the broad focus is on meaning, and the concepts provided by other dimensions can be cross-pollinated to further develop this. From the perspective of the

Specialization dimension, for example, "[d]ifferent emphases on epistemic relations or social relations are associated with different types of meaning" (Doran, in press, p. 11). Coupled with semantic density (the degree of condensation of meaning), epistemic relations and social relations produce different flavours of semantic density, reflecting the different possible bases for condensation of meaning. Maton (2014) distinguishes, therefore, between

- *epistemological condensation*, where the condensing of meanings (from other concepts or empirical referents) emphasizes epistemic relations; and
- *axiological condensation,* where the condensing of meanings (from affective, aesthetic, ethical, political and moral stances) emphasizes social relations. (Maton, 2014, p. 153)

The notion of condensation of meaning is important for the reason that it highlights that meanings are relatable to one another, and semantic density highlights that some meanings can subsume other meanings. At the same time, groups or clusters of meaning (located within ideas, perspectives, artefacts, and so forth) are able to be positioned according to systems of beliefs and values, what Maton (2014) refers to as *cosmology*[13]. Maton (2014) contrasts cosmology with Gellner's (1959) definition of *ideology*: "a system of ideas with a powerful sex appeal" (p. 2). Maton (2014) notes that Gellner's conception of ideology is applicable to "all systems of ideas or practices, including scientific theories: each has more or less 'sex appeal'" (Maton, 2014, p. 152). Cosmology, by contrast, is about the valorising or downplaying of ideas, practices, and theories. Maton (2014) provides the following explanation:

> Gellner did not propose, however, a word for that which makes
> one set of ideas and practices sexy and another not so hot. This
> I shall define as a cosmology. Cosmologies are constitutive
> features of social fields that underlie the ways actors and

---

[13] Although cosmology is not used in the astrological sense, the distinction between astrology (an axiologically-oriented field) and astronomy (an epistemologically-oriented field) is a helpful metaphor for clarifying cosmologies as values systems (see Maton, 2014, p. 170).

practices are differentially characterized and valued. A cosmology is the logic of the belief system or vision of the world embodied by activities within a social field. (Maton, 2014, p. 152)

Returning to the work of Wrigley (2005), a useful example of this concept in the context of musical performance assessment relates to what Wrigley termed "ideological resistance musically to the quantification of quality" (p. 3). Wrigley describes a particular ideological position characterised by a critique of segmented approaches for assessing musical work (see Chapter 3). The positioning of this idea is given by underlying cosmologies, which may valorise or downplay the idea that segmented assessment practices are inappropriate for assessing musical work: for example, philosophical beliefs about the authenticity inherent in a holistic artwork, or the view that the features of an artwork can be systematically broken down without harming the integrity of the appraisal. Cosmology is a helpful conceptual tool, though it is dealt with very little in this thesis in the sense of explicit discussion of the concept. It is however a relatively necessary gateway subject to the concept of *constellations*, which describe structures, or "networks" (Doran, in press, p. 11) of meaning (Maton, 2014). Constellations are produced through interactions within and between clusters of "ideas, practices, beliefs and attributes—or, for brevity, 'stances'" (Maton, 2014, p. 152). In describing the theoretical nature of constellations, Maton (2014) explains:

> These clusters may become arranged into more strongly or weakly integrated constellations that are more strongly or weakly bounded from other constellations… The nature of constellations and their arrangement are, of course, neither essential nor invariant. In astronomy a 'constellation' historically referred to a grouping of stars that make an imaginary picture in the sky. For a particular group of viewers, they may appear to have a necessary basis to their coherence but this need not be the case. For example, Pleiades is an open cluster of stars that are gravitationally bound to one another and which appear in the constellation of Taurus, though they have no substantial

astrophysical relationship to its other stars. Similarly, constellations are understood here as groupings that appear to have coherence from a particular point in space and time to actors with a particular cosmology. Different cosmologies may generate different constellations. Thus, which stances are included in a constellation, and relations within and between constellations, may vary according to different actors, change over time and be the subject of struggles. (p. 152)

In LCT, observation and discussion of constellations is enabled through the use of *signifiers*. These signifiers include "*central signifiers* around which constellations are constructed" (Maton, 2014, p. 154, italics original), and *associated signifiers*, which are more specific examples associated with central signifiers. Educational assessment, for example, could be positioned as a central signifier, and associated signifiers might include concepts such as formative and summative assessment, standards, measurement, outcomes, and so forth. To anticipate Chapters 7 and 8 of the thesis, the kinds of constellations discussed in this research are anchored by central signifiers including musical knowledge and skills and assessment knowledge and skills (Chapter 7), and traditional conservatoire education and university education (Chapter 8). Additional conceptual description of constellations is provided later in the thesis as is relevant to the given discussion.

4.4 Summary of the Conceptual Framework

This chapter was the last of three chapters to describe the conceptual framework of the thesis. Although the conceptual framework is divided across three chapters, each part of the framework functioned as an integrated part of the research, and the research itself is in many ways the product of a conceptual dialogue between these parts. Chapter 2 began the description of the framework by outlining some of the broader contextual positions informing the study, including a description of the framework itself, as well as a description of some general theoretical ideas upon which the research is premised. Chapter 2 also described the substantive context of the research and situated *musical performance assessment* and *higher music education*.

The focus in Chapter 3 was more specifically on situating the three research questions introduced in chapter 1. Chapter 3 highlighted the main conceptual gaps toward which the research questions are oriented, which included limitations of extant approaches for conceptualising musical performance assessment criteria, the foregrounding of assessment practices over assessment participants, and the need for fresh means of visualising and enabling dialogue between positions on assessment practices.

The main focus in this chapter was on describing Legitimation Code Theory, a specific theoretical framework that was adopted in the research. Section 4.2 explained the background to the selection of the framework and outlined its basic structure. LCT is modular in nature and comprised of multiple dimensions which contain concepts that can be used individually or combined. Some concepts, such as cosmologies and constellations are not contained within specific dimensions, being overlaid across the framework. Section 4.3 described the main parts of the framework that were adopted in the study, including:

- The Specialization dimension (comprised of epistemic relations and social relations)
- The Semantics dimension (comprised of semantic gravity and semantic density)
- The Autonomy dimension (comprised of positional autonomy and relational autonomy)
- The 4-K Model (comprised of subsets of epistemic relations and social relations)
- The concepts of cosmology and constellations

A useful complement to the conceptual framework is the personal memo included in Appendix A, which provides additional insight into the development of various aspects of the framework over the course of the project. Having concluded the discussion of the framework, the following chapter explains the methodological approach followed in this work, including an explanation of how LCT is operationalised in the research.

Chapter 5: Methodology

5.1 Introduction

The first part of this thesis explained the conceptual framework of the research. It described the general context of the work (Chapter 2), the more specific framing of the research questions (Chapter 3), and the theoretical points of departure selected to aid in the exploration of these questions (Chapter 4). This chapter outlines the methodological approach taken in the research. It begins by outlining the general methodological perspectives that inform the study (Section 5.2). Section 5.3 describes the way in which data are perceived, as well as the kinds of data that inform the project. Section 5.4 explains the analytical strategy of the research, and includes an explanation of how theoretical concepts from Legitimation Code Theory (LCT) were enacted in the research. Section 5.5 addresses the trustworthiness of the methodology.

5.2 Overview of the Methodological Approach

This section describes the broad ideas upon which the methodology of the project is premised. The previous chapters all have implications for the methodological approaches taken in the project. Chapter 1 introduced the broad framing of the research, highlighting the general proposition that the field of assessment is in need of new conceptual approaches (see Shay, 2008a; Boud et al., 2016) and determined an exploratory direction for the work. Chapter 2 described some of the basic theoretical positions upon which the work is predicated and provided a broader context for the work. The theoretical positions described in Chapter 2 have several important interactions with other parts of the project, including setting out the basic theoretical standpoint for exploring the research questions, determining a general logic for the adoption of LCT, and framing data analysis. Chapter 3 sharpened the lens of the research by situating the three research questions that guide the research and serve as points of departure for an explorative study of conceptual approaches for musical performance assessment. Chapter 4 described the suite of theoretical concepts from LCT that were selected for use in responding to the research questions.

Since the emphasis in this project is on exploration a qualitative research strategy was adopted (Creswell, 2007; Maxwell, 2013; Robson, 2011; Silverman, 2011). This is not to downplay the value of quantitative approaches, but rather, to acknowledge the respective strengths and weaknesses of qualitative and quantitative designs. While quantitative approaches suit an emphasis on operationalising data, there is a risk that in focusing on "how much or how many" (Merriam, 2009, p. 5) the phenomenon under observation itself may remain obscured (Silverman, 2011). The reason for employing qualitative approaches in this project relates to their emphasis on the phenomenon, rather than the measurement of variables related to it—indeed, the strength of such approaches lies in the invitation of a rich array of variables rather than the controlling of them (Holliday, 2007; Silverman, 2011).

5.2.1 Three Positions on qualitative research

This study was approached according to a number of general principles of qualitative research. The outline presented below integrates insights from several well-regarded methodological sources to form a robust foundation for the project. The general strategy reflects an emphasis on deriving meaning from qualitative data (Alasuutari, 1995; Braun & Clarke, 2013; Merriam, 2009; Robson, 2011; Silverman, 2011), and aspires to an adaptive, reflexive, and integrated approach (Braun & Clarke, 2013; Maxwell, 2009, 2013; Robson, 2011). Three general positions informing this strategy are outlined here.

5.2.1.1 Qualitative sensibility

This research adopts the premise that good qualitative research requires a certain "sensibility" (Braun & Clarke, 2013, p. 9) which involves acknowledging effective research dispositions. These include:

- A sensitivity to working with humans as research participants and the associated management of rapport (Braun & Clarke, 2013), effective communication skills (Robson, 2011), and ethics (Maxwell, 2013)
- A flexible and enquiring perspective, comfortable with ambiguities (Robson, 2011)

- An awareness of self and the willingness to step outside personal perspectives—a key approach to managing bias in qualitative research (Braun & Clarke, 2013; Merriam, 2009; Robson, 2011)

5.2.1.2 Reflexive and interactive by design

This research acknowledges that prescribing an inflexible or narrow approach can limit potential learnings (Maxwell, 2013). The approach taken recognises the value of:

- Multiple data collection methods (Maxwell, 2013; Robson, 2011)
- Reflexivity—that is, a willingness on the part of the researcher to consider developments in the research and respond accordingly, often by modifying the plan (Maxwell, 2013; Merriam, 2009; Robson, 2011)
- Interaction—that is, qualitative research involves multiple parts which do not occur in direct sequence but rather influence one another regularly (Maxwell, 2013)

5.2.1.3 Emphasis on meaning

In this research the notion of *meaning* is central (Braun & Clarke, 2013; Merriam, 2009) and involves considering several factors which influence how observations about meaning are made within the study:

- A considered stance on knowledge (epistemology) and reality (ontology) (Braun & Clarke, 2013; Merriam, 2009)
- An acknowledgement that qualitative data can be considered according to a variety of perspectives, each of which carries inherent strengths and weaknesses (Alasuutari, 1995; Silverman, 2011)
- While meaning can be made both inductively and deductively, the emphasis is on the former (Merriam, 2009)
- Subjectivity is viewed as a strength of the research design (Braun & Clarke, 2013)
- Meaning is made of data through the lens of the researcher, whose characteristics influence what is concluded (Braun & Clarke, 2006; Maton & Chen, 2016; Maxwell, 2013; Merriam, 2009)

5.2.2 Dialogic approach

As has already been mentioned in several parts of the thesis, the research is informed overall by a dialogic approach wherein the different parts of the work remained connected to one another throughout the development of the study (see Chapter 1, Section 1.3). This means that the methodological approaches discussed here both informed and were informed by interactions with the other parts of the project described in the previous chapters. To this end, the methodological design should not be construed as having been fixed in place and enacted in a linear fashion. Indeed, to follow such an approach would have been counter-intuitive to both the aims and the nature of the research (Maxwell, 2009). Similarly, the methodology is not a transformation *of* the research questions, but is rather *for* enabling the study of those questions—it is an important device in serving and enabling the other components of the project (including the research questions) to interact.

5.3 Data

This section describes the approach to data taken in the project. Firstly, Section 5.3.1 explains the perspectives on data drawn upon in the project. These perspectives reflect the positions taken in the research about the role of data and views according to which they are made sense of. Secondly, Section 5.3.2 explains the approach taken in data collection and outlines the kinds of data collected to inform the research. Thirdly, Section 5.3.3 explains the ethical considerations informing the project.

5.3.1 Perspectives on data

5.3.1.1 The role of data

Revisiting the aims of the project provides important context for the role played by data in the work, which in turn influences the kinds of data collected in the project and the means by which they are collected. At the basic level, Maxwell (2013) asserts that the kinds of data available to qualitative studies "can include virtually anything that you see, hear, or that is otherwise communicated to you while conducting the study" (p. 87), and that "[t]he development of a good data collection plan requires creativity and insight, not a mechanical

translation of your research questions into methods" (Maxwell, 2009, p. 236). This view has important implications for an explorative theoretical work, for the reason that it highlights the range of possible resources available for meaning-making activities. In this research the objective was not to strictly control the range of data drawn upon, but rather to remain responsive to and purposefully select (Robson, 2011) data which served the aims of the research. Indeed, including a range of kinds of data (discussed below) to achieve triangulation is a key strategy in enhancing the trustworthiness of qualitative research (see Section 5.5). This approach also explicitly recognises that data are not restricted to categories, but also occur as connections (Maxwell, 2009, 2013). This point is emphasised by Maxwell (2013) who cites Glaser (2001):

> All is data…what is going on in the research scene is data, whatever the source, whether interview, observations, documents, in whatever combination. It is not only what is being told and the conditions of being told, but also the data surrounding what is being told. (Glaser, 2001, p. 145, in Maxwell, 2013, p. 87)

The role of data in this project is closely related to the explorative, theory-generative aims of the research (Chapter 1). Since this research is about exploring conceptual matter, data were viewed as resources in this exploration. Although they helpfully inform on practices, and the positions evident in data in some cases have strong implications for practice in and of themselves (for example, by highlighting compliance issues), the data are specifically useful to the aims of the research in the sense of providing empirical referents to which theoretical ideas can be meaningfully related. This is important for responding to the emphasis of the research on *meaningful* conceptual approaches (Chapter 1, Sections 1.2 & 1.3). This is one rationale for limiting the data collection, and related to this is the position that this research is not intended as a case study of a tightly defined research object. The objective was not to capture a detailed perspective on assessment practices in a specific location (or locations), but rather, to *purposefully* (Maxwell, 2013; Robson, 2011) seek a broader range of heterogenous data

that provided a greater range of perspectives and possibilities. The theoretical focus of the research coupled with the richness of qualitative data meant that many possible analytical trajectories were available even for a small sample. The need to remain cogent of balancing the accrual of data with ongoing analytical activities became evident during the middle of the project (see Appendix A), and the amount of data collected was limited to what would actually be of practical use within the scope of the project.

5.3.1.2 Viewing data

In this research, the data were considered analytically according to two primary perspectives which constituted "viewpoints from which qualitative data can be approached in order to actively produce observations" (Alasuutari, 1995, p. 46). In practice, the decision to apply these viewpoints was not segregated from the other activities of the project—rather, their applications developed over time as the research unfolded. These perspectives are not mutually exclusive, allowing for the fact that data can be considered from multiple viewpoints simultaneously (Alasuutari, 1995; Silverman, 2011).

The first perspective applied to the data corresponds to what Alasuutari describes as a *factist* perspective (1995). In this perspective, the content of the data is treated as information which informs on reality. For example, in the factist perspective, curricular documentation would be held to inform on the curriculum it describes. An important feature of this viewpoint is the subsequent consideration of whether data provided by interviewees or collected from policy documents truthfully describes that which it discusses at any given point. As Atkinson and Coffey (2004) note, "we cannot treat records – however 'official' – as firm evidence of what they report" (p. 58). Considering the truth of the messages imparted by the data is very useful for the reason that it provides points of comparison across data: Depending on what other information we choose to compare a given assertion with, we may have cause to question the accuracy of the assertion. In this research the data collected includes a strong proportion of claims about different aspects of assessment— the factist perspective provides a helpful basis for interactions with these claims, emphasising that their accuracy is not without question. Importantly, the data do not provide an unfettered line of sight to the thinking of the

participants (Silverman, 2011). There are, for example, many possible influences on interviewees' responses: they might be intimidated by the topic of discussion and wish to provide an answer they believe is likely to be seen as correct, or socially acceptable. Conversely, the safety of anonymity may serve to embolden critiques of other social groups, practices, and institutions. The consideration of data from a factist perspective usefully distinguishes a factual interpretation of data from other types of insight. In particular, it corresponds with the second perspective on data informing this research, which Alasuutari (1995) refers to as the *specimen* perspective.

The purpose of the specimen perspective is to examine data on the basis of what they are, rather than what they are about. In this perspective, the data cannot lie, as it represents only itself (Alasuutari, 1995): perspectives from interviewees, for example, are treated *as* perspectives in and of themselves, rather than as a source of information about some aspect of the world beyond the interview. I refer again to Atkinson and Coffey who assert that "[d]ocumentary sources are not surrogates for other kinds of data" (2004, p. 58). Similarly, Silverman (2011) cautions against employing peoples' perspectives as explanations: "How could anybody have thought this was the case in social science?" (p. 445). Considering data for what they *are,* however, is analytically fertile. In this respect, data are useful for theory-building, in that they provide a perspective or information to be responded to which we can be relatively confident has some anchorage in the social world, and which therefore offers an important resource in developing analyses that express generalisability across contexts (though not of the kind traditionally associated with statistical quantitative research, see Section 5.5). Atkinson and Coffey (2004), for example, explain in justifying the analysis of texts that

> [t]his recognition or reservation [about the factual accuracy of data] does not mean that we should ignore or downgrade documentary data. On the contrary, our recognition of their existence as social facts (or constructions) alerts us to the necessity to treat them very seriously indeed. We have to approach documents for what they are and what they are used to accomplish. (2004, p. 58)

Together, factist and specimen perspectives offer a pair of complementary lenses for social research, highlighting that data can be usefully put to work to multiple ends simultaneously, and that their limitations and strengths are determined in part by the context and aims of the research itself.

5.3.1.3 A note about "giving voice" to data

This project is not about simply *giving voice* to the research participants and the non-human resources included in the study. Rather, it is conducted on the premise that the researcher cannot in fact achieve so basic a function (Braun & Clarke, 2006; Maxwell, 2013). This research adopts instead the view advocated for by Braun and Clarke (2006), which is that "even a 'giving voice' approach 'involves carving out unacknowledged pieces of narrative evidence that we select, edit, and deploy to border our arguments'" (Fine, 2002, p. 218, cited in Braun & Clarke, 2006, p. 80). As Maton and Chen (2016) observe: "While not always obvious to the noviciate, one's specialised gaze is always active in research" (p. 39). The correlative is that the voice of the researcher is unavoidably intermingled with the data, and with the perspectives drawn upon in other parts of the research.

5.3.2 Data collection

This research drew principally on three kinds of data to explore the research questions. All data collected in the project were purposefully collected (Maxwell, 2013; Robson, 2011). This section describes each of these kinds, which are academic literature and theories (Section 5.3.2.1), interviews (Section 5.3.2.2), and non-scholarly documents (5.3.2.3).

5.3.2.1 Literature and theory

One of the central resources drawn upon in this research is extant scholarly literature. Writing of the role of literature in research, Maxwell (2009) emphasises that it can serve a range of purposes beyond the provision of theory. Maxwell asserts, rather, that

> it can be a source of *data* that you can use to test or modify your
> theories. You can see if existing theory, the results of your pilot

research, or your experiential understanding is supported or challenged by previous studies… [Y]ou can use ideas in the literature to help you *generate* theory, rather than simply borrowing such theory from the literature. (Maxwell, 2009, p. 227)

In this research, scholarly literature—particularly literature about musical performance assessment—provides a range of useful resources that support the exploration of conceptual frameworks. Some examples include broad standpoints on assessment, propositions about ways of assessing musical performances, the distinctions between different kinds of criteria, observations about the relationship between assessment and other aspects of higher music education (including the people involved), and claims about what is preferable in assessment practices, all of which offer points of departure for conceptualising assessment practices. The idea that theory can be considered a form of data has overlap with the positioning of literature as data. A view of theory as data in this project is important for the reason that it treats theory as a resource for meaning-making beyond the provision of tools for understanding or taking positions on phenomena. Since this research is about conceptualising assessment practices, concepts themselves (for example, kinds of criteria discussed in research publications) provided information to be viewed analytically.

5.3.2.2 Interviews

In this research, semi-structured interviews were used to generate data which served as a source of insight into the ways in which a range of assessment participants perceive musical performance assessment. Although using semi-structured interviews to generate data is well-established practice in qualitative research (Robson, 2011), this research heeds Silverman's proposition (2011) that it is incumbent on the researcher to consider why a "contrived" (2011, p. 166) device such as the semi-structured interview would be used in favour of attempting to capture naturally occurring data. There are two primary reasons. Firstly, the social contexts in which data informing on assessment participants' perspectives on assessment might naturally occur are frequently difficult to

access for both logistical and ethical reasons. While naturally occurring talk in situations such as assessment committee meetings or in actual student exams would be illuminating, they would also require substantial additional ethical clearance, and it is likely that the social intrusion might influence both the data collected—thereby diminishing the value of such data as natural—and influence the processes under observation, which may themselves have significant implications for other actors. For similar reasons, naturally occurring text-based data which might provide useful insight—such as e-mail chains, or written examination reports—are difficult to access. The second reason for which I have chosen to use semi-structured interviews to generate data pertaining to actors' knowledge is that the contrived nature of the interviews does not necessarily prohibit the usefulness of the resulting data for this research. While the data collected were likely to be different from that which might occur in natural contexts, the interviews do however remain authentic as pre-meditated conversations. Indeed, the very contrivance of an interview situation may have played a role in generating useful data that could very well be justified as having naturally occurred within the interview context. Further, in placing the emphasis explicitly on participants' perspectives on assessment, the interviews might also have captured perspectives that would not be readily available in natural assessment discourse. Additionally, the kinds of positions and claims made by participants under interview conditions are useful beyond an attempt to understand actual practice by providing insight into the kinds of stances with which the participants explicitly identified. In this way, the interview data offer resources for studying both assessment practices and the perspectives of participants themselves.

Given the exploratory nature of this study a diverse group of interview participants was sought that included both students, assessors, and a small group of administrative staff from multiple Australian institutions of higher music education. For the assessors and students, it was a criterion of participation that they be actively engaged in undergraduate coursework that involved a substantial emphasis on musical performance and its assessment. Initially, a more even divide between students and staff was sought, and it was expected that student participation would be more straightforward to acquire. During the data gathering phase of the project, however, it became apparent

that higher music educators were more inclined to participate in the interviews than students. This may have been influenced by the mode of access to participants, which was through communication with heads of institutions in the first instance. Students may have been less inclined to respond to advertisements for participation from institutional leadership and administrators than staff members. While the student perspectives collected were useful data for the study, on reflection the staff perspectives provided ample material to work with within the scope of the project, and the focus could have been narrowed to include only staff. The administrative staff members were interviewed for the reason that their roles involved engaging with both students and assessors, and thus they were able to lend an additional perspective on to the project. Twenty-six participants were selected purposefully from a group of six Australian institutions (all within universities) of higher music education, with the aim being to field a heterogeneous group of perspectives (Robson, 2011). The participant group included eighteen assessors, six students, and two administrative staff members. The group included a mix of men and women and a range of ethnicities, and many participants also had international music education experience. Table 5.1 describes the basic characteristics of these participants.

Table 5.1 Description of research participants

Name	Description	Institution
Assessor A	Casually employed jazz performance teacher with some experience teaching at other institutions	1
Assessor B	Full-time leader of a classical disciplinary area and a member of institutional executive, also a performance teacher at the institution	1
Assessor C	Casually employed jazz performance teacher and postgraduate student	1
Assessor D	Full-time classical performance teacher	1
Assessor E	Casually employed jazz and contemporary/popular music performance teacher	2
Assessor F	Full-time member of institutional executive with a background in classical music, also a performance teacher at the institution	2
Assessor G	Full-time member of institutional executive with a background in classical music, also a performance teacher at the institution	2
Assessor H	Casually employed jazz performance teacher	2
Assessor I	Full-time music lecturer with responsibilities for oversight of assessment processes	2
Assessor J	Full time multi-disciplinary department leader performance teacher	2
Assessor K	Full-time leader of a disciplinary area within a classical department—a performance teacher within this area	3
Assessor L	Full-time leader of a disciplinary area within a classical department—a performance teacher within this area	3
Assessor M	Full-time leader of a disciplinary area within a classical department—a performance teacher within this area	3
Assessor N	Full-time member of institutional executive with a background in classical music	4
Assessor O	Casually employed jazz performance teacher	4
Assessor P	Full-time contemporary/popular music teacher with a background in jazz education and teaching	5
Assessor Q	Full-time member of institutional executive with a background in classical music, also a composition teacher at the institution	5
Assessor R	Full-time member of institutional executive with a background in classical music, also a performance teacher at the institution	6
Student A	Classical percussion student	1
Student B	Jazz drum kit student	1
Student C	Contemporary/popular music voice student	2
Student D	Classical piano student	4
Student E	Contemporary/popular music student	5
Student F	Classical performance student	6
Student Administrator A	Full-time student administrator at the institution	2
Student Administrator B	Full-time student administrator at the institution	4

5.3.2.3 Documents

The data collected in the research included non-scholarly documentary materials that frame performance assessment in Australian higher music education. The three main groups of documents that were considered for inclusion in the research were documents specific to individual units of coursework (for example, course outlines), documents that govern institutional assessment practices at the broader university level (for example, publicly available policies and procedures), and documents from the Australian regulatory landscape (for example, the Australian Qualifications Framework; the Higher Education Standards Framework). A pilot collection was undertaken early in the project to establish the relevance of different kinds of documents, the scope of their contents, and their accessibility. The documents used in this project collectively provided insight into a range of aspects of assessment practices in higher music education, including:

- Information about the basic design of assessment tasks
- Information about the intended purposes of assessment
- Information about the kinds of learning outcomes that assessment tasks are linked to
- Information about criteria and other guidance notes for assessment in coursework
- Information about the structure of curriculum and the location of assessment within this
- Information explaining the basic principles according to which assessment is conducted

Differences in the scope and range of the resources collected enabled a broader understanding of how performance assessment in Australian higher music education is framed and what claims are made about it—some of these findings are included in the contextual outline provided in Chapter 2 (see Sections 2.4 & 2.5). Invariably, however, the documents were usually fairly specifically focused on one particular topic, and the various documents available frequently specified different kinds of information about the same assessment system. For example, where regulatory and institutional policies

generally provide generic guidelines for the design of assessment, unit-specific documentation usually provides specific details of the assessment tasks associated with a unit of study, and the learning outcomes that they are intended to test. Collectively, however, they provide a fuller picture of what types of performance assessment are used in courses of study, what learning outcomes they test, and what principles inform their design.

Documents also differed substantially in terms of accessibility. In Australia, regulatory policies and reference materials associated with them are generally publicly available, as are the governing policies of universities. Coursework materials, by contrast, are often not publicly available. While some details about assessment in individual units of study are usually available on institutional websites, they are often brief and limited in scope, and rarely indicate which learning outcomes are associated with which assessment tasks. Helpfully, a small number of Australian providers of higher music education allow public access to authoritative documents such as course profiles (also known as course outlines or unit outlines), which are legally-binding documents that describe the curricular scope of individual units of study. In particular, these documents typically include an explanation of unit assessment tasks and the learning outcomes associated with these tasks.

5.3.3 Ethical considerations

Several measures were taken to ensure that this project was conducted in an ethical manner. Firstly, ethical approval to conduct interviews was ratified by Griffith University's Human Research Ethics Committee (reference no. 2016/865) prior to the commencement of data collection. In particular, consideration was given to the implications of collecting conversational data about educational practices and personal perspectives relating to them. The identities of all participants were obscured to avoid placing interviewees in positions that might be professionally, educationally, or otherwise socially vulnerable. In some instances, this also involved redacting other identifiers: explicit references to their instrumental specialisms, or working relationships, for example. The interview excerpts quoted in this thesis were provided to the respective interviewees prior to publication in order to ensure any concerns regarding anonymity were addressed.

Secondly, in view of the potential for data-gathering activities to offer reciprocal benefits to the research participants (Griffith University, 2016a), the interviews were conducted in ways that offered opportunities for the participants to develop their own knowledge about assessment. Primarily this was by including a conversational segment at the end of the interviews wherein participants were invited to ask any questions they might have. Many interviewees were interested in my personal perspectives on various assessment topics, and this offered an opportunity for me to respond in-kind. Several interviewees also made subsequent contact about assessment-related topics.

While the benefits to the individual participants were relatively small, this research benefits the broader community of music students, teachers, and assessors by contributing to contemporary understandings of the ways in which assessment is thought about by those who do it. In particular, the practical recommendations made in this thesis may be useful to those actively involved in performance assessment in higher music education.

5.4 Analytical Approach

This part of the chapter describes the analytical approach employed in the research. The approach described here is primarily informed by the work of Maxwell (2009, 2012, 2013), Robson (2011), Braun and Clarke (2006, 2013), and Maton and Chen (Chen, 2010; Maton & Chen, 2016, in press). The first section (5.4.1) provides an overview of the analytical approach and identifies its key components. The second section (5.4.2) describes the approach taken in coding the data, while the third section (5.4.3) describes strategies for connecting meaning across data. Section 5.4.4 introduces *translation devices* to explain the means by which concepts from LCT are analytically linked to data, and Section 5.4.5 provides some additional details related to software used in the research as well as the role of quotations in the thesis.

5.4.1 Overview of the analytical approach

This research involved a considered approach to analysis that involved thinking carefully about what analysis actually refers to and what its purpose is. Maxwell (2009) describes data analysis as what is done with information "to

make sense of it" (p. 234). In realist research (see Chapter 2, Section 2.3), Robson (2011) tells us that "seeking an answer to the question… 'How can I understand what is going on here?' (in terms of 'What mechanisms are operating?') is central (p. 409). A hallmark of analysis in qualitative research is that it produces products and insights along the way which may inform further work (Robson, 2011), and hence both Robson and Maxwell emphasise a view of analysis as a central, integrated research activity that takes place as the research unfolds (Maxwell, 2009, 2013; Robson 2011). Maxwell (2009) explains:

> Analysis is often conceptually separated from design, especially by writers who see design as what happens *before* the data are actually collected. Here, I treat analysis as a part of design (Coffey & Atkinson, 1996, p. 6), and as something that must itself be designed. Every qualitative study requires decisions about how the analysis will be done, and these decisions should influence, and be influenced by, the rest of the design. A basic principle of qualitative research is that data analysis should be conducted simultaneously with data collection (Coffey & Atkinson, 1996, p. 2). This allows you to progressively focus your interviews and observations, and to decide how to test your emerging conclusions. (Maxwell, 2009, p. 236, italics original)

Analysis is also closely linked to the concept of *interpretation.* One view is that "[t]he process and products of analysis provide the bases for interpretation" (Robson, 2011, p. 408), however Robson (2011) questions the distinction between the two. He tells us:

> The traditional, and still widely used, terminology is to refer to 'analysis' of data, whether quantitative or qualitative. Taken literally, analysis is a 'breaking up' of something complex into smaller parts and explaining the whole in terms of the properties of, and relations between, these parts. Not only is this, necessarily, a reductionist process but it is also seen by many as necessarily reliant on the particular form of statistical reasoning

where hypotheses are based on probability theory applied to sampling distributions… The major research traditions in flexible design research [Robson's suggested successor to the term *qualitative research*] are incompatible with the approach. Interpretation carries very different conceptual baggage. Whereas the purpose of analysis is often seen as a search for causes… interpretation is considered to be about shedding light on meaning. (p. 412)

In many ways the aims of this research are firmly anchored in interpretation—the research questions introduced in Chapter 1 are all focused on interpreting different aspects of musical performance assessment practices. Analytical strategies such as coding wherein the data are segmented are nonetheless a very useful means of exploring the contents of data in the sense of developing a nuanced view of the various themes which are present, and for revealing themes that might otherwise be obscured through the premature application of theoretical concepts (Maton & Chen, 2016). A solution to the kind of division between analysis and interpretation described by Robson (2011) is to deploy a research strategy that involves multiple approaches for making sense of data, including approaches that are less reductionist in nature. To this end, Maxwell (2009, 2013) identifies three main kinds of qualitative research strategy which provide different kinds of insights:

- *Categorising strategies*, which involve allocating data to categories which can be referred to as codes or themes
- *Connecting strategies*, which focus on the connections between data
- *Memos and displays*, which provide a means of visualising data, analyses, and interpretations, to keep record of insights and enable data to be seen in different ways

This research draws upon all three kinds, to varying degrees, however the focus here is on the first two[14]. The distinction between these two strategies

---

[14] Examples of memos and displays are included in different parts of the thesis. Appendix A contains an extended memo, and data are displayed in a range of ways in later chapters.

reflects what Maxwell (2013) describes as "two different modes of relationship: *similarity* and *contiguity*" (p. 106, italics added). Maxwell (2013) explains:

> Similarity relations involve resemblances or common features; their identification is based on comparison, which can be independent of time and place. In qualitative data analysis, similarities and differences are generally used to define categories and to group and compare data by category [categorising strategies]… Contiguity relations, in contrast, involve juxtaposition in time and space, the influence of one thing on another, or relations among parts of a text; their identification involves seeing actual *connections* between things, rather than similarities and differences. Contiguity relationships may also be identified among abstract concepts and categories, as a subsequent step to categorizing analysis of data [connecting strategies]… Neither of these strategies can be assimilated to the other; they are based in different forms of relationship in your data, although it is possible to combine the two strategies. (p. 106)

Since this research has explorative and generative theoretical aims, these different means for analysing and interpreting data provided a complementary frame for a multifaceted approach for making sense of the data. The main features of this approach are described here.

5.4.2 Coding

To quote Maxwell (2013), "[c]oding is a typical categorizing strategy in qualitative research" (p. 106). In the context of this research, coding is

> different from coding in quantitative research… the goal of coding is not to produce counts of things but to 'fracture' (Strauss, 1987, p. 29) the data and rearrange it into categories

---

Displays of various kinds were also used throughout the project to help visualise and interpret data in different ways.

that facilitate comparison between things in the same category and between categories. (Maxwell, 2009, pp. 236–237)

There are a range of well established (and conceptually overlapping) approaches for coding qualitative data, including strategies from grounded theory (see for example Corbin & Strauss, 2008) and interpretive phenomenology (Robson, 2011), amongst others. The approach taken in this research is most strongly inspired by Braun and Clarke's (2006) well known description of thematic analysis, Maxwell's (2013) discussion of different kinds of codes, and the approach described by Maton and Chen (2016), who used thematic analysis in a study that also incorporated LCT. Guiding this aspect of the analysis was a set of distinctions proffered by Maxwell (2013) who defines three main types of category:

- *Organisational categories*, which are "broad areas or issues" (Maxwell, 2013, p. 107) that are often realised prior to data analysis (for example, from earlier research or study): for example, authenticity, learning, criteria

- *Substantive categories*, which are descriptive, and more explicitly reflect the content of data: for example, frustrations with assessment criteria, value for technical skills, teachers positioned as legitimate assessors

- *Theoretical categories*, which are more general or abstract categories adopted from a separate framework or developed through the research, typically representing "the researcher's concepts… rather than denoting participants' concepts" (Maxwell, 2013, p. 108): concepts from Legitimation Code Theory (for example, stronger/weaker epistemic relations) fall into this kind of category

The researcher's analytical gaze is unavoidably coloured by prior knowledge (for example, of concepts or theories, or personal experiences), and while this is a strength and important influence on the trustworthiness of qualitative research, it also needed to be accounted for in allowing insights to emerge from the data as free from pre-established organisational or theoretical categories as possible (Maton & Chen, 2016). To this end, early analyses

involved a very close open coding of the data using substantive categories produced through reading the data and developing fine-grained categories "based on what data… seem most important" (Maxwell, 2013, p. 107, see also Corbin & Strauss, 2007). Maton and Chen (Chen, 2010; Maton & Chen, 2016) refer to this process as *substantive coding*.

Initial iterations of substantive coding of interview data produced hundreds of very fine-grained descriptive themes. Some examples of these types of codes are included in Table 5.2. After coding several interviews, these themes were condensed to form a smaller number, and this process was repeated many times throughout the research. As themes became more condensed and less descriptive they became more reflective of organisational categories (see Table 5.2). Some of these categories were clearly influenced in their terminology and conception by familiarity with concepts from outside the interview data (for example, assessment concepts from the literature on assessment). While this process produced many interesting insights and a broad range of possibilities for topics to focus upon, this was also a challenging point in the research. A fundamental issue that emerged was the question of what to actually *do* with the insights that were being produced through analysis. Although many interesting empirical observations were available it remained difficult to connect these to the theoretical aims of the thesis, which placed the focus on synthesising conceptual approaches for assessment rather than re-conceptualising practices in specific contexts. It became clear part way through the project, amid repeated attempts to make sense of the data, that the aims of the research needed to be tightened to enable a more focused analysis. As analytical trajectories and dead-ends were routinely encountered it became clear that a recurring tendency to become overly focused upon substantive themes, as well as themes that were locked into the context of assessment (for example, formative assessment, summative assessment, authentic assessment, and so forth) was in fact missing the broader aim of the research. In order to progress, the more specific and/or substantive assessment concepts that emerged needed to be escaped temporarily to avoid elaborate re-description. This prompted a search within the literature on assessment for broader organisational themes which could be used to focus the study—this was inspired by Chen's approach (see Maton &

Chen, 2016), wherein her data were organised according to a set of three theoretical categories derived from an existing framework. One of the issues here was that the broad range of concepts from the field of educational assessment offered too many possibilities for useful analysis, and further, that there seemed to be little means of linking these concepts together (see Appendix A). A partial solution to this issue emerged from the work of Broadfoot (2012) and Lebler and Harrison (2017) who both describe key questions for assessment practices. These questions were effective in moving beyond the more specific categories that had been defined and provided a means of tightening the focus of the analysis. They also bore a closer relationship to the concepts from LCT, which provides concepts for realising the questions independent of their relationship to assessment. This process was accompanied by a number of revisions to the research questions, and I have since described it as the *unriddling* phase of the research. A fuller account of this unriddling is included in Appendix A. An overview of the coding structure developed through this process is provided in Table 5.2. What is important about this set of codes in the context of this work is not that it reflects practice in a specific assessment context, but that it provides multiple ways of *viewing* assessment practice.

Although Table 5.2 depicts a coherent set of categorical constituents, developing useful distinctions between these took many repeated iterations of analysis. One question here is why concepts such as authenticity, learning, progress, and so forth, are allocated to the organisational categories column rather than the theoretical categories column. The reasoning behind this is most easily explained by introducing an additional concept. Another way to visualise this set of categories is using semantic gravity, which was introduced in Chapter 2 (see Section 2.4). In view of the differences between the kinds of data collected in this research, this became in important way of conceptualising their relationships to one another. The basic premise of semantic gravity is that meaning can be conceptualised as more (stronger semantic gravity) or less (weaker semantic gravity) tethered to a context.

Table 5.2 Categorical structure for data coding

More theoretical categories			More substantive categories
Specialization	Design	Assessment content	Assessment can be informal
Stronger social relations	Participants	Assessment format	Assessment of musical performance can take
Weaker social relations	Purpose	Assessors' learning	place in different contexts
Stronger epistemic		Assessors	Assessment of musical performance might be
relations		Challenges	"thought" in a different way outside of higher
Weaker epistemic		Change	education
relations		Credentials	Assessment needs to take place in order for
4-K Model		Criteria	students to receive a quantifiable grade
Stronger subjective		Culture	Everybody in higher education needs to engage
relations		Emotions	with assessment in some way
Weaker subjective		Feedback	Intergenerational performance has declined
relations		Industry	Students have musical performance lives
Stronger interactional		Institutions	outside of institutional study
relations		Interactions	Actors can be expected to downplay
Weaker interactional		International perspectives	assessment
relations		Motivation	Culture can influence assessment
Autonomy		Music education	High enrolments in music education can
Stronger positional		Outcomes	influence its worth
autonomy		Policies	Assessment is most often poorly implemented
Weaker positional		Progress	Assessment commands value in the world
autonomy		The real world	Assessment as a practice can be philosophically
Stronger relational		Standards	related to musical performance
autonomy		Students' learning	Emphasis on formal qualifications as criteria for
Weaker relational		Students	employment in higher music education can
autonomy		Subjectivity	downplay the value of educators' holistic selves
		Trustworthiness	A primary function of assessment is to provide
			feedback
			Assessment can serve some functions more
			than others
			Communication can take place between actors
			about assessment
			Music can be easy to assess
			Technical work can be separated conceptually
			from other performance work
			Degree of integration between coursework and
			external musical pursuits can influence the
			educational experience
			Students are knowledgeable about navigating
			assessment
			Assessment should represent real-world
			situations
			Different faculties may take different approaches
			to managing assessment
			Music is hard to assess

Note 1. The columns of this table are vertical lists, there is no horizontal correlation between individual units of text. Several other tables in the thesis also share this format (see Tables 6.1, 6.3, 6.4, 8.1).

Note 2. The elements listed closer to the left side of the table are more theoretical in the sense that they are further abstracted from the data; those closer to the right side of the table are more substantive in the sense that they are less abstracted from the data—they are more descriptive than theoretical.

Hence, substantive categories express the strongest semantic gravity, in that they are both closer to the data (given by their more specific description), as well as to the reality described by data (given by their description of more specific foci). The theoretical categories, by contrast, express the weakest semantic gravity, in that they are the least bound to data

Table 5.3 Depiction of semantic gravity for analytical categories

Continuum of semantic gravity	Strengths of semantic gravity	Description	Example codes
SG- ↕ SG+	SG--	Codes most distant from data	Specialization Autonomy 4-K Model Design
	SG-	Codes somewhat distant from data	Participants Purpose Authenticity Trustworthiness
	SG+	Codes somewhat closer to data	Criteria Policies Culture can influence assessment
	SG++	Codes closest to data	Assessing music is hard Assessment should represent real-world situations

5.4.2.1 A note on the question of what to code

Not all data collected in this project were formally coded to the extent described above, since the aim of coding the data was ultimately to provide an interpretive framework. As the broader framework crystallised it became possible to quickly ascertain links between the categories within it and data within a given artefact (document, interview transcript, or piece of academic literature). This proved to be an important resource in identifying broader relationships both within and across data, and provided a means of connecting them.

5.4.3 Connecting data

To complement the categorical analysis described above, connecting strategies were used to realise meaning across data. They are particularly important in this thesis, for the reason that they represent a key mechanism in the building of theory (Maton & Chen, 2016; Maxwell, 2013). Further to the definitions provided in Section 5.4.1, connecting strategies can be realised in a few different ways that are relevant to this research. One of these reflects a focus on connections between data within a single source (for example, an individual interview transcript, or a journal article). Following the connections made within a single source of this nature can help reveal aspects of the data that might otherwise remain hidden in segmental analysis—for example, an interviewee's change in mindset over the course of an interview, or the development of an argument that isn't explicitly named in the text. This approach has a certain amount of overlap with the coding processes described above, in that it provided nuance to the set of codes described in Table 5.2. The main limitation of this kind of connecting analysis is, to quote Maxwell (2009), that "a question about similarities and differences across settings or individuals, or about general themes in your data, cannot be answered by an exclusively connecting analysis" (p. 239). Hence, the second meaningful way in which connecting strategies are realised in this work is in the connecting of concepts across data, codes, and theory. Chapters 6, 7, and 8 of this thesis are directly concerned with this, and it explains the absence of simple reportage of data in the following chapters. Although a standard approach in qualitative studies is to include a large amount of rich description of data, this is not the primary objective in the following chapters, which instead focus on discussion of the connections between data and concepts to provide detailed responses to the research questions.

Since connecting analyses are concerned with the relationships between different parts of the text they are "often seen as holistic" (Maxwell, 2013, p. 112). In departing from a coding-type strategy as outlined above, they offer a means for grappling with one of the common issues associated with qualitative analysis, which is that the meaning associated with given data may change depending upon the size of an individual component of a given sample (for instance, a paragraph in a journal article, as distinct from the entire paper,

see for example Campbell, Quincy, Osserman, & Pedersen, 2013). This is usefully illustrated from a linguistics perspective by Martin (1992) who explains that "linguistic phenomena usually have to be viewed from a number of complementary angles to be fully understood" (p. 10). Martin[15] highlights that some meanings are available in individual words, others across small passages, and others still occur as a result of longer stretches of text. This has implications not only for connecting analyses, but also for coding, and is the rationale for why the same text can be interpreted in multiple ways.

5.4.4 Building theory through translation

This section explains how concepts from LCT (Chapter 4) can be realised in data analyses. This is an important step in traversing the discursive gap (Chapter 2, Section 2.3) that lies between theory and data. Both coding and connecting strategies offer insights as to how theory and data may come together—the core business of analysis—however they do not in themselves make this explicit. Indeed, Maton and Chen (2016) contend that while much research possesses powerful internal conceptual languages (internal languages of description, see Section 2.3), there is often a lack of an external language that can productively relate theory to data, therefore making the traversal of the discursive gap more explicit. This has implications for the trustworthiness of research, since "without making explicit one's theory and the principles of its enactment, and in ways that enable others to recreate the analysis for themselves, the veracity of one's knowledge claims remains obscured" (2016, p. 28). A solution to this problem is to develop a more explicit external language of description that connects theory and data (it is therefore related to the second kind of connecting strategy described above). It has become relatively standard in studies working with LCT to realise external languages of description in the form of *translation devices*. Translation devices are a means of relating theoretical concepts and data which, in so doing, challenge a dichotomous inductive/deductive view of analytical meaning-

---

[15] A fuller explanation of Martin's (1992) approach is beyond the scope of this chapter, however this multi-faceted view of meaning construction across text builds on earlier work by Halliday (1979) and Pike (1982) who uses an analogy from physics to describe text in terms of particles (smallest units), waves (larger units), and fields (larger units again).

making (Maton & Chen, 2016). They do this by comprising "a concept, indicators of how the concept is realized empirically, criteria against which relations with data can be decided, and examples from the data" (pp. 30–31). Translation devices can be formed in a variety of ways, however one of their general features is that they provide multiple points of reference for interpreting the way in which meaning is made between theoretical concepts and data—a simple example was provided in Table 5.3. Translation devices do not necessarily need to assume the form of a table, although this is a common approach.

The structural properties of LCT mean that translation devices using concepts from the framework will often include some common features. These include an indication of how a concept (for example, epistemic relations) manifests, as well as distinctions between areas of relative strength and weakness where continua are involved. Since LCT is an interpretive framework, analysis is given meaning by determining relative distinctions between different locations on continua (for example, a distinction between stronger and weaker epistemic relations). In this, the coding and connecting strategies outlined above are an integral resource. Translation between theory and data is a central interest in this thesis, and each of the following three chapters responds to the research questions by exploring this. Each of these chapters could be conceived of as an elaborate translation device, bringing together theory and data, and in response to the aims of the research, linking the two together to develop conceptual approaches for interpreting different aspects of musical performance assessment practices.

5.4.5 Additional notes

5.4.5.1 Software

The research was aided by use of several software packages, including Nvivo 11, DevonThink Pro, and Microsoft Excel. Initial substantive coding of several interview transcripts was completed in Microsoft Excel which provided a simple way to code the transcripts line per line. Once an initial set of codes had been refined, the interview data were transferred to Nvivo 11 which offers useful coding functionalities and text query tools. While it proved difficult to keep track

of the many early substantive codes in Nvivo 11, once the codes had been refined the application became a more useful container for this data. From the beginning of the project, other documents and academic literature were compiled in DevonThink Pro, which is a powerful digital filing application that can handle a large amount of documents and provides detailed text query tools. Microsoft Excel was used intermittently throughout the project to develop finer grained analyses of specific sets of data (for example, a passage from an interview), and provided a useful way to visually realise coding alongside text.

5.4.5.2 Quotations within the thesis

Although it is relatively standard practice to include a large number of quotations from data in the qualitative research, the approach taken in this thesis is slightly different. Given the theoretical aims of the work, the following three chapters are each focused on developing a theoretical response to the research questions. Each describes a process of translation between theoretical concepts and data to develop these responses, to which end the focus is *not* on thematic reportage of the kind that would commonly be found in many types of case study research. Rather, data are included explicitly in theory-building capacities. The aim in these chapters is not simply to re-describe data, but to use them as an empirical base for the development of further theoretical ideas. In this, many of the quotes included are somewhat extended to more fully provide for the illustration of concepts within the discussion.

5.5 Trustworthiness

Throughout this chapter I have discussed at various points the applications and limitations of different aspects of the research design. In doing so, my intention has been to address points of methodological *trustworthiness* within the methodological context. This section discusses the notion of trustworthiness, which is used in a similar way to Robson (2011)—as a general term for collecting and discussing constructs such as validity and reliability which have an effect upon the quality of this research. Although several of these concepts are discussed individually here, I shall continue to use trustworthiness as a general collective term. I feel it is necessary to note again

a general consensus amongst qualitative methodologists that concepts such as validity—which can have very specific definitions in quantitative research paradigms—can be seen to function in different ways in qualitative research contexts (Maxwell, 2013; Robson, 2011; Silverman, 2011). Scholars of qualitative methodology have proposed a variety of viewpoints regarding the constituent elements of a trustworthy qualitative research design, a number of which are discussed here.

5.5.1 Validity

Several prominent methodologists agree that defining validity is difficult in qualitative research paradigms. Robson (2011) notes that "it is something to do with being accurate or correct, or true" and that it refers to consideration of "what might be appropriate bases for judging something to be credible" (p. 156). Maxwell (2013) takes a similar approach, proposing that validity refers to "the correctness or credibility of a description, conclusion, explanation, interpretation, or other sort of account" (p. 122). Maxwell (2013) makes two further points: Validity cannot be guaranteed by methods, and validity "does not imply the existence of 'objective truth'" (p. 122). Rather, it is about testing our research outcomes "against the world, giving the phenomena that we are trying to understand the chance to prove us wrong" (pp. 122–123). From this perspective, a useful way of approaching validity in qualitative research is in terms of how it might be threatened (Robson, 2011; Maxwell, 2013). The threats to validity outlined below are derived from Robson (2011) and Maxwell, (2013).

- Description: This threat refers to fallibilities in how the data is represented—Robson (2011) emphasises the value of employing approaches such as audio-taping and videoing can reduce the severity of this threat
- Interpretation: For Robson (2011), this threat refers to the ways in which the interpretation of data may be protracted by the imposition of theories and frameworks and that the appropriateness of use of such tools should be re-examined and modified as necessary throughout the

research—one approach to counteracting this threat is to maintaining a clear articulation of how the conclusions were reached (Mason, 1996)

- Theory: This threat refers to the dangers inherent in not considering alternative perspectives on the data in question and can be offset by seeking examples within the data that run contrary to the explanatory theories formed by the researcher (Robson, 2011)

- Bias: For Maxwell (2013) this threat is characterised by the "selection of data that fit the researcher's existing theory, goals, or perceptions, and the selection of data that 'stand out' to the researcher" (p. 124)—it can be counteracted by identifying and making clear one's possible biases as the researcher

- Reactivity: This threat refers to the influence the researcher has on the data—this is not always resolved necessarily by seeking to minimise one's influence as the researcher (though approaches such as avoiding leading questions remain valuable) insofar as acknowledging how the data was influenced in the first place

5.5.2 Reliability

Robson (2011) and Silverman (2011) both note that the usual quantitative reading of reliability poses problems for qualitative researchers. While reliability in quantitative research is concerned generally with the repeatability of the findings by other researchers employing the same methods (Merriam, 2009; Robson, 2011; Silverman, 2011), in qualitative research "[a]t a technical level, the general non-standardization of many methods of generating qualitative data precludes formal reliability testing" (Robson, 2011, p. 159). Some methodologists argue that reliability in qualitative research is, rather, an impossibility. This perspective posits that the relationships between individual researchers and that which is researched are fundamentally unrepeatable, and further that attempting to impose a framework that can be tested for reliability may adversely influence qualitative flexibility of the work (Merriam, 2009). Indeed, Merriam argues that "[t]he more important question for qualitative research is *whether the results are consistent with the data collected*" (2009, p. 221, italics original). Conversely, others (Silverman, 2011, for example) argue that reliability must be a consideration for the qualitative

researcher, and that it can often be enhanced by employing multiple researcher perspectives to the ends of testing whether the same deductions are made categorically or in using the existing theory. This research adopts a middle-ground approach, and alternative viewpoints were available for cross-checking the data usage throughout the project. These viewpoints came from both the supervisors of the research project, as well as in the form of supporting feedback from other colleagues specialising in the fields of educational assessment, musical performance assessment, and LCT (see Appendix A for further description). Nonetheless, one point that methodologists appear to agree upon is that a method of enhancing the reliability of qualitative research is to leave a clear trail as to how the conclusions were reached (Merriam, 2009; Robson, 2011).

5.5.3 Generalisability

As with validity and reliability, the view of generalisability informing this research departs from aims of statistical representation that are common to quantitative paradigms (Maxwell, 2009, 2013; Robson, 2011; Silverman, 2011). The notion of generalisability that informs this study refers to how one might go about "extending research results, conclusions, or other accounts that are based on a study of particular individuals, settings, times, or institutions to other individuals, settings, times, or institutions than those directly studied" (Maxwell, 2013, p. 136). In qualitative research, statistical generalisability is frequently limited because the relevant sampling approaches are not practically feasible (Silverman, 2011). While this observation is certainly true of this research, it is not necessarily a problem: As Alasuutari (1995) notes, many qualitative projects (including this one) are not about testing specific hypotheses, but rather pursue more open-ended explorations of given subjects.

Maxwell (2013) proposes two types of generalisability: *internal* and *external* (see also, Robson, 2011). Here, internal generalisability refers to the "generalisability of conclusions within the setting studied" while its external counterpart is generalising "beyond that setting" (Robson, 2011, p. 160). Both Maxwell (2013) and Robson (2011) observe that internal generalisation is frequently problematic if the context of the object of study is not addressed

representatively—a key concern for qualitative research given that "it is impossible to observe everything even in one small setting" (Maxwell, 2013, p. 137). Qualitative research can however work to develop explanatory theory—a focus in this project project—that helps "in understanding other cases or situations" (Robson, 2011, p. 160). As Maxwell (2013) puts it:

> [T]he generalizability of qualitative studies is usually based not on some explicit sampling of some defined population to which the results can be extended, but on the development of a theory of the *processes* operating in the case studied, ones that may well operate in other cases, but that may produce different outcomes in different circumstances. (p. 138)

Additionally, because of its inherent depth, qualitative research may also be suited to the task of identifying what Popper (2005) refers to as *black swans*—ideas that do not appear thinkable until such a point as they are uncovered. This is to say that one of the potentials of research such as this is to identify that which runs counter to existing hypotheses (Silverman, 2011). Similarly, qualitative research may not seek to generalise across unstudied contexts, but simply to be clear about what has been concluded from the empirical data (Alasuutari, 1995; Maxwell 2013; Robson, 2011; Silverman, 2011).

5.5.4 Limitations

In reviewing the methodology of the research, several main limitations are apparent. Firstly, this research is beholden to the general limitations of most interpretive qualitative designs, which has particular implications for the kind of validity and generalisability it is capable of—the research does not provide statistically generalisable conclusions. While the purposeful data sampling strategy that was selected is useful in view of the explorative research aims, it does mean that the data may not be as easily re-purposed for future work. Related to the validity of the study, another limitation is that the research was conducted by an individual rather than a team, meaning that inter-rater coding was not possible beyond the feedback that was available from the supervisory team or from professional colleagues—further information on this is provided in Appendix A.

A third limitation is that the work is highly theoretical. While this is a deliberate aspect of the work, it does mean that some calibration of the ideas developed in this thesis will be needed to facilitate their implementation in more specific contexts or other research studies. Related to this are limits inherent in the interpretive nature of the analysis—beyond the initial explorative phases of analysis, subsequent analysis and theorising was influenced by foci of the research questions. Since the research questions are about *ways of theorising* rather than questions about *how things are* at a particular research site, later analysis was mostly focused on finding intersections between theoretical ideas and the data. Since the data do offer a range of useful insights into actual assessment practices and thinking, re-analysing the data using a stricter content analysis or discourse analysis approach might reveal different insights which could be useful in directing further research.

5.6 Summary

This chapter described the methodological approach taken in the research. The first part of the chapter (Section 5.2) provided an overview of this approach, and outlined the basic perspectives upon which the approach is predicated. Section 5.3 described how data are conceptualised in the research and explained the kinds of data that inform the study—these include transcripts from interviews with assessment participants, academic literature, and publicly available documents (including policy, procedure, and coursework documentation) that shed light on assessment practices in higher music education institutions. Section 5.4 explained the analytical approach followed in the research. This approach includes categorising strategies (coding) and as well as connecting strategies which place an explicit emphasis on building meaning across data. Section 5.4 also introduced translation devices, which are the principal means by which concepts from LCT are related to data. Section 5.5 concluded the chapter by discussing issues related to trustworthiness in the research. Having described the methodological approach of the research, the following chapter is the first of three to respond to the research questions.

Chapter 6: Conceptualising Bases of Achievement for Musical Performance Work

6.1 Introduction

This is the first of three chapters to respond to the research questions of the study. The following chapters aim to bring theory and data together in dialogue to develop conceptual possibilities in response to these questions. This chapter develops a response to the first question: How can bases for achievement in musical performance work be meaningfully conceptualised? Concepts from the Specialization dimension of Legitimation Code Theory (LCT) provided theoretical tools for interpreting and exploring the question which led to the development of a theoretical device for conceptualising criteria in musical performance assessment practices.

The first part of the chapter (Section 6.2) outlines the main theoretical points of departure for the discussion developed through the chapter. It reviews the basic interpretation of the question and explains the approach that was taken in using concepts from LCT to explore the question. As part of this, Section 6.2 locates criteria as the main object of focus for studying bases of achievement. Section 6.3 explains an approach for empirically realising concepts from the Specialization dimension and illustrates connections between these concepts and more substantive criteria for musical performance assessment. Section 6.4 leverages the conceptual ideas developed in the previous sections to explore perspectives on criteria for musical performance assessment collected through interviews with higher music education staff and students. The final part of the chapter (Section 6.5) offers some discussion to highlight the implications and possible applications of the work.

6.2 Theoretical Points of Departure

6.2.1 Achievement

The research question that anchors this chapter is about bases for *achievement* in musical performance assessment. Although achievement as a term can have many possible interpretations its meaning in education literature is often loose and taken for granted (Sadler, 2010). Sadler (2010) asserts that it is "rarely defined, explored as a concept, or listed in book indexes" (p. 729). In response, Sadler (2010) offers a re-examination of the terminology based on common meaning and etymology:

> In ordinary discourse, to achieve means to bring to fruition some significant act not previously accomplished, or to attain some significant performance status not previously reached. It means being successful in bringing about a desired end as a result of substantial effort, and clearly involves challenge. Normally, the magnitude of an achievement is proportional to the challenge involved. Climbing Mount Kilimanjaro, rescuing a company from looming insolvency, or winning a major literary award would be significant achievements for most people. In the higher education context, completing a degree counts as a significant achievement for most students, as does having a scholarly article accepted in a reputable journal for an academic. In each case, an act is completed, finished or finalised and the result is clearly evident. Whether something should be classified as achievement depends to some extent on the context. Something that is accepted as an achievement for one person may not be so for another. Thus, threading a needle would not count as an achievement for the ordinary person, but could be a significant achievement for a person with poor fine motor coordination who has struggled to master the technique. (p. 729)

Sadler (2010) offers a basic definition for achievement which reflects "[a] goal or level reached; an enterprise completed, accomplished, attained

successfully, or brought to a successful end—especially by means of exertion, effort, skill, practice, or perseverance" (p. 730). In academic education contexts, achievement can therefore include many things, including the degree to which students achieve a particular quality of response to a task, the quality of learning, the quality of teaching, promotions or advancement through coursework, an improvement in social standing, and so forth. The basic concept of achievement in the more specific context of musical performance work still accrues many possible points of reference: it might be about how well a student performs a particular work, how far a student has come in their development, the effective uptake of feedback, overcoming performance nerves, and so on. In the context of assessment, it is usually the case that some kinds of achievement are considered acceptable points of reference for grading students' work, while others are considered to be contaminants, or what Sadler (2010) refers to as *non-achievements*. To begin with, this chapter is concerned with students' achievement in musical performance assessment in the more general sense—that is, without the imposition of what is or isn't considered to be a legitimate achievement. The reason for this is the focus on *underlying* bases of achievement—the issue of non-achievement is considered later in the chapter (Section 6.5.4).

6.2.2 Situating criteria

Judgements about students' achievement are informed by criteria (Chapter 3). From a theoretical standpoint, criteria can be viewed as bases for achievement in that they form points of reference by which to ascertain a qualitative judgement about degrees of achievement, and they are similarly characterisable in terms of their relative legitimacy (see Chapter 3). While studies of criteria often tell us in descriptive terms about the *kinds* of substantive criteria available (and desirable), they tell us less about *underlying* bases for achievement in the sense of more basic organising principles that can describe a fuller range of possible criteria. Although some studies get closer to these underlying bases by discerning categorical and theoretical commonalities between criteria, these are usually locked into descriptive typologies (Chapter 3). As a response to this situation, this chapter considers how concepts from the Specialization dimension of LCT can be enacted to

offer a more dynamic means of interpreting bases of achievement in musical performance assessment.

To focus this part of the research, criteria were taken as the main object of study with which to conceptualise bases of achievement in analysis. Specialization was selected as the main theoretical point of departure for exploring the question on the basis of earlier thought experiments and pilot analyses (Maxwell, 2009). It is useful for the reason that it highlights an analytical interest in how criteria are specialised: that is, an interest in the underlying principles of criteria themselves which helps us to think through what particular criteria are actually *about*. It provided a basis for developing a more theoretical language of description for musical performance assessment criteria. The discussion developed in this chapter explores the translation between theoretical concepts from Specialization and substantive criteria for musical performance assessment. The development of a translation device in Section 6.3 provides a starting point for theorising ways in which bases of achievement in musical performance assessment can be meaningfully interpreted to illustrate the underlying principles embodied by criteria selection and emphasis. This device provides a means of realising the manifestation of theoretical concepts in substantive perspectives on musical performance assessment criteria (Section 6.4) to infer underlying positions that frame students' achievement. This section begins by reviewing the basic features of Specialization (Chapter 4) and the approach by which it was enacted in the research (Chapter 5).

6.2.3 Specialization

Specialization is useful for the reason that it is explicitly purposed towards theorising bases of achievement in social practices (Chapter 4). Its purpose, in part, is to highlight relationships between empirical possibilities and to move beyond the use of descriptive terms that remain close to practice. Specialization was introduced in Chapter 4, and its basic features are briefly revisited here. The dimension begins from the premise that all practices, beliefs, and knowledge claims are "oriented towards something and by someone, and so sets up an epistemic relation to an object and a social

relation to a subject" (Lamont & Maton, 2010, p. 5). The basic constituents of Specialization are

- *epistemic relations* [ER] between practices and their object or focus (that part of the world towards which they are oriented); and
- *social relations* [SR] between practices and their subject, author or actor (who is enacting the practices) (Maton, 2014, p. 29, italics original).

Taken together, these two sets of relations generate a space of possibilities (Figure 6.1).
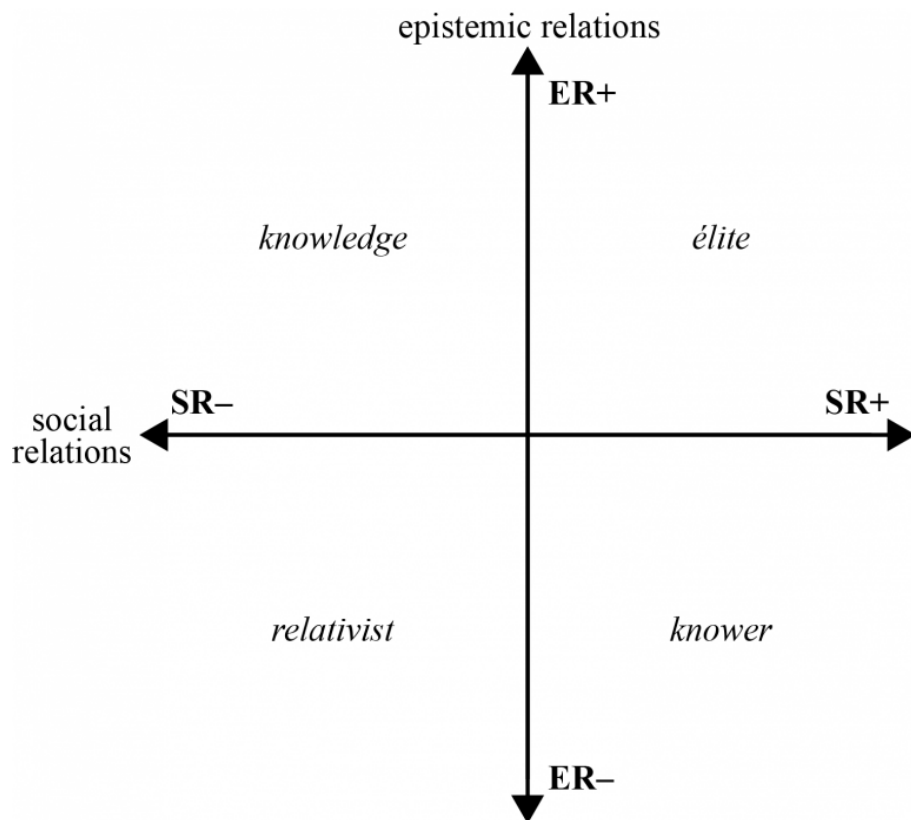
Figure 6.1 The Specialization plane (Maton, 2014, p. 30), reprinted with permission

The basic modalities of this space are

- *knowledge codes* (ER+, SR-), where possession of specialized knowledge, principles or procedures concerning specific objects of study is emphasized and the attributes of actors are downplayed
- *knower codes* (ER-, SR+), where specialized knowledge and objects are downplayed and the attributes of actors are emphasized as

measures of achievement, whether viewed as born (e.g. 'natural talent'), cultivated (e.g. 'taste') or social (e.g. feminist standpoint theory)

- *élite codes* (ER+, SR+), where legitimacy is based on both possessing specialist knowledge and being the right kind of knower; and

- *relativist codes* (ER-, SR-), where legitimacy is determined by neither specialist knowledge nor knower attributes—'anything goes'. (Maton, 2016a, pp. 13, italics original)

6.2.4 Approach

This part of the research draws mainly on discourse around criteria for musical performance assessment and is informed by a mix of data, including scholarly literature, perspectives from research participants, and institutional documents. Since LCT is an explanatory framework (Chapter 2), it offers generic concepts that need to be calibrated to context (Chapter 4). The main aim here was to explore the potentials of the Specialization dimension insofar as it could offer a productive perspective on bases of achievement for musical performance assessment. Since the main aim of using LCT is to exploit its potential as an *explanatory* framework a more specific translation device (Chapter 5) was developed to help explain the meaning through which the theoretical concepts were realised empirically, and vice versa. The development of the device was an iterative process involving:

- Initial development of an interpretation based on existing knowledge of musical performance assessment criteria and theoretical concepts, as well as early engagement with data (including scholarly literature)

- Refinement of the device based on repetitive interaction with data— including interview data, institutional documents, and scholarly literature—and theory

An important part of the process for developing the device was the accumulation of examples of substantive realisations of epistemic relations and social relations. A useful resource to this end was extant literature on musical performance assessment criteria which provided a helpful starting point. The following section explains the basic process by which these criteria were coded into the translation device. Although the structure of this chapter

paints a relatively linear picture of the process, the work involved many micro-iterations of refinement as interpretations shifted over the course of the research. What is documented here is, therefore, an intentionally simplified representation of the process that describes the outcome in as straightforward a manner as possible.

6.3 Translating Concepts

The literature on musical performance assessment usefully offers a range of possibilities for criteria which have already, to varying extents, been analysed and effectively pre-coded into categories by researchers. To illustrate the translation between extant criteria for musical performance and the concepts from Specialization this section draws on contributions from Blom and Encarnacao (2012), McPherson and Schubert (2004), and Wrigley (2005). Each piece of scholarship provides a considered summary and discussion of substantive criteria for musical performance assessment. Blom and Encarnacao (2012) propose eight main regions of criteria including:

- Technical skills
- Analytical skills
- Appreciative skills
- Personal skills
- Interpersonal skills
- Organisational skills
- Hard to measure, on-stage criteria
- Criteria particular to the rehearsal process

McPherson and Schubert (2004) describe four categories[16], including:

- Technique
- Interpretation
- Expression
- Communication

---

[16] McPherson & Schubert (2004) include these within a broader *musical* category: they also propose others which are not included here within *non-musical* and *extramusical* categories.

Wrigley (2005) distinguishes a range of categories associated with instrument groups, including:

- Strings
  - Technique
  - Musical understanding & performance
- Piano
  - Technical mastery & control
  - Sound quality
  - Convincing musical understanding
- Woodwind
  - Technical control
  - Sound production
  - Musicality & interpretation
- Brass
  - Technical preparation
  - Sound production
  - Musical interpretation
- Voice
  - Technique
  - Interpretation
  - Musicality

These broader categories provided some clear points of departure for realising the conceptual space of Specialization. Preliminary coding grouped those with a clear emphasis on knowledge and skills as reflective of epistemic relations, those with a clear focus on people as emphasising social relations, and the remainder in an undefined category. Repetitive themes (for example, technique and technical skills) were grouped together (see Table 6.1). This broad analysis offered a starting point for finer-grained coding. Blom and Encarnacao (2012), McPherson and Schubert (2004), and Wrigley (2005) each include a range of examples within these broader categories, and access to this finer level of nuance enabled a more descriptive analysis.

Table 6.1 Broad realisations of epistemic relations and social relations

Epistemic relations	Social relations	Undefined
Technique	Personal skills	Appreciative skills
Analytical skills	Interpersonal skills	Organisational skills
Sound quality and production	Expression	On-stage criteria
	Communication	Criteria particular to the rehearsal process
Musical understanding	Musicality	Musical understanding and performance Interpretation

The availability of these distinctions is a useful analytical resource for the reason that they offer gradations of explanation—the finer-grained categories are closer to substantive practice, in that they generally express possible empirical realities more clearly than the broader concepts. Realisations of these categories are therefore more readily available to the imagination of the analyst who is familiar with the field. Table 6.2 reconceptualises the broader categories that were previously coded under epistemic relations using these finer-grained examples. This in turn provides a more nuanced view of this region of criteria and highlights subtler distinctions between those criteria that are more strongly reflective of techniques and skills (epistemic relations), and those that include a broader latitude for a basis in personal taste or other personal attributes (social relations).

Table 6.2 Finer-grained analysis of constituents of epistemic relations

Less likely to include a basis in social relations	Greater capacity to include a basis in social relations
Playing technique	Phrasing
Fluency of performers	Sense of line and sensitivity to harmonic movement
Accuracy of playing notation	
Intonation	Balance/blend
Playing in time	Expressive qualities
Tempo	Emotional impact
Articulation	An understanding of the work
Dynamic variation	Interpretive qualities
Tone quality	Style
Timbral variation	Character
Breathing/posture	Conviction
Projection	Collaboration
Awareness of style/performance practice	Cohesion
Rhythm	
Style	
Clarity	
Coordination	

This finer-grained analysis enables distinctions within the broader category of technical skills to highlight possible subtleties within technical skills themselves. Repeating the analytical exercise with the remaining categories produces a pair of lists of criteria that constitute substantive realisations of epistemic relations and social relations. For those categories that did not express a clear association with either epistemic relations or social relations—for example, phrasing—additional terminology was added to clarify the difference between analytical possibilities with the given category. For example, phrasing can be further distinguished to include skilful or idiomatic phrasing (a stronger basis in epistemic relations), and phrasing based upon personal taste (a stronger basis in social relations).

Table 6.3 Detailed realisations of epistemic relations and social relations

Concept	Broad examples	Finer-grained criteria		
epistemic relations	Technique Analytical skills Sound quality and production Musical understanding	Playing technique Fluency of performers Accuracy of playing notation Intonation Playing in time Tempo Articulation Dynamic variation Tone quality Timbral variation Breathing/posture Skilful phrasing	Rhythm Idiomatic style Clarity Coordination Phrasing Technical and stylistic knowledge of the work Balance/blend Skilful cohesion Sense of line and sensitivity to harmonic movement Projection	Awareness of style/performance practice Informed interpretation
social relations	Personal skills Interpersonal skills Expression Communication Musicality	Reliability Attendance Remembering equipment Sense of involvement Sense of individual personality Emotional impact Personal understanding of the work Tasteful cohesion Verbal interaction Trust and respect	Sharing of decision-making Achieving integration Social factors including cooperative skills Negotiation of shared musical goals Expressive qualities Expressive phrasing Collaboration Personal interpretation Character Conviction Personal style	Confrontation and compromise Feeling of artistic contribution Confidence Imagination

Repeating this analysis with additional data invariably produces further possibilities and is one of the key strategies in developing a more nuanced interpretation of the field (hence the inclusion of multiple sources here). Comprehensive analysis of the field is not the objective here, however, and would be beside the aims of this part of the thesis. Rather, this analysis demarcates a basic and fairly generalised interpretation of the territory, and its main use is as a starting point for developing further analyses of practice (as in the following section).

Building on this analysis, Table 6.4 incorporates the continua of strengths of epistemic relations (ER+, ER-) and social relations (SR+, SR-) to synthesise a translation device for theorising musical performance assessment criteria with Specialization. The inclusion of indicators for stronger (denoted by '+') and weaker (denoted by '-') manifestations of epistemic relations and social relations offers a starting point for a topological view of assessment criteria given by the Specialization plane (Figure 6.1). The range of possible Specialization codes embodied by criteria practices are given by the combination of criteria enacted in an assessment situation and the relative emphases placed upon them. Assessments that place a stronger emphasis on performative, analytical, and interpretive techniques, skills, and understandings (epistemic relations) to the exclusion of criteria relating more closely to personal attributes (including taste, self-expression, and relationships with others; social relations) embody knowledge codes (ER+, SR-). The inverse—an emphasis on taste and personal interpretation— produces knower codes (ER-, SR+). The presence of both produces élite codes (ER+, SR+), while a lack of emphasis (or de-emphasis) upon both produces relativist codes (ER-, SR-). Important here is the idea of the *criteria environment* for an assessment. As discussed in Chapter 3, musical performance tasks typically involve a synthesis of multiple criteria, both in the designing and enacting of work as well as in the appraisal of the work. Hence, singular examples of specific criteria would seem inadequate indicators of Specialization codes for assessment events—rather, their presence may serve to signify the possibility of a code. Confirming the presence of a code requires investigation of a range of data (here, representations of criteria) and the degree of emphasis represented in a given context. The following section

builds upon the analytical groundwork developed here to consider perspectives on criteria for musical performance assessment arising from interviews with research participants.

Table 6.4 Translation device for realising epistemic relations and social relations in musical performance assessment

Concept	Manifests as	Broader examples	Substantive possibilities		Strength of relations	Indicated by
epistemic relations	Emphasis on performative, analytical, and interpretive techniques, skills, and understandings	Technique Analytical skills Sound quality and production Musical understanding	Playing technique Fluency of performers Accuracy of playing notation Intonation Playing in time Tempo Articulation Dynamic variation Tone quality Timbral variation Breathing/posture Rhythm Idiomatic style Projection	Awareness of style/performance practice Clarity Coordination Phrasing Technical and stylistic knowledge of the work Balance/blend Skilful cohesion Sense of line and sensitivity to harmonic movement Informed interpretation Skilful phrasing	Stronger epistemic relations (ER+)	Performative, analytical, and interpretive techniques, skills and understandings are emphasised as the basis for achievement
					Weaker epistemic relations (ER-)	Performative, analytical, and interpretive skills and understandings are less important in determining students' achievement
social relations	Emphasis on personal attributes including taste, self-expression, and relationships with others	Personal skills Interpersonal skills Expression Communication Musicality	Reliability Attendance Remembering equipment Sense of involvement Sense of individual personality Sharing of decision-making Achieving integration Social factors including cooperative skills Negotiation of shared musical goals Verbal interaction Confrontation and compromise	Feeling of artistic contribution Confidence Imagination Personal style Collaboration Emotional impact Personal understanding of the work Tasteful cohesion Conviction Personal interpretation Character Expressive qualities Expressive phrasing Trust and respect	Stronger social relations (SR+)	Personal attributes including taste, self-expression, and relationships with others are emphasised as the basis for achievement
					Weaker social relations (SR-)	Personal attributes including taste, self-expression, and relationships with others are emphasised as the basis for achievement

6.4 Enacting Theory to Explore Perspectives on Bases of Achievement in Musical Performance Assessment

This part of the chapter builds upon the previous section to develop a thematic discussion of data collected from interviews with staff and students involved in musical performance assessment at a range of Australian higher education institutions. It describes observations produced through application of the translation device described in Table 6.5 to the interview data. The broader focus in these interviews was on interviewees' perspectives on musical performance assessment, and each interview included questions relating to criteria for musical performance assessment. The discussion developed here is the product of an analytical dialogue between the concepts developed in the previous section and findings from the interviews—it illustrates how these concepts can be directly related to empirical data. Prior familiarity with theoretical concepts from both LCT and the field of assessment, coupled with early analytical ventures, meant that some analytical preconceptions existed before these concepts were applied to data from interviews. This was both a strength and limitation of the process. To help mitigate the risks associated with applying abstract theoretical concepts prematurely in the analysis process, several iterations of coding were performed on interview data collected earlier in the project to capture a range of substantive themes independent of concepts from LCT (see Chapter 5). This early analysis provided a broad range of substantive themes. To theorise the data, themes relating to the conceptualisation of criteria were subsequently analysed using the Specialization dimension of LCT, and the basic starting points for realising concepts in analysis were introduced in the previous section. Through dialogue between the theoretical concepts and interview data a more nuanced set of observations emerged. The discussion developed here summarises these observations across three themes.

6.4.1 Epistemic relations as a basis for achievement

This section summarises perspectives from data that highlighted epistemic relations as a basis for achievement in musical performance assessment. Analysis of the interview data indicated a stronger emphasis on epistemic

relations (ER+) given principally by the value accorded to technical skills. Many interviewees emphasised the basic importance of technical skills as foundational to success in musical performance assessment. One Head of School (Assessor Q) provided a detailed explanation:

> I think whatever the style of music there are a number of non-negotiables… for example, if it's an instrumental performance it's the accuracy, it's the technical ability of the performer to be able to get across the material accurately and well. Is the tempo constant, or are they slowing down because the music is difficult? Is the piece actually being played from start to finish without breaks?... That's a fundamental issue. Intonation is a fundamental issue; being in tune or not. The clarity of diction for a vocal performance I think is absolutely critical.

Several participants highlighted that the contextual setting of higher music education presupposed "an assumed technical level" (Assessor B, a classical department leader) for students' performances, and a responsibility to uphold a certain standard of technical expertise. As one participant commented: "Of course we're an institution, so we do need to be looking at the technical versatility of someone" (Assessor J, multi-disciplinary department leader). Related to this was the view that technical work is more transparently assessable. Although participants agreed that assessing musical performance work is complex, several participants felt that technical work was simpler to assess. One classical department leader, for example, explained that he had developed a numerical system for "actually marking each individual scale with a tick or a cross… [to] come out with a percentage" (Assessor B). Assessor B expanded: "The only time I use… a rubric of criteria is for tech exams, which is very easy: you can give a cross or a tick, can they lip trill? Can they transpose? Five scales, which ones are correct?". This participant characterised this as "a much more, sort of, summative assessment, as opposed to what we do with the performance [recital]". Being focused on more tightly defined skills, more in-depth technical assessments were construed by some as valuable, though more difficult to enact for logistical reasons:

It's a little harder to do but, you can assess them [students] on more specific things… At [an institution] we used to have the technical assessment—let's all get together and play our scales. If you wanted to—and this is all where it becomes more labour intensive of course, and universities don't really want to do it because they have to pay people to do it—I could get someone in a room and say, 'can you play me this chord progression in this key using… guide tones and making the harmony evident… don't play it like Eric Dolphy, play it like Joe Pass… I want to hear the chords [in your improvisation]'… It could be something even more prescribed [such as]: 'Have a look at this chord progression. Play me all the arpeggios and the chords. Or play me all of the chord scales that go with these chords'. (Assessor P, a jazz and popular music teacher)

Related to the previous quote, one student administrator noted distinctions in emphasis on technique across disciplinary areas within an institution:

Performance assessments for classical are more like solo assessments. They're all across the board looking at technique obviously. But classical is extremely technical and I think the students feel a lot of pressure being up there by themselves. But then again, the [students in other disciplines] don't want to let each other down… Somewhere in the middle is jazz, where there's all this technical pressure, plus they're working together in an ensemble. (Student Administrator A)

Others made similar observations in relation to specific disciplinary areas. For example, an executive leader leader at one institution remarked:

it's a difficult thing to actually sort of… [come] together with common structures and common goals. Even though we do have rubrics and graded sort[s] of criteria of what makes [a] distinction [and] what makes [a] high distinction, you'll tend to find that certain areas will just naturally mark higher. For example, strings.

Strings and pianos I find that… they tend to mark the students that much higher up on the scale. As [for] other areas… brass for example, [brass staff member] is quite a hard task master and nothing's good enough so he will tend to bring it down. Everyone will be looking at the same graded descriptors, but they'll be kind of interpreting them in a slightly different way. (Assessor F)

Although participants agreed that technical work is easier to assess, some noted caveats. As one participant observed "you've got other layers [beyond the set criteria], what's it sound like or something?" (Assessor B, a classical department leader). For Assessor B, the solution to this was to provide additional space in the rubric to account for this: "There's room in the comments to say, your sound qualities aren't refined or refined—you can have a broader comment around the criteria of a tech exam". For another participant, however, additional complexity lay in his personal experiences of the technical work being assessed:

Even the technical stuff is hard to assess because I think, sometimes… I have to sit there and ask my kids, play me such and such a scale. Now, I would have immense difficulty with that, playing stuff out of context—you know, play me an F Phrygian, play me a… Lydian augmented, play me an E Lydian. This stuff is hard to play out of context, to just—to come up with, you know? (Assessor A, a jazz performance teacher)

While participants generally agreed that a threshold level of emphasis on technical skills is valuable, some were also critical of overemphasising technical skills. A multi-disciplinary department leader at one institution, for instance, commented:

We need to make sure that they understand at whatever level they're at, their instrument, their repertoire, the style. From that point onwards I try to look at or try to listen and see the performance from a purely musical experience, keeping all those other things in mind… I know in the past I've been happy to

perhaps give a higher mark for a little bit of technical fault [where] it's been more musical… I think that's more important. (Assessor J)

The interpretation of epistemic relations developed in response to the perspectives of interviewees includes two main features. Firstly, the notion of a base value for technical skills conveys that, for this group of participants, student achievement is usually predicated upon being able to fulfil some threshold standard of technical execution. Analytically, this is distinguished as a consistent baseline emphasis on epistemic relations (ER+). The data suggest that this emphasis is tightened in certain scenarios, including assessment events designed to focus specifically on technical skills, and in situations where the personal values of those involved lead them to place greater emphasis on students' technical capabilities (stronger epistemic relations, ER++). To locate examples of weaker epistemic relations it may be necessary to explore other regions within the higher music education sector. Two possibilities here are popular music and contemporary art music. Although these fields are not without emphasis on technique, both could conceivably offer scenarios in which technical skills are comparatively less important relative to the context of classical and jazz performance degrees. One interviewee who had experience teaching in a popular music context explained:

[T]he thing about teaching in a pop culture course is… Even though the music's simpler the trajectory of 'okay now you know this, let's go onto this. Now you can do that, let's go onto that'— that's a lot easier in jazz. Pop music's so wide that there's no clear trajectory the way there is in jazz education… [where you might say], 'you're over those two five ones in the major key, now let's do them in a minor key'… It's a lot harder to plot a course for pop music. And it's a lot harder to say... Well for instance… As a jazz player if you're playing a chord progression, you're either successful at that or you're not. Whereas it's a little harder

to judge [in pop music]. (Assessor P, a jazz and popular music teacher)

6.4.2 Social relations as a basis for achievement

Analysis of the interview data also highlighted ways in which social relations were emphasised as a basis for students' achievement. Related to the substantive examples discussed in Section 6.4, interviewees generally valued criteria related to interpretation and engagement (SR+). As some of the data documented above illustrates, some participants valued criteria of this kind equally if not more than technical forms of criteria—to quote one assessor: "The most important [layer] for me is the element of enjoyment and engagement from the performer—because there's an assumed technical level—then it goes down beyond that" (Assessor B, a classical department leader). Comments from one student participant highlighted, however, ways in which stylistic repertoire choices could have a substantial influence (SR+) on students' capacity to succeed.

> I can think of… one notable example, where ... a good friend suffered a little bit in his markings because he chose a piece that we did together… in both of our exams… He was a double bassist and I wanted to play this piece with him which was a modern Finnish piece, from like, 1976 or '73… It's about colours and sounds and soundscapes and effects and it was a beautiful piece, really fantastic writing. But for the double bass it touches on a lot of extended techniques, a lot of very different artistic thoughts… we performed that in his exam… they [the examiners] didn't often want scores of pieces that people [students] were playing—because they were playing very standard classical repertoire… Bach and the usual concerto that everyone's done. One of [the examiner's] main comments was, "we would've liked a score for this piece so we could understand a bit more of what's going on." And I think they marked him down… he didn't do as well as he'd hoped... It was a very good performance, a very accurate performance of his piece, and a very colourful and

emotive performance of this piece, but they didn't like it as much. I think, they found that they just couldn't really interpret it or take away what they wanted to take from it… We played it in my recital for the percussion staff, and the performance got a very high mark. They really appreciated and thought that, artistically, what we demonstrated interpretively from this piece was very strong and was very accurate and very sophisticated. But the string faculty felt very differently. And so we had very different marks, I think because the string department might have been looking for something else. (Student A, a classical percussion student)

Notwithstanding interpretive criteria, other social criteria were highlighted by interviewees. In particular, students' progress carried a strong emphasis for a number of participants. As one student explained: "If I were to start a music school or come up with some new idea I feel like you would be assessed… maybe half of it would be on your progression through lessons throughout the semester" (Student F, a classical brass student). The importance of students' progress was highlighted in several different interviews, and linked by some participants to the quality of students as learners. One departmental head emphasised: "We are not running a you know… Australia's Got Talent… we are an educational institution. We want we want to reward progress. We want to reward hard work" (Assessor K, a classical performance lecturer). One participant elaborated on the subject at length:

We're always balancing you know, 'okay, that kid sounded good today, but he sounded like that six months ago when we auditioned him. Whereas that [other] kid sounded pretty awful, but significantly better than he did when we auditioned him'… it is an issue in music because people often say, 'oh, you just have to assess people on their performance on the day', which sometimes is not indicative at all of them having done any work throughout the semester… If you've come to the end of grade 12 [at school] and you sound pretty good and you don't intend… on ever learning any more or doing any more work, then just go out

and do gigs. Don't come to university if you don't intend on doing any more work or learning anything more… We've got quite a lot of students like that, happens all the time… It's how good a student they are you know. Do they show up? Do they practice what you tell them to? Or do they already think they're too good to… pay any attention to you? They're the extremes. (Assessor P, a jazz and popular music teacher)

Personal relationships between staff and students were seen to be important in other ways as well. Some participants communicated mixed messages in this respect, underscoring the complexity of reconciling a perceived need for objectivity in assessment on the one hand, with personal knowledge and relationships with students on the other. For example, a jazz teacher at one institution (Assessor C) provided the following contrasting explanations:

I think I've learnt that I have to have… a bit of separation in those situations… especially if you've taught someone, you know? You have to really put all that stuff that you've built over time—that rapport that you might have with that student… it all has to sort of sit on the sideline while you critically analyse, according to the criteria, how they're doing that day in that moment… It's hard because you might go 'oh my gosh, I've heard them sing that thing so many times in their lesson a million times better and right now it's just not happening'.

You can usually tell if someone's just really nervous as well, so you can... then you do take that into account. Like if you can really see... or if you do know someone quite well—that's a good thing about having that rapport with someone, you can tell and take that into account… I think people, at least in our department, do kind of go 'well… that would be like, you know, a C+ but I could tell they were nervous and that affected everything so maybe we should be a bit kind and be like B- or B or something like that'. Because that [the student's nerves] affected all those other things, you know?

6.4.3 Clashes

Sections 6.4.1 and 6.4.2 used the concepts of epistemic relations and social relations to highlight contrasting bases of achievement for musical performance assessment derived from interviewees' perspectives. These analyses are not intended in any way as broad generalisable reflections on musical performance assessment practices, but are rather a next step in fitting the theoretical concepts developed earlier to a data context. The contrast between epistemic relations and social relations as underlying bases of achievement sets the scene for a complex interplay between the values of those involved in assessment. Multiple participants highlighted the issue that different assessors might apply different criteria in an assessment of the same performance. To quote one classical department leader: "There's lots of different layers [of possible criteria] and unfortunately everybody sitting on a panel might have different concepts of how they will assess said student" (Assessor B). Three excerpts from the interviews are provided below that emphasise this issue.

Example 1: The most important [layer] for me is the element of enjoyment and engagement from the performer—because there's an assumed technical level—then it goes down beyond that: is that student of a technical proficiency to be an [acceptable] standard for that level?… And then finally whether they've fulfilled all the assessment criteria, if there is an assessment criteria, which includes the provision of program notes and all that sort of stuff… Criteria for a performance exam is really difficult. I just see it as a terribly grey area because, if you have three assessors for a senior student's exam they all may have completely different criteria they want to hear. I personally want to hear musical maturity, but again that's a subjective opinion… What I think is musical someone else might not think is musical. It's very difficult so I try and be a little broader with my internal criteria. You want to see progression—like if I've heard that student over three

years I want to hear progression. You want to hear the ability to convey an idea, which is another part of the criteria for me. Then of course you've got, you know, rhythmic integrity, sound quality, things that are identifiable; but then I've sat on panels where someone else has said 'look I don't care that the rhythm was a bit dodgy I really loved it,' whereas, for somebody [else], depending what genre it is you want rhythmic integrity—if it's Stravinsky you want rhythmic integrity, if it's Brahms oh there's a bit of, you know, you can move it around a bit. (Assessor B, A classical department leader)

Example 2: It [17] was an American-written piece and she [the interviewee's teacher] had a very strong basis [in] American study and was drawing upon the different inspirations that contributed to this work's composition... When I performed that she was on my panel, as well as another teacher whose training background was from Versailles in France… He had actually a lot to say about that piece... what more I could make of it as a musical demonstration in terms of, often, a lot of contrast and colour and character. Very different interpretive contexts… I think I recall he marked me lower for that piece because I think he felt, interpretively, I could've demonstrated more as a piece of music. Even though perhaps the technical aspects of... there were very technically difficult things in that piece that were done correctly, but I think he felt, interpretively, that [the piece] was lacking. So there's that little bit of separation. (Student A, a conservatoire percussion student)

---

[17] A shorter excerpt from this quote was included earlier in Section 3.3.1. In that section it was used to highlight the importance of social relations in framing the second research question. Its purpose here is different—this is clarified by the analysis that follows.

Example 3: I remember one instance where I was on a panel of, I think it was three people, for a violin recital… It was interesting that myself and another non-string person were really captured by the musicality of this performance, and it was really engaging. The whole thing from start to finish really had us engaging in the music… engaging in the world that they were presenting, in the genre they were presenting. The string specialist at the time wanted to mark down significantly because they knew the piece so well. They said, 'oh, [the student] didn't use this fingering. Missed out that. That phrase wasn't executed properly', etcetera, etcetera. It was quite a heated debate for a while, and it wasn't one of those ten [to] fifteen minute discussions after a recital... It went on for almost half an hour, forty minutes. It was a long discussion. (Assessor J, A multi-disciplinary department leader)

Each of these quotes illustrates a kind of clash between perspectives on students' performance work—it is clear that different criteria may be accorded different value by different people in different contexts. Determining the underlying bases for criteria in terms of epistemic relations and social relations enables one form of analysis of the kinds of clashes that might occur between these interpretations. Construed on the Specialization plane, contrasting emphases on epistemic relations and social relations can be realised as code clashes (Figure 6.2). Example 3, for instance, illustrates a clash between a knowledge code (ER+, SR-) and a knower code (ER-, SR+): The two perspectives represented in this quote reflect opposing emphases on musical technique (ER+) and holistic interpretation (SR+). The clash in the second example is less pronounced, but also includes a technical position (ER+, SR-) and an interpretive position (ER- SR+). The basis of both perspectives in the differing musical tastes of the assessors indicates an overall common ground in interpretation (social relations) for both perspectives represented, hence the clash can be analytically distinguished from the first as occupying a stronger overall position in social relations.
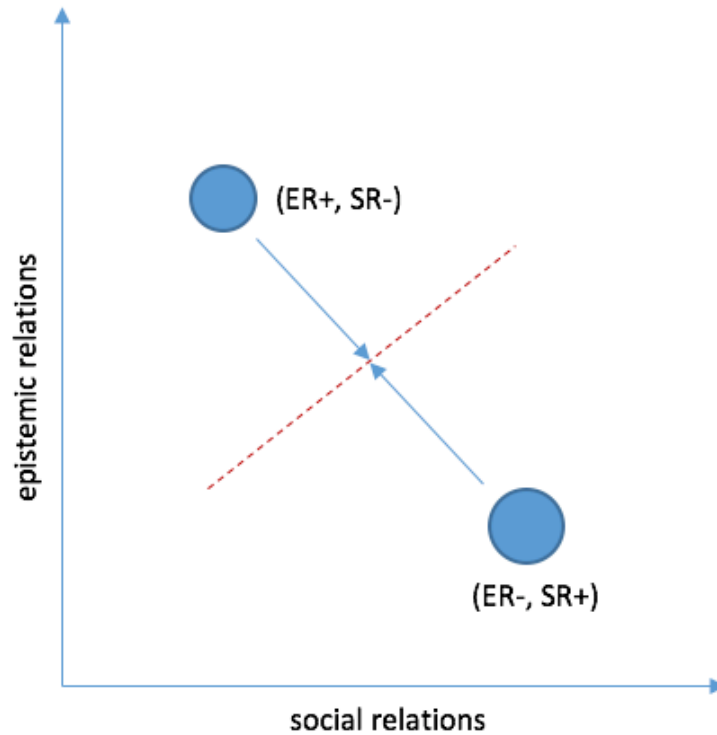
Figure 6.2 A visual depiction of a code clash between knowledge and knower codes

6.4.4 Summary

This part of the chapter considered data collected from interviews with assessment participants (including student administrators) through the lens of the translation device developed in Section 6.3. This section complements the device by realising the concepts in substantive data. Where perspectives from the literature (Section 6.3) tend towards the allocation of singular meanings to criteria, the interview data highlight that criteria are more nuanced—they can be analysed in terms of both epistemic relations and social relations to interrogate the more specific positions encapsulated within individual concepts: the "ability to convey an idea" (Assessor B), for example, can subsume both technical and social characteristics. Further, the underlying bases of criteria may differ depending upon the broader contexts of the quotes from which they are derived. This means that a concept which would be ambiguous in isolation or otherwise allocated a single analytical position—for example, musicality—can be interpreted more meaningfully as having a basis in one or both sets of relations depending on the context in which it is invoked.

6.5 Making Use of the Concepts

This chapter has so far been focused on realising a device for conceptualising criteria in musical performance practices using the Specialization dimension of LCT. Section 6.3 began the development of this device, drawing on a sample of criteria from musical performance assessment literature. Section 6.4 contextualised perspectives from research participants using these concepts to illustrate their application in the analysis of data. This part of the chapter considers how the conceptual perspective adopted from the Specialization dimension of LCT provides some useful analytical tools that can be put to practical use in music education settings.

6.5.1 Locating the space of practice

Sections 6.3 and 6.4 have explored a range of possible orientations to Specialization within discourse on musical performance assessment criteria. The focus here is on the relative positioning of these orientations. Clashes (Section 6.4.3) signify the presence of contrasting knowledge and knower codes, and these can be further located by considering the general borders of the theoretical space within which these positions are be taken up. The basic value for technical skills (ER+) throughout provides one possible boundary that distinguishes musical performance assessment practices in higher education from other possibilities outside this context. Similarly, the general value placed upon factors such as taste, interpretation, and engagement points to a broader emphasis on social relations (SR+). Taken together, these two general observations point to the presence of an élite code underlying the general context of musical performance assessment (Figure 6.3). Locating the examples developed previously as finer-grained positions within this broad region helps to explain how practices and practitioners can move forward from a common ground to occupy more specific locations and orientations. In other words, distinctions between knowledge and knower codes, highlighted by code clashes, are nonetheless predicated much of the time on an underlying system of value for constituents of both epistemic relations and social relations (see Figure 6.3). Within this general space, the basic modalities of the dimension can be reconceptualised to reflect the context of musical performance assessment, where

- *knowledge codes* (ER+, SR-), reflect an emphasis on musical techniques, skills, and understandings of concepts (for example, stylistic interpretation); the attributes of performers are less important
- *knower codes* (ER-, SR+), reflect less emphasis on musical techniques, skills, concepts, and the attributes of performers are emphasised as the basis of achievement
- *élite codes* (ER+, SR+), reflect an emphasis on both social attributes and the possession of particular techniques, skills, and conceptual understandings
- *relativist codes* (ER-, SR-), specify something other than social attributes or the possession of skills techniques, and conceptual understandings as the basis of achievement
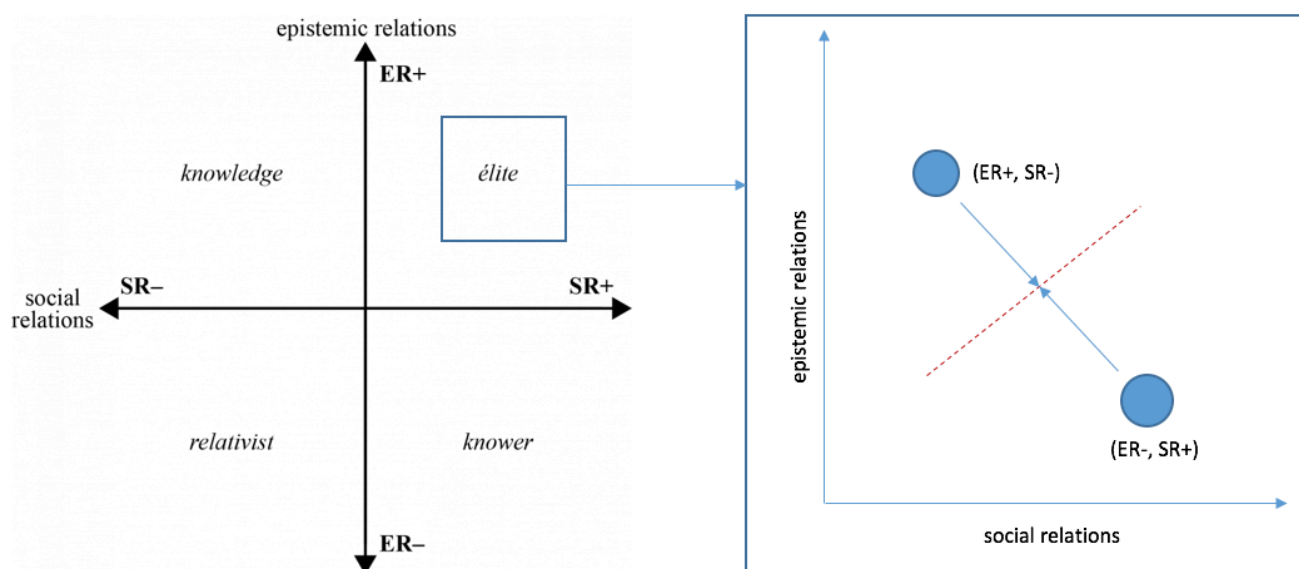
Figure 6.3 A code clash (see Section 6.4.3) located within a broader élite code

6.5.2 Specialization pathways

In addition to clashes between codes, *shifts* and *tours* (Maton, 2014; Maton & Howard, 2018) between codes usefully illustrate movement between positions over time. One possible shift reflects the move from a knowledge code—in which techniques and conceptual understandings are gatekeepers to achievement—to an élite code, wherein the holistic experience of the music is accorded the most importance (Figure 6.4). To illustrate how this kind of logic can underpin broader curriculum, consider the following set of criteria specified for musical performance coursework at one Australian institution where these

criteria are repeated for subsequent iterations of a *Practical Studies* series of courses. They state:

> You will be marked on how well you can:
>
> 1. Demonstrate a level of musical craftsmanship* relevant to the stage of your degree
>
> 2. Demonstrate a level of musicality and artistry^ within a solo and/or collaborative performance context relevant to the stage of your degree
>
> *Musical Craftsmanship to include: Rhythm; tonal quality; posture; presentation; flexibility; memory; cantabile; touch/bowing/articulation, timbral variation, economy of motion, intonation, breathing, direction (to be applied as relevant to each performance area).
>
> ^Musicality and Artistry to include: Expressive control; grasp of style; interpretive capability; dramatic/performance flair; program construction; sophistication in performance; musical insight. In Years 1 and 2, greater importance will be given to Basic Craft. In later years, Musicality and artistry will be accorded increasing importance. (Griffith University, 2020b)

Similarly, the practice of distinguishing and alternating between technical examination type tasks and performance tasks more broadly focused on art-making (for example, recitals) can underlie shifts in the dominance of different codes. Perspectives from interviewees highlighted ways in which these broader movements can have significant effects. A classical performance lecturer at one institution, for example, explained:

> I don't think that even the most clever teacher could really come up with a way to do our way of examining in the way that I would like to teach, where you're actually concurrently doing the technical work that supports your… technical development, such that you can then… apply that to musical performance. You

know, it's the idea of working on one's craft such that their craft can enable them to make art, and it's just like craft, craft, craft, craft, craft, craft, craft. This is boring. And then it's like art, art, art, art, art, and then often, the house of cards falls over. You know, you see… string players in particular get repetitive strain injuries and you've got brass players with embouchure problems, and singers, you know… they're sick… you're not taking care of the nuts and bolts, because you're like overwhelmed just trying to learn all these sonatas and whatnot… so I think it's too much music. Too much repertoire too really, particularly for first years, and the dissociation of craft and art… My whole pedagogical approach is really working on students' craft such that they can be musical, but I think divorcing them is dangerous. (Assessor L)

Locating this issue in a topological conceptual space (Figure 6.5) provides a means of visualising an underlying structure that may otherwise remain only partially evident through the visible influence of the effects produced by those structures. In addition to providing a framework for the analysis of extant practice, calibrating the framework to context also enables generative possibilities (Figure 6.6) which can be used to actively challenge issues in existing practices by highlighting alternative possibilities. In other words, this approach offers one means of mapping a range of possible structures for curriculum and assessment over time, and for highlighting the kinds of specialisation provided and assessed within music education programs.
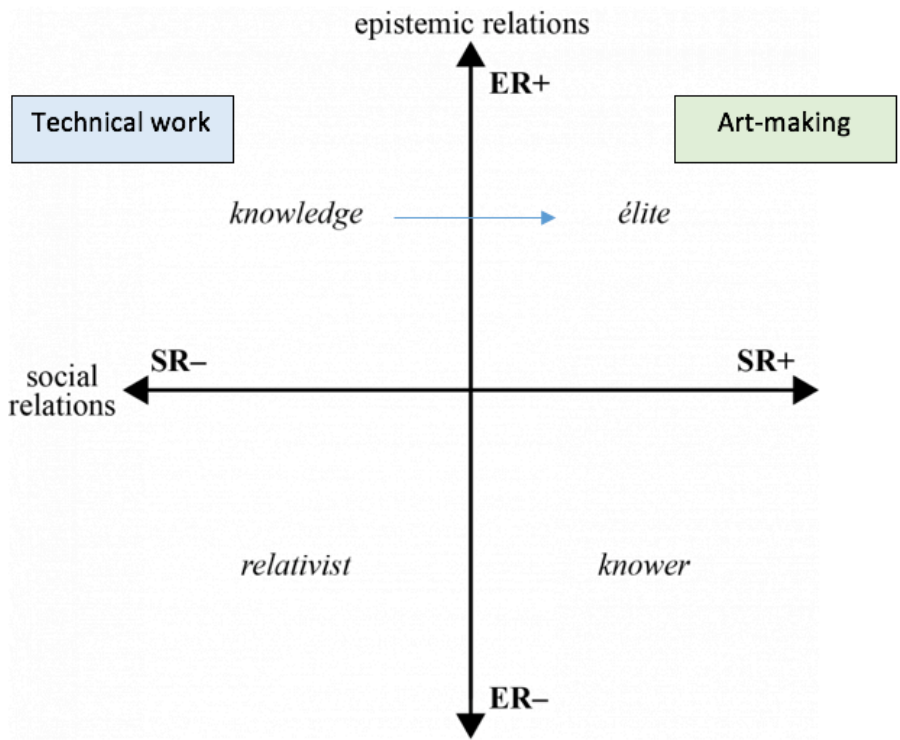
Figure 6.4 A code shift from a knowledge code to an élite code over time
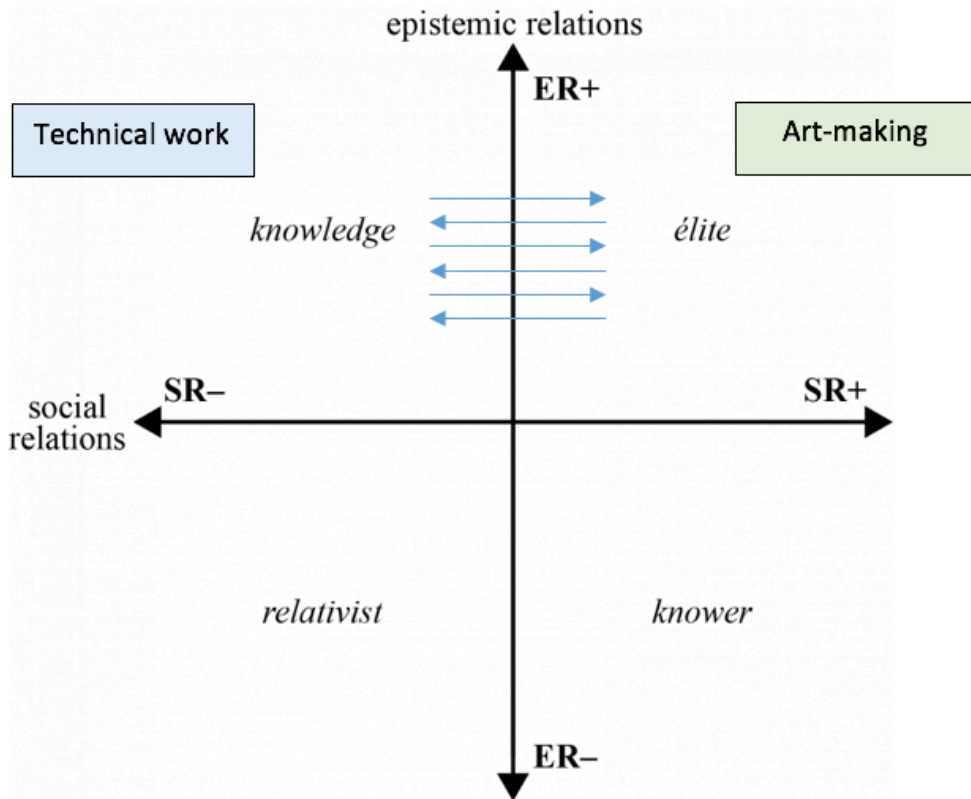
Figure 6.5 Return trips between codes over time
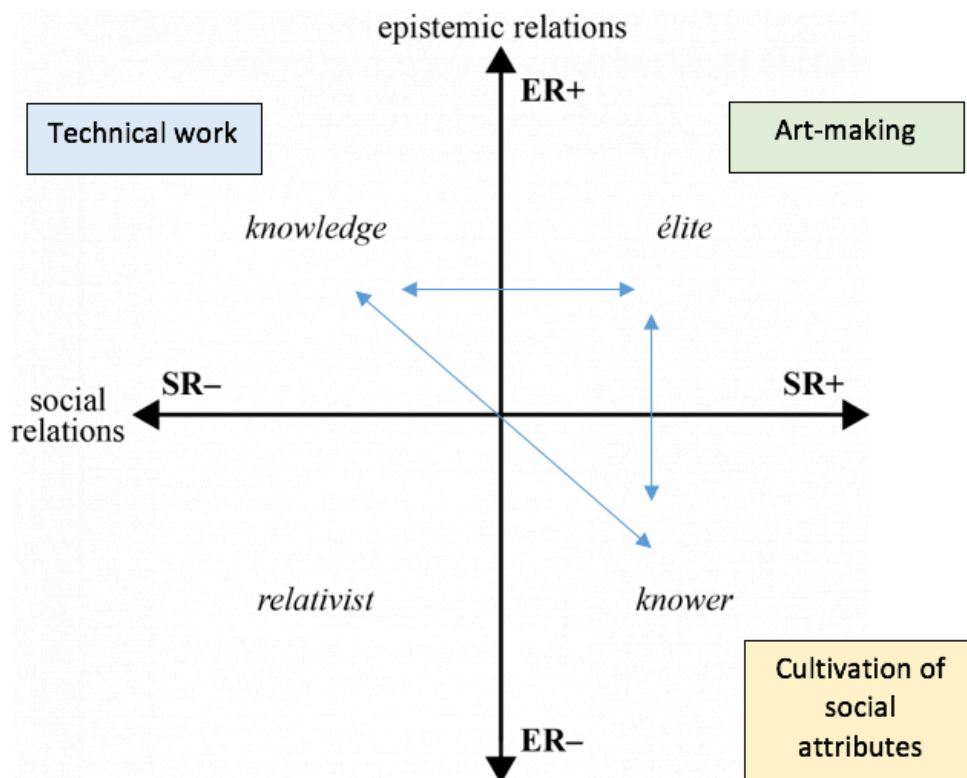
Figure 6.6 Space of possibilities for the specialisation of musical performance curricula and assessment

6.5.3 Implications in a qualifications framework system

The more specific codes that underlie musical performance practices reflect regions of practice embedded in a broader educational landscape that includes other kinds and breadths of claims about learning, teaching, and assessment. At the broader level of higher education, Wolff and Hoffman (2014) assert that "[h]istorically, the two relations (epistemic relations and social relations) illustrate the divisions between the natural sciences and the humanities" (p. 81). The authors further suggest that "no matter what the discipline, the acceptance by Higher Education of the responsibility to inculcate Graduate Attributes in 21st century curricula firmly establishes the increasing importance of knower code aspects, the traditional focus of the humanities" (Wolff & Hoffman, 2014, pp. 81–82). In the Australian context, as elsewhere, graduate attributes are important for the reason that they reflect claims about the outcomes that graduates of higher education qualifications will have achieved (Allais, 2012a, 2012b; see Chapter 2). These types of statements typically include broad references to both skills and knowledge (a basis in epistemic relations) and social attributes (a basis in social relations). A set of graduate

attributes from one institution (Griffith University, 2016b), for example, asserts that graduates of all programs should be:

- Knowledgeable and skilled, with critical judgement
- Effective communicators and collaborators
- Innovative, creative and entrepreneurial
- Socially responsible and engaged in their communities
- Culturally capable when working with first Australians
- Effective in culturally diverse and international environments

In Australia, these claims are important at the level of individual coursework units for the reason that the specific learning outcomes determined for a unit of coursework are typically linked to the broader graduate attributes set by institutions. The correlative is a claim for both epistemic relations and social relations as bases for achievement at the levels of both whole programs of study as well as for individual coursework units. The learning outcomes determined for an individual unit of coursework are usually more discipline-specific and related to the substantive content of the given unit of coursework. Below is an example of a set of learning outcomes that guides musical performance coursework at one institution. Upon completion of the given coursework students should be able to:

1. Present reliable technique supported by appropriate practice habits and diverse pedagogical knowledge
2. Express musical and stylistic sensitivity within a solo and/or collaborative performance context
3. Produce an engaged performance incorporating musical creativity and personal style
4. Critically evaluate personal musical progress as an emerging and lifelong learning musician (Griffith University, 2020a)

These unit-level outcomes are usually linked individually to graduate attributes in course documentation—they are effectively taken as realisations of the broader attribute in question. At a finer-grained level again, different assessment tasks will typically be purported (in course documentation) to

reflect a selection of outcomes, such that a student who passes a given piece of assessment will be held to have demonstrated achievement of the related outcome.

Navigating the various gradations of translation involved between the levels of learning outcomes that govern higher music education in qualifications framework environments is a complex process that has strong implications for the integrity of coursework, both from a regulatory standpoint, and from the standpoint of delivering on educational claims about what students are gaining through their purchase of education. Part of this relates to the complexity of codifying achievement standards that are by definition abstract (Sadler, 2014). Scholars such as Sadler (2014) have argued that meticulous description of achievement standards and outcomes places us at an unavoidable loggerhead with the inability of these statements to actually capture what it is they seek to describe. In response to this issue, one possible way forward involves leveraging theoretical tools to overcome the descriptive gap. In offering a means of theorising specific elements of practice in an explanatory fashion (see Chapter 2, Section 2.3), this approach offers a resource for supporting the calibration of standards in a given context by providing a language of description for the underlying principles of standards and outcomes rather than the standards and outcomes themselves. The proposition here is to remove the issue of objective categorical specification by looking at the problem differently: Rather than focusing on which criteria to choose from an infinite range of possibilities and combinations, focusing on the underlying principles that inform the selection of criteria provides a means for retaining an authentic latitude in the expression of criteria while retaining a transparency about the bases of achievement. The materials developed in this chapter offer a *starting point* for this kind of calibration in music education contexts, and the approach explored here is simply one possible response to the research problem. While it offers some concepts that could be employed in practice, the bigger contribution of the exploration is to illustrate an alternative means of conceptualising the space of possibilities for assessment criteria that takes on a more topological flavour and thereby offers a greater degree of nuance for interpreting bases of achievement in practices.

6.5.4 On academic non-achievement

Returning to the subject of non-achievement briefly introduced earlier in this chapter (Section 6.2.1), distinguishing bases of achievement from processes of legitimation is one means of explaining difficulties in establishing trustworthy assessment practices. Specialization codes highlight possible bases for achievement in social practices, however they do not in themselves make any sort of claim about which codes are legitimate or dominant in a context—they simply represent a topology of possible more or less legitimate positions. The legitimation of particular bases of achievement is a separate (though intimately related) process, and many examples of legitimating language have been provided throughout this chapter. Academic non-achievements, in this view, can be construed as a by-product of co-occurring or competing acts of legitimation where different bases of success are valued in different ways by members of the same assessment community. In Australia, for example, it is not considered acceptable to assess students on the basis of their effort, enthusiasm, or progress, from a regulatory standpoint, unless effort, enthusiasm, or progress are included in relevant course learning outcomes together with indications of how standards will be applied to their assessment. Comments from interview participants in this research, however, communicate a clear value for these qualities. Making the bases of achievement available for assessing students' work explicit seems likely to be a useful step by which to interrupt inconsistent assessment practices, however the tools discussed in this chapter are not necessarily adequate for that task on their own. The reason for this is that both legitimate and illegitimate criteria may be characterised by the same Specialization code. The analysis conducted in this chapter, for instance, positions criteria related to both personal interpretation and effort as having a basis in social relations. From a regulatory standpoint, personal interpretation can serve as a legitimate criterion while effort cannot. The issue here is less a question of achievement basis and more so a question of legitimation for elements of assessment practice (in this case, criteria selection). To this end it is a departure from the focus of this chapter, however this issue is considered further in Chapter 8 which is more directly focused on conceptualising the legitimation of aspects of assessment practice itself.

6.5.5 On the relativist code

The relativist code has received little attention in this chapter. It is an interesting theoretical space for the reason that it prompts the question: What are possible bases of achievement that reflect a weaker emphasis on both knowledge and skills (epistemic relations) and peoples' attributes (social relations)? The work of McPherson and his colleagues (McPherson & Thompson, 1998; McPherson & Schubert, 2004) provides us with some possibilities: for example, environmental conditions, or performance order. These have little to do with either knowledge or skills (ER-) or social characteristics (SR-) reflected in students' performances. Another useful example relates to students' attendance during assessment events. Although an unexplained failure to present for a formal assessment generally results in a failing grade, the basis of achievement (or failure, in this case) has little to do with the student's abilities or social characteristics (ER-, SR-); rather, the outcome is determined by a mechanism of local policy. In some situations this is not a bad thing—to receive a formal higher education qualification necessitates participation in assessment tasks that are designed to provide insight into students' achievement of particular outcomes. This analysis also helps to explain the tenuous validity of some non-achievements: for example, curricular designs wherein students are graded on the basis of their attendance or participation in particular classes or other learning activities rather than their completion of work that provides insight into their attainment of learning outcomes. Although the relativist code should not be conceived of as an inherently negative space, in the context of assessing students' learning it can help to highlight principles of assessment that lead to spurious grades in that it may prompt us to ask the following question: If an assessment design or event downplays both knowledge (epistemic relations) and knowing (social relations), what might be grading processes actually be based upon? One possibility here is the application of something like a bell curve within a normative-referenced assessment approach (see Section 3.4.3.1), wherein students' achievement (as symbolised by grades) is influenced by some factors unrelated to their demonstrated knowledge or attributes. This example highlights that different codes may be used to theorise *parts* of an assessment design, as distinct from overall designs.

6.6 Summary

This chapter explored a response to the first guiding research question of the study. The focus in this chapter was on meaningfully conceptualising bases of achievement for musical performance assessment. To this end, this chapter described the means by which concepts from the Specialization dimension of LCT were employed to theorise criteria for musical performance assessment.

Section 6.2 outlined the main theoretical points of departure for the chapter. While a range of approaches have been taken in conceptualising bases of achievement for musical performance work, approaches are commonly limited by their transferability between substantive contexts (due to rich contextual description) and their capacity to actually describe practices (in the case of typological models, see Chapter 3). To explore a possible way forward, concepts from the Specialization dimension of LCT were employed as tools for conceptualising musical performance assessment criteria in terms of epistemic relations and social relations. To this end, Section 6.3 discussed a means of translating between these concepts and more substantive criteria that more closely reflect real-world terminology. A translation device was developed to illustrate the realisation of epistemic relations and social relations in substantive disciplinary forms and vice-versa. This device enabled a depiction of musical performance criteria as a field of specialisation wherein achievement is regulated by the dominant Specialization codes. Locating which codes are emphasised or downplayed provides a means of explaining the underlying bases for assessment practices. The translation device developed in Section 6.3 is offered as a starting point for enacting Specialization in analyses of musical performance assessment criteria. It informs the analysis of interviewee perspectives on musical performance assessment criteria developed in Section 6.4.

In applying the concepts from the translation device to interview data (Section 6.4) a greater level of conceptual nuance is made available to the interpretation of the device. This section establishes further empirical realisations of social relations and epistemic relations to enable the theorising of code clashes (Section 6.4.3). The purpose of this analysis is not to offer an absolute theoretical characterisation of criteria for musical performance

assessment, but rather, to make empirically visible the linkage between the theoretical concepts and data (in this case the perspectives of interview participants). This in turn enables the discussion of further potentials for the framework (Section 6.5) including

- A theorisation of higher musical performance education as underlain by an élite code
- A means of translating fuzzy concepts (for example, flair, originality) across disciplinary and sub-disciplinary contexts
- The realisation of Specialization pathways in musical performance curriculum, and the generation of alternative possibilities
- The explication of bases of success for both students and educators
- Implications for navigating practice within the auspices of qualifications frameworks, including conceptual clarification of what is being assessed, and how this aligns with broader educational outcomes

The main contribution of this chapter lies in the conceptual generalisability of this discussion. The material developed in this chapter provides a general starting point for enacting Specialization in music assessment contexts, as well as providing some suggestions as to how these concepts can be used to realise and reflect upon practices. Although the specific characteristics of individual contexts are invariably unique, the devices and perspectives developed here are general enough to be moulded to more specific contexts or investigated further in subsequent research studies. These ideas provide a starting point for an acute application of the framework at a deeper level in a more tightly specified context—for example, a single departmental area in an institution. For substantive studies in tightly defined contexts, the concepts offered here can be refined to enable a higher level of nuance.

Chapter 7: How Who You Are Matters

7.1 Introduction

This is the second chapter of the thesis to discuss the outcomes of the research. It is a response to the second guiding research question: How can bases of legitimation for assessment participants be meaningfully conceptualised? The context for this question was described in Chapter 3 (Section 3.3). Where the previous chapter focused on criteria for assessing students' work, this chapter shifts the focus to the people involved in musical performance assessment practices. More specifically, this chapter explores how *who you are*—as a student or an assessor—can matter in musical performance assessment practices.

This chapter is structured similarly to the previous one, and begins with an overview of the main theoretical points of departure for the discussion that follows (Section 7.2). Section 7.2 explains how assessment participants' characteristics can be interpreted using the concepts of epistemic relations and social relations, which were elaborated on in the previous chapter. Here, they offer a means to distinguish between the embodied characteristics, capitals [18], and resources—collectively, *attributes* [19] —that assessment participants bring with them. Sections 7.3 and 7.4 distinguish *epistemic attributes* and *social attributes* to illustrate ways in which participants' attributes can help or hinder their ability to effectively participate in assessment practices. In so doing, Sections 7.3 and 7.4 offer some simple theoretical perspectives to make visible bases of legitimation underlying this participation. The final part of the chapter (Section 7.5) takes a generative stance, and adopts concepts from LCT to sketch out an approach for making explicit the kinds of *gaze* adopted and embodied by assessment participants which have implications for successful participation in assessment. This section discusses how making these gazes explicit offers a means by which business-as-usual

---

[18] See Chapter 3, Section 3.3.1.

[19] The notion of assessment participants' attributes as it is used here bears some similarity to Biggs' (1993) concept of presage, which refers to the factors related to students and teachers that are in place before learning and teaching occurs. Connections between the concepts are discussed further in Chapter 9.

can be productively interrupted in assessment contexts by highlighting the ways in which participants' attributes influence their involvement in practice.

7.2 Theoretical Points of Departure

The research question that guides this chapter emphasises the *legitimation* of assessment participants. In this, it recognises that assessment participants are not equally equipped or otherwise positioned to successfully participate in assessment practices. As outlined in Chapter 3, the notion of *successful participation,* as it is interpreted here, is broader than the getting of good grades for students, or of grading accurately for assessors. Rather, it is about how well people are actually able to fulfil (or are allowed to fulfil) their various roles and intentions in assessment contexts. The discussion developed here is focused mainly on two kinds of assessment participants, students and assessors.

The theoretical points of departure for this chapter build on those introduced in the previous chapter. In Chapter 6 the Specialization dimension of LCT was used to distinguish a space of possibilities for conceptualising assessment criteria, produced through relative emphases on technical skills and understandings (epistemic relations) and performance factors related to taste or other social attributes (social relations). In this chapter, Specialization is revisited to conceptualise the kinds of attributes possessed by assessment participants (including skills, knowledges, traits, taste, philosophies, and so forth), and how possession of these attributes can influence participants' abilities to successfully participate in practice.

7.2.1 A focus on social relations

The concept of *social relations* (Chapter 3) highlights the attributes of people as a basis for legitimacy in social practices. These attributes themselves can express bases in both epistemic relations (knowledge, skills, techniques, etc.) and social relations (taste, traits, identities, dispositions, gender, etc.). Therefore, while the *focus* in this chapter is on the attributes of people (a focus on social relations), the analytical interest here is in the underlying *bases*[20] of

---

[20] Explanations for focus and basis were provided in Chapter 4, see Section 4.3.

the attributes in question (epistemic relations or social relations). Organising the possible constituents of participants' attributes according to these bases produces a pair of conceptual *constellations* (see Section 4.3.5) which are interconnected regions of conceptual territory where the constituents of a constellation are unified by a common theme and relate to each other in particular ways (Maton, 2014). A constellation is a useful conceptual metaphor for the reason that it recognises that conceptual territories are given metaphorical shape by the relationships between the constituents of those spaces (hence, constellations instead of clouds). This chapter is principally concerned with two kinds of constellations that represent the attributes of actors: an *epistemological constellation* that brings together attributes underpinned by epistemic relations, and an *axiological*[21] *constellation* that brings together attributes underpinned by social relations.

An important distinction between epistemological constellations and axiological constellations is that their structural qualities are not necessarily the same. Where epistemological constellations are typified by "integrated sets of technical, 'content' meanings describing and explaining an object of study" (Doran, in press, p. 11), axiological constellations "involve nuanced sets of values, positions, evaluative stances and interpretive frames" (Doran, in press, p. 11). The implication of this distinction is that the ways in which the constituents of the different constellations relate to each other (within the respective constellations) may differ. Doran (in press, 2019) provides a useful illustration, contrasting meaning-making in physics with meaning-making in poetics[22]. Doran (in press, 2019) shows that where physics builds meaning through an integrated set of well-defined concepts (an epistemological constellation), poetics builds meaning through the positioning of value-laden positional language (an axiological constellation). A broader observation of this work is that epistemological constellations may tend toward taxonomic structures, while axiological constellations do not necessarily comprise

---

[21] Axiology is a philosophical discipline concerned with values—the adjective term *axiological* is often not explicitly defined in studies that use it; here it can be considered in the broadest sense as pertaining to values, however in the context of LCT it generally signifies an association with the concept of social relations.

[22] The study of poetry and/or other literary forms

relational structures in the same way, relations between concepts being generated instead through an interplay of values and beliefs (Doran, in press, 2019).

The relevance of this distinction emerged through the research, as it became clear that interpreting epistemic attributes and social attributes required slightly different analytical approaches. In studying assessment practitioners' epistemic attributes (Section 7.3), a pair of epistemological constellations emerged, one relating to musical knowledge and skills and the other relating to knowledge and skills for educational assessment. These constellations were relatively straightforward to organise, and reference more definable bodies of knowledge and more defined relationships between those bodies. Studying assessment participants' social attributes (Section 7.4) was initially more challenging for the reason that distinct conceptual themes were less evident. Simply mirroring the analytical approach led to broad ephemeral themes that were difficult to interpret meaningfully. To solve this issue, some finer-grained concepts from the 4-K model (introduced in Chapter 4, Section 4.3.4) offered a means of revealing subtler distinctions—these theoretical tools are revisited in Section 7.4. More broadly, however, repeated analytical attempts highlighted that for this type of research analytical strategies are not one-size-fits-all, perhaps for the reason that objects of study with different underlying bases may require different analytical approaches to achieve meaningful outcomes.

7.3 Epistemic Attributes

The concept of epistemic relations highlights the possession of particular knowledge, skills, techniques, and technical understandings as a basis of legitimation in social practices. Here, the interest is in the degree to which legitimate participation in assessment practices is influenced by the possession of more or less specialised knowledge and skills. Where possession of specialised knowledge or skills is important, epistemic relations are stronger (ER+), and where less specialised knowledge or skills are necessary to legitimately participate in practices, epistemic relations are weaker (ER-). The previous chapter argued that some musical performance assessment practices in higher education are informed by an underlying

emphasis on technical knowledge or skill (see Section 6.5). This is evidenced in part by the prevalence of expert musical performance practitioners being engaged as assessors. Conversations with interview participants in this research highlighted that possession of disciplinary knowledge could have a substantial influence on the ways in which assessment participants view one another (as evidenced in the previous chapter, see Section 6.4.3 for example), or the ways in which they perceive assessment approaches themselves. A classical performance lecturer at one university explained:

> I'm pretty comfortable now knowing where our students sit in relation to our cohort on the [instrumental area taught by assessor L], and I struggle a little bit more when I get to the other areas, even instruments that I know intimately well. My [partner plays another instrument] so I hear [that instrument] every single day of my life. I know what's good, I know what's bad. Do I know what's a 73 as opposed to 72? Not so much… We're musicians. We know what music is… I can objectively say "oh, your rhythm was bad," or "isn't it funny how you played Bach like it was Jobim,"… But I don't know with guitar… There was a famous time where there was a guitarist playing and a different… teacher was listening… It was good… the [non-guitarist] said, "well yeah, but they have frets. Isn't that easier?... [They're] excellent musicians, but [don't have] a super in-depth understanding of instruments other than their own. So the idea of having a specialist on there is important. (Assessor L)

Part of the complexity of assessment in higher education is that a particular set of disciplinary knowledge and skills may not be the only set of knowledge and skills that are relevant in assessment contexts (Dawson et al., 2013). The data collected throughout the project collectively suggested two prominent epistemological constellations underlying participation in musical performance assessment practices. The first of these constellations constitutes the network of musical knowledge and skills that are available to assessment participants, while the second of these constellations constitutes a network of knowledges

and skills related to educational assessment itself. To this end, an *epistemological musical constellation* can be distinguished from an *epistemological educational assessment constellation*. Table 7.1 summarises signifiers (see Chapter 4, Section 4.3.5) for this pair of constellations identified in the data. The interviews conducted through the research in particular highlighted distinctions between the two in striking ways. Two examples are discussed here to situate the differences between these constellations in the context of assessment participants' perspectives.

Table 7.1 Two epistemological constellations

Musical knowledge and skills	Assessment knowledge and skills
Technical knowledge of music theory, e.g.: • Harmony • Dynamics • Intonation • Rhythm Knowledge of particular instruments, e.g.: • Bb Trumpet (e.g. fingerings and tuning considerations) • Cello (e.g. bow and vibrato techniques) • Piano (e.g. reasonable hand reach) Knowledge of musical idioms and styles, e.g.: • Distinctions between the features of Dixieland and Bebop styles of jazz music (as written about in history books on the subjects) Ability to perform in particular ways, e.g.: • Ability to pitch accurately • Ability to play in time • Ability to hear harmony • Ability to perform with stylistic accuracy Combinations of the above, e.g.: • Idiomatically accurate jazz improvisation on the Bb trumpet	Understandings of assessment concepts, e.g.: • Criteria • Feedback • Formative/summative assessment • Grade integrity • Standards • Validity Ability to construct assessment tasks using assessment concepts, theories, and tools, e.g.: • Continuous assessment • Peer assessment • Self-assessment • Holistic/analytic assessment • Rubrics • Constructive alignment Ability to deliver feedback in different ways, e.g.: • Written feedback • Verbal feedback

7.3.1 Example 1: Musical knowledge and skills

The previous chapter highlighted the range of ways in which musical performances can be specialised. For students, the degree to which their performances emphasise technical capability (epistemic relations) or tasteful artistic choices (social relations) can influence their ability to successfully participate in assessment practices. The kinds of musical resources students may be able to draw upon are personal to each student, and accessing these resources at appropriate times is one means by which students are able to succeed (either in performing or in other capacities). As one department leader summarised:

> I think there's two types of students. I mean there's actually a bit more than that, but if you want to really simplify it. There's the student that comes in who plays with good enough technique and plays accurately enough that they tick certain boxes… they do reasonably well in exams but sometimes they're a bit limited because of their expressiveness, or their creativeness, or musicianship. Then you have the other student that perhaps technically is not so great but has more imagination and more creativity… You of course don't get necessarily a black and white situation. Often you get… students with a bit of this and a bit of that and crossover. Sometimes you'll have two students walking in to do assessments, say two flute students… After they've both finished [you think] gee actually if you took the best of that one and that one and put it together you'd have someone really fantastic. You're always weighing that up and I think sometimes that—not always—but I think sometimes students who have a neat technique often get favoured a little bit. Sometimes I think the students who have more imagination and creativity don't get recognised as much as they should. (Assessor M)

Some research participants highlighted that a similar logic can apply to assessors, as illustrated at the beginning of Section 7.3. Other participants also highlighted ways in which possession or lack of particular musical

knowledges could influence the ways in which assessors are viewed. Some of the quotes included in the previous chapter offered simple examples of inter-assessor judgement. Others, however, pin-pointed the issue of musical knowledge and skills more specifically. For example, a classical performance teacher at one institution commented:

> I feel very strongly as [an instrumentalist] that it's a great disadvantage to [instrumentalists] and to their marks if they're not assessed by a specialist… I think that that is a real flaw at a lot of tertiary institutions because they don't want people to assess their own students, so they put in someone from another instrument… As [an instrumentalist, my instrument] is first of all highly idiomatic. We don't have the big wonderful standard repertoire that a lot of other instruments have. It's almost impossible for a [different instrumentalist] to tell whether a piece is difficult or not… You find cases of people having good students and giving them a really difficult piece and the piece is very [outwardly] unimpressive, and the student gets a really bad mark… The student that played the [outwardly impressive piece] who actually can't play very well—and just got that piece because it's easy—gets a better mark. I know that happens a lot… The teacher has to say 'hang on guys'… I'm really glad that the way we do things [at this institution] is that the specialist of that instrument is assessing that instrument. (Assessor D)

7.3.2 Example 2: Assessment knowledge and skills

Where the data collected in the course of this project provided some acute examples of the importance of musical knowledge, references to assessment knowledge were less explicit. For the interview participants, one possible reason for this is that, in contrast with musical concepts, their access to educational assessment concepts may have been more limited. Many interviewees discussed topics for which conceptual terminology exists in the field of assessment research (for example, authentic assessment, and the relative merits of analytic and holistic approaches to grading), however very

few participants indicated any kind of awareness of this terminology. One reason for this may have to do with the experiential and inherited manner in which educators often develop their knowledge of assessment (Bloxham & Boyd, 2007; Boud, 2010). Another factor in this may be the kinds of access to knowledge provided through various employment situations. As one casual staff member in a jazz department observed:

> I'm in two minds about it [assessment] especially given that I'm a casual staff member… I only mentioned that because maybe if I knew more about it, maybe if I was in a full time role like [a department leader]… When you can see things in a bigger picture [with] all the parts that go into a degree… I don't know whether it's because I've got this lovely bubble [where] I go in and I teach and I try to support the student and then leave… Then I go, 'oh it's such a shame that they, feel like crap when it comes to assessment'. But maybe… when you're around it all the time in a full-time role and you work in the university and know how it all works maybe… Maybe it's just normal? I don't know. (Assessor C)

The growing field of assessment literacy (see Chapter 3, Section 3.3) highlights a broader professional concern for assessment practitioners' possession of knowledge and skills relating to educational assessment concepts. Although frameworks geared towards musical performance assessment in higher education settings do not yet appear to be readily available, assessment literacy standards for broader music education have been advocated for by scholars such as Brophy (2019). The value of assessment literacy was highlighted in comments from several staff members from a jazz department at one institution who noted that access to assessment skills could influence self-perceptions and confidence:

> Well, I used to really worry about, you know, (pause for thought) being able to, say all the right things and tick all the right boxes and give the absolutely the right mark… Knowing that I'm responsible for every single thing I do, right, used to be quite

worrying, now I'm starting to trust my own, (pause for thought) basic intuition… like I was saying to you earlier you know if someone can play. (Assessor A)

I was quite, nervous about meeting you because I'm like, 'wow, what if I'm doing all the wrong things?'. (Assessor C)

Students are also active participants in the constellation of knowledge and skills related to educational assessment, and access to this knowledge can influence they ways in which they participate in assessment. This is particularly apparent in the evidence that students adopt strategies in assessment (see for example, Barratt & Moore, 2005). Although assessment literacy appears to have been less focused on students, the recent emergence of feedback literacy (Carless & Boud, 2018) bridges this gap by placing a more direct emphasis on students' capabilities in making use of feedback. Additionally, participatory forms of assessment (for example, peer and self-assessment) are now characterised by ever-growing bodies of knowledge.

7.3.3 Conceptualising legitimation

The examples included above illustrate just some of the ways in which possession of particular knowledge and skills can influence participation in assessment practices by acting as points of legitimation for those involved and mediating access to different parts of given constellations. Systems of legitimation and access are complicated by the presence of overlapping constellations of knowledge and skill, and so the two discussed here are likely to be overlaid with many others. Conceptualising these constellations in terms of relative strengths and weaknesses of the common underlying basis (epistemic relations) provides one means of analysing and interpreting the bases of legitimation in a given context (Table 7.2). The translation device developed here reflects four possible points along continua of epistemic relations, and some examples are provided derived from the topics discussed so far. Differences across contexts—in local procedures, values, assessment design, and so forth—will inevitably influence legitimation processes, and as with all specific translation devices provided in this thesis the usefulness of this device is as an illustration and a starting point for further application in specific

Table 7.2 A translation device for two epistemological constellations

Strength of epistemic relations	Musical knowledge and skills	Manifestations in data	Assessment knowledge and skills	Manifestations in data
ER++	Legitimacy is contingent on specialised knowledge, e.g. a specialist understanding of a musical instrument	I feel very strongly as a guitarist that it's a great disadvantage to guitarists and to their marks if they're not assessed by a specialist. (Assessor D)	Legitimacy is contingent on specialised knowledge of educational assessment principles and procedures	"Assessment in music education should be discussed using commonly accepted definitions of assessment, measurement, and evaluation". (Brophy & Fautley, 2017, cited in Brophy, 2019, p. 918)
ER+	The basis for legitimacy is slightly more flexible, e.g. including expert musicians with access to a high level of musical knowledge who may not necessarily possess an intimate knowledge of the musical instrument on which the assessed is performing	"According to Fiske (1975), raters need not be specialists in the area. In this study, for example, one of the percussion evaluators was a brass faculty member." (Bergee, 2003, p. 148)	Legitimacy is somewhat contingent upon specialised knowledge of assessment principles and procedures	I was quite, nervous about meeting you because I'm like, 'wow, what if I'm doing all the wrong things?'. (Assessor C)
ER-	Legitimacy is not contingent upon in-depth musical knowledge, though some knowledge of music is required	Example unapparent in data	Legitimacy is only weakly influenced by specialised knowledge of assessment principles and procedures	The actual terms of assessment are maybe not as even as important as who's assessing… You don't need to have a list of criteria that you need to tick off…You just need to have a vague awareness of… the type of broad level that should be… assessed. (Assessor A)
ER--	Possession of musical and/or skills knowledge has a limited influence on legitimacy	In the real world… When you go out to play people assess you anyway but informally… You're being assessed by the public. (Assessor O)	Legitimacy is not contingent on specialised knowledge of educational assessment principles and procedures	"Knowledge of assessment is passed on as a folk practice, and is essentially unexamined and taken for granted". (Boud, 2010, p. 255)

contexts. Reflecting the mixture of data sampled in this research, this device draws on examples from both interviews with research participants as well as quotations from the scholarly literature to provide indicators for all regions.

7.3.4 Additional notes

More knowledge is not necessarily equivalent to better practice or greater legitimacy in assessment situations. The previous chapter, for instance, illustrated that more nuanced specialist knowledge of a given instrument in assessment conditions could lead to judgements that might be perceived by other assessment participants as over-meticulous and ultimately less valid as a means of representing the holistic worth of a performance. Additionally, Dawson et al. (2013) assert that the accumulation of knowledge about assessment does not mean that assessment practices themselves will be enhanced. At the same time, the constituents of constellations may interact with one another in ways that have implications for systems of legitimation. For instance, quotes provided throughout this thesis indicate a perception of disconnect between music and assessment. Similarly, conceptual ideas from the field of assessment de-valorise some of the constituents of music education constellations (for instance, assessment *of* learning, see Partti, Westerlund, & Lebler, 2015). Maton (2014) refers to this positioning as *charging*, in the sense that different constituents of constellations may be "charged with legitimacy" (p. 150). To quote Maton (2014):

> [M]eanings condensed within practices may be *charged* differently. For example, a social scientific concept condensing a range of political meanings… may be portrayed positively, neutrally or negatively (along a continuum) in comparison to other meanings. Constellations can thus condense more or less epistemological and/or axiological meanings that are charged positively, neutrally and negatively to different degrees. (p. 153, italics original)

The notion of *condensation* is useful here as well, in that it reflects the idea that different amounts of knowledge are possible in a given area. Condensation is related to the concept of semantic density (see Chapter 4),

and its relevance is evident in the observation that that different amounts of knowledge can influence legitimate participation in assessment practices. Charging, condensation, and semantic density are useful theoretical ideas, however a more detailed discussion of these concepts is generally beyond the scope and aims of this part of the thesis—exploring these further is a direction for future research.

7.4 Social Attributes

The previous chapter highlighted some ways in which social relations could be used to characterise students' performance work and criteria for assessing it, however many of the examples in Chapter 6 also highlighted that social relations matter in the sense that they can have implications for the legitimation of people themselves. For instance, while students' effort and enthusiasm (bases in social relations) might function as determining criteria in the formal grading of their work (whether rightly or wrongly), the expression of these attributes of the students can also lead to judgements about the students themselves. The same is also possible in many ways for staff, who might be characterised in many ways (for example, as harsh, easy going, lazy, diligent, traditional, precocious, popular, and so forth).

Where analysis of the interview data revealed clearly discernible epistemological constellations, an axiological constellation generated by social relations was more difficult to meaningfully distinguish. Many possible themes were evident (for example, gender, musical specialism, amount of experience) and while these offered useful indicators of social relations, these themes did not group together in as straightforward a fashion as those described in the previous section. Part of this could be explained as a limitation of the data collection, and future research could focus more specifically on constellations around any of these individual topics. Since this project explicitly sought heterogeneous perspectives, a trade-off is that detailed insights into individual aspects of musical performance assessment were generally less available. As this project is an explorative theoretical work, this does not disrupt the overall aims of the work, however it did necessitate some analytical problem-solving to discern a meaningful approach to conceptualising social attributes. In this respect, being forced to confront this limitation actually enhanced the overall

research outcome by prompting further analytical exploration. To resolve the aforementioned issue, the concept of social relations was revisited, and some finer-grained tools from within this aspect of LCT were adopted to develop the analysis.

7.4.1 Revisiting social relations

The premise of social relations highlights that legitimate participation in practices can be more or less dependent upon being the right *kind* of person. The stronger the emphasis on social relations the more tightly bounded the criteria for legitimate participation. In other words, what it is to be the right kind of participant can be more tightly defined (SR+) or less tightly defined (SR-). In this depiction, the weaker end of the spectrum denotes a space wherein social attributes are relatively unimportant for legitimate participation in practices, while on the stronger end being the right kind of person is critical.

Figure 7.1 Social relations bounding legitimacy

To realise a greater degree of nuance, Maton (2014) distinguishes two derivatives of social relations, which are relations to categories, or *subjective relations*, and relations to interactions, which are *interactional relations*. These concepts were introduced in Chapter 4 but are revisited here to sustain clarity in the discussion. The definitions provided for these concepts in Chapter 4 were

- *subjective relations* (SubR) between practices and the kinds of actors engaged in them
- *interactional relations* (IR) between practices and the ways of acting involved (List adapted from Maton, 2014, p. 184, italics original)

To clarify these concepts further, Maton (2014) highlights that

> [t]hey thus describe how strongly knowledge claims bound and control legitimate *kinds of knowers* (subjective relations) and

legitimate *ways of knowing through interactions with significant others* (interactional relations). Both these social relations can take a multitude of forms. There are numerous potential bases for subjectivity – social class, sex, gender, race, ethnicity, sexuality, religion, region, etc. – and thus categories for defining legitimate knowers. Similarly, there are numerous 'significant others' (which may be objects or subjects) and means of interaction that may serve to define legitimate ways of knowing. These are as diverse as therapy with a psychiatrist, external stimuli and the human mind, master– apprentice relations, parent–child interactions, and so forth. (p. 185, italics original)

A more extended example is discussed in the next section to illustrate the application of these concepts.

7.4.2 Example: The influence of gender in assessment participation

To show one way in which nuance within social relations matters, this example considers a series of comments from a student who was interviewed as a part of the project. Comments from this interviewee—a woman in her final year of jazz performance studies, majoring on the drum kit—illustrate that the implications of social attributes for participating in assessment practices are not limited to the actual performing of music. These comments highlight one important way in which social attributes can influence the outcomes of an assessment event well in advance of the event actually taking place. An illustration of finer-grained thematic analysis of these comments using subjective relations and interactional relations (amongst other coding concepts) is included in Table 7.3.

> I think that definitely being a lady… It's definitely a different world to a dude drummer. If I was a guy and I upload[ed] the same videos, exact same content [to Instagram] I don't think it would be as popular… Unless there was something really special about it. Because I'm a girl it's, "oh whoa it's a girl", that sort of look on things, which is very weird and different, but… I have a

marketability… A lot of the comments are based on, "you're a girl, you're playing drums", rather than, "whoa, that technique's really good", or, "that fill was cool." I feel there's definitely still a gap in the conservatory, in the female drumming world… When you come in [for] classes I'd say it's 80% male, 20% ladies, and usually 20% is singers, so then your instrumentalists are down to two or three per year level, which then I found creates a kind of bro club within the jazz student area… When they need players for bands and when they need players for recitals, they'll go towards their friends, which are the guys. Whether there's a girl at that same level that they can ask, they won't consider, or they won't even think of. We needed to replace a sax player in one of my bands recently, and in the Facebook chat it was all, "listen to all the guy players from the [institution]", even though there's two or three good ladies, they weren't even mentioned… When I [said] "oh how about this girl, how about this girl?", they're like, "oh, I didn't even think of that." And so, I think it's not… forefront-thinking... It's not in the minds of people to pick equally. I think they just go with, their gut and their friends. Assessment-wise I would like to source out more female players for my bands… I would go towards getting more girl players in my band for assessment. One because they're good at their crafts, but two because I'm friends [with them] and generally we get along better… That's kind of bad but I think it's good having [a female staff member] as an [assessor], for that different opinion. I think if it was two guys, say [male staff member] and another guy, I think it would be different results-wise. (Student B)

A number of research studies have recognised the influence of gender in musical performance assessment (see for example, Elliot, 1995/1996; McPherson & Schubert 2004). Generally, the focus of research in this area is on how gender might influence the ways in which students are assessed, in the sense of the grading of their actual performance on a given task. To this end, concepts such as gender have previously been considered to be non-

musical factors (McPherson & Schubert, 2004) in the sense that they sit outside of the category musical performance skills themselves and "produce unfair biases" (p. 74). While this view may not present acute issues in some fields of musical performance, the lines are blurred in others where social attributes such as gender may play a significant role as *part of* the musical performance product itself (in popular music, for example). In this example, however, we see that the effects of social attributes can influence participation in assessment beyond the substantive assessment event. Enacting concepts from the 4-K Model offers a means of distinguishing and theorising an underlying conceptual interplay. Placing these concepts on continua of strength offers a means of distinguishing a topological space for legitimation according to social attributes (Tables 7.4 and 7.5). The final section in this chapter examines this topological space more closely to explore the conceptual gazes that occur within it.

Table 7.3 Coding subjective relations and interactional relations in a line-by-line analysis

Text				
I think that definitely being a lady… It's definitely a different world to a dude drummer.		SubR+ (the categories of "dude" and "girl" drummers are seen to matter)		
If I was a guy and I upload[ed] the same videos, exact same content [to Instagram] I don't think it would be as popular… Unless there was something really special about it. Because I'm a girl it's, "oh whoa it's a girl", that sort of look on things, which is very weird and different, but… I have a marketability… A lot of the comments are based on,				
"you're a girl, you're playing drums", rather than, "whoa, that technique's really good",	Focus is ER, basis in SR			
or, "that fill was cool." I feel there's definitely still a gap in the				
conservatory, in the female drumming world… When you come in [for] classes I'd say		Manifestation of SubR (the female drumming world) drummer		
t's 80% male, 20% ladies, and usually 20% is singers, so then	Manifestation of IR ("bro club" produced through high % of male students)			
our instrumentalists are down to two or three per year level, which then I found creates				
a kind of bro club within the jazz student area… When they need players for bands			IR+ (interactions with others—friends, or those in the "club"—influence participation)	
and when they need players for recitals, they'll go towards their friends, which are				
the guys. Whether there's a girl at that same level that they can ask, they won't consider,				
or they won't even think of. We needed to replace a sax player in one	SubR+, ER- (in this explanation, gender matters, while technical skills are downplayed by the male students)			SR+ (overall emphasis on social relations)
of my bands recently, and in the Facebook chat it was all, "listen to all the guy				
players from the [institution]", even though there's two or three good				
ladies, they weren't even mentioned… When I [said]				
"oh how about this girl, how about this girl?", they're like, "oh, I didn't				
even think of that." And so, I think it's not… forefront-thinking...				
It's not in the minds of people to pick equally. I think	SR+ (overall emphasis on social relations)		IR+ (interactions with friends and gut feeling matter)	
they just go with, their gut and their friends.				
Assessment-wise I would like to source out more female players for my bands… I would	SubR+ (gender matters)			
go towards getting more girl players in my band for assessment. One because they're	ER+, SR+ (SubR+, IR+) (Emphasis on skills gives ER+, emphasis on "girl players" gives SubR+, emphasis on "getting along" gives IR+)			
good at their crafts, but two because I'm friends [with them] and generally we get				
along better… That's kind of bad but I think it's good having [a female staff member] as		SubR+ (gender of assessor matters)		
an [assessor], for that different opinion. I think if it was two				
guys, say [male staff member] and another guy, I think it would be different results-wise.				

Note. Colours reflect relations identified as most emphasised (others may be present), text provides more detailed description of analysis

Key. ▢ social relations (SR), ▢ subjective relations (SubR), ▢ interactional relations (IR)

Table 7.4 A translation device for subjective relations and interactional relations

Concept	Description	Possible manifestations	Strength of relations	Indicated by	Example (Gender)
subjective relations	Emphasis on social categories	Gender Instrument specialism Students Assessors Genre specialism Age Ethnicity	Stronger subjective relations (SubR+)	Membership of particular social categories meaningfully influences assessment participation	Participation in assessment practices is strongly influenced by identification with a particular gender
			Weaker subjective relations (SubR-)	Social categorisation is less influential upon assessment participation	Participation in assessment practices is weakly influenced by identification with a particular gender
interactional relations	Emphasis on interactions with significant others	Personal educational experiences Interactions with peers Exposure to music Interactions with materials	Stronger interactional relations (IR+)	Interactions with significant others meaningfully influences assessment participation	The legitimation of certain gender types is more dependent on actors' interactions with one another and/or materials
			Weaker interactional relations (IR-)	Interactions with significant others is less influential assessment participation	The legitimation of certain gender types is less dependent on actors' interactions with one another and/or materials

7.5 Gazing at Assessment

This chapter has so far been focused on developing some conceptual tools for making visible bases of legitimation underlying assessment participants' capacities to successfully participate in assessment practices. This work offers an opportunity to distinguish between some general ways of viewing those involved in assessment practices that highlight different bases from which to perceive assessment practices and different claims about what attributes are important. The previous section revisited the concept of social relations and discussed the more nuanced applications of subjective relations and interactional relations. The focus here is on the modalities—or *gazes*— produced through the intersection of these relations, visually depicted by the social plane (Figure 7.2). This view highlights that stronger or weaker emphases on social relations can manifest in multiple ways. Taken together, these relations on continua of strength generate a space of possible positions within the general modalities that offers a more nuanced view of the kinds of social relations at play in a given context. Gazes were introduced in Chapter 4 (see Section 4.3.4), however a summary of their main characteristics are reproduced here:

- Where legitimacy is based on knowers possessing a *social gaze*, practices relatively strongly bound and control the kinds of knowers who can claim legitimacy but relatively weakly limit their ways of knowing (SubR+, IR–). For example, standpoint theories base legitimacy on membership of a specific social category (social class, gender, ethnicity, etc.), regardless of knowers' past or present interactions.

- Practices that base legitimacy on the possession of a *cultivated gaze* weakly bound and control legitimate categories of knower but strongly bound and control legitimate interactions with significant others (SubR–, IR+). These often involve acquiring a 'feel' for practices through, for example: extended participation in 'communities of practice' (Lave & Wenger, 1991); sustained exposure to exemplary models, such as great works of art; and prolonged apprenticeship under an acknowledged master.

- Practices that define legitimacy in terms of possessing a *born gaze* relatively strongly bound and control *both* legitimate kinds of knowers and legitimate ways of knowing (SubR+, IR+), such as religious beliefs of an act of God towards a chosen person or people, and claims to legitimacy based on both membership of a social category and experiences with significant others (e.g. standpoint theory that additionally requires mentoring by already-liberated knowers in consciousness-raising groups).
- Practices that relatively weakly bound and control both legitimate kinds of knowers and legitimate ways of knowing (SubR–, IR–) are characterized by weaker social relations that, alongside different strengths of epistemic relations, may form part of either a knowledge code (ER+, SR–) underpinned by a *trained gaze* that emphasizes the possession of specialist knowledge and skills, or a relativist code (ER–, SR–) that offers a *blank gaze.* (Maton, 2014, pp. 185–186)
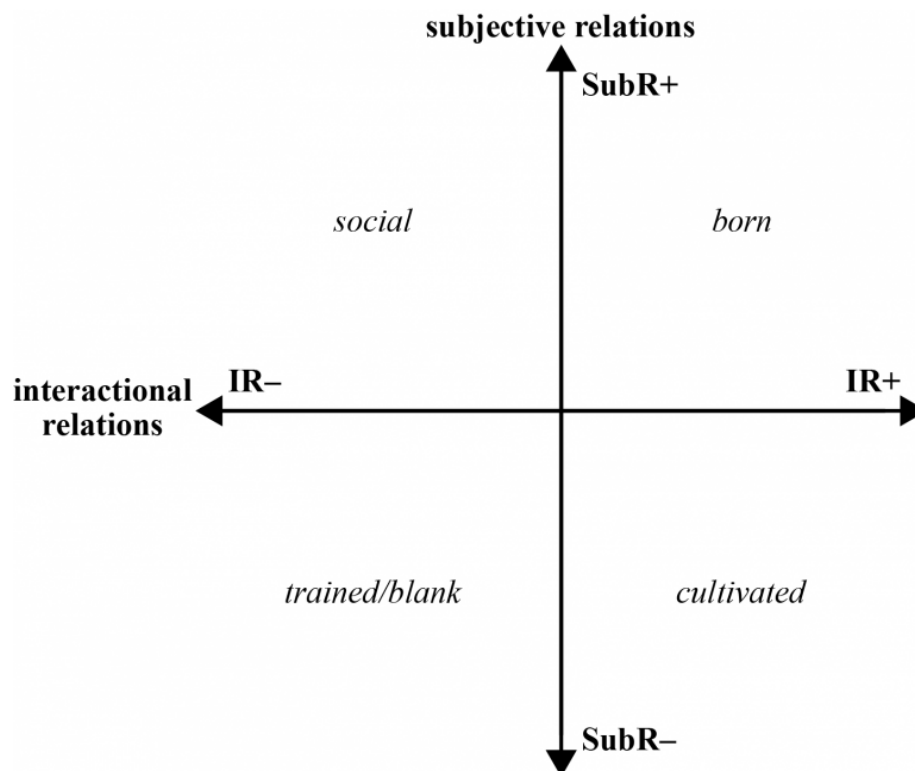

Figure 7.2 The social plane (Maton, 2014, p. 186), reprinted with permission

These gazes reflect ways of seeing and evaluating the world, and therefore constitute possible positions both for enacting legitimation as well as for being

legitimated (where possession of a particular gaze is construed as a person's attribute). In other words, gazes provide a framework for distinguishing ways in which emphases (or de-emphases) on social relations can serve as bases for legitimation. The discussion developed in this chapter so far highlights that different kinds of gaze may be viewed as more or less legitimate, depending on a given assessment context. This section discusses each gaze in turn to consider how participants might come to acquire a particular gaze, and how gaze might be explicitly addressed to generate enhance the capacities of assessment practitioners to participate in assessment. As with the other topological modalities discussed so far in this thesis, gazes are not construed here as mutually exclusive categories. Rather, assessment practitioners may adopt different gazes, or formulate new ones depending on their particular contexts.

7.5.1 The trained and cultivated gazes

Both the trained[23] and cultivated gazes reflect a weaker emphasis on social categorisation. Reflecting an emphasis on epistemic relations rather than social relations, trained gazes describe perspectives underpinned by the possession of knowledge and skills acquired through explicit training. Examples of training processes include staff training days, the explanation of procedures to staff or students, and the provision of clear feedback about the quality of one's musical performance. In these situations, pedagogic transfer is relatively explicit and related to concepts that are well-theorised. Cultivated gazes, by comparison, reflect a stronger emphasis on social relations, and in particular emphasise interactions with significant others (interactional relations; for example, with friends, colleagues, assessment events, musical performances, etc.) as the basis of these perspectives. For those involved in musical performance assessment, the distinction between the two is important for the reason that they reflect different modes of *knowing* assessment—one is more closely aligned with formal training, while the other is more closely aligned with the development of personal perspectives through meaningful

---

[23] The trained gaze is given by an emphasis on epistemic relations rather than social relations, however in the absence of an emphasis on epistemic relations a *blank gaze* is also possible, which reflects a relativist code (see Chapter 4, Sections 4.3.1 and 4.3.4).

interactions over time. In the cultivated gaze, knowledge is less explicitly an object of study. Further comments from one research participant previously quoted in this chapter provided a useful example that highlights this distinction. In the following quote, the interviewee is describing the importance of having instrumental specialists assess students:

> [T]he problem with… instruments that don't have multiple teachers at an institution is that they're the ones that end up… if the teacher isn't marking… getting marked by non-specialists. So if you've got four [specialist teachers of one kind] and they mark each other's students, I think that's much better than if you've got [one kind of specialist], and then [another kind of specialist marks their students]. Unless they know a lot about the instrument I always feel that it's a disadvantage to the [students]. Here [a different kind of specialist] has been marking [students that play my instrument] for the last 25 years, so even though he [doesn't play our instrument] he knows a lot about [it], and I've found that all the assessments the we did last year, we completely agreed on everything. We actually also agreed on the [marks for his students], interestingly enough… I think [it's important] for there to be communication between the examiners, particularly in a situation where… you can ask, 'is this piece difficult?' Because if you don't know, you can't possibly make a judgement… I just was hosting the performance extension workshop and one [student who plays a different instrument] got up and played a piece that we play a lot on [my instrument]. It's a piece that I would give someone in second year and expect that they would be able to play it… She's in fourth year and she's one of the best [at her instrument] and… she's really struggling to play it. I would never have known that had she not told me… that can be a challenge if you're assessing other instruments, and I think you do need to have that awareness of how difficult something is… I think that's a

big [and] really important factor. (Assessor D, a classical performance teacher)

The distinction between training and cultivation is given here in the description of the non-specialist teacher who has extended experience in assessing a particular instrument. The kinds of inter-examiner communication described by the interviewee usefully illustrate cultivation processes as distinct from process that might be considered to be more formal kinds of training (for example, participation in a workshop about the particularities of assessing different instruments). Examples of explicit training of this kind were difficult to locate in the interview data, however many Australian institutions do offer assessment training and it may be that the participants simply had not opted (or been required) to take available training. Conversely, comments from a number of participants in this research did seem to more clearly indicate an absence of formal training in educational assessment practices, and some examples of this have already been provided in this chapter (see for example Section 7.3.2). Although many possessed access to a trained musical gaze (provided through their experiences of formal music education), the data suggested that understandings of assessment developed through cultivation processes over time. For example, Assessor G, a Head of School, explained that his "convictions [about global assessment] have become more established" through his experiences. He explained:

I've tried analytical marking keys… You know, I think that they can work if they're very genre-specific or very instrument-specific… but I think that one of the problems is trying, is trying to get one that's universally applicable. It's a disaster.

Assessor J, a classical department leader, explained that his thinking about assessment had changed over time, emphasising that "it really has". Where assessor J "used to be a real stickler for the technical… because of [his] upbringing" he found that "over time in performing as much as adjudicating and assessing… the performances that are perfect are the performances that are sometimes less interesting". Assessor J explained that he now feels "imperfections are… the things that make us who we are" and that this change

in mindset influences his own playing: "From a personal point of view… that idea has taken the pressure off my own performances". Assessor J further explained that his approach in assessment has "moved from that [technical] point of view", however he emphasised that "we still need the descriptors, we still need the technical guidelines, we still need those things so that we have a benchmark… I think it's the good stuff after that".

Assessor B, a classical department leader at another institution, felt that his thinking about assessment had changed "100%". He explained:

> As a young academic… I had criteria sheets that I made for all our staff… ten [marks] here… rhythm out of ten and, accuracy out of ten, sound quality out of ten—subjective sound quality—blah blah until it adds up to 100 and I thought that was great, really smart. Then [an assessment publication] came around and I realised you can, criteria the hell out of everything, you can over assess to the point of, stifling the creative process… That whole solo recital… That's your opportunity to express yourself to us and show us who you really are, and that's why I don't believe that it should be assessed… I've moved away from it, and it's funny, as I've become more actively involved in the executive element of the university it's made me go further away from seeing assessment as a requirement for performance, because you see the minutiae of it and how you can get so caught up in it.

The idea that assessors cultivate their perspectives on assessment over time is supported by perspectives from the literature on assessment. Boud (2010) for instance, characterises assessment as a "folk practice" (p. 255), where knowledge is "essentially unexamined and taken for granted" (p. 255). Bloxham and Boyd (2007) assert that "[w]e learn the craft of assessment informally through being assessed ourselves and through being part of a community of practice" (p. 3). In the music assessment context, Bergee (2003) agrees (after Fiske, 1975) that "raters need not be specialists in the area" (p. 148). Bergee (2003) posited that more experienced assessors could help

newer faculty members understand assessment processes (a clear example of interactional relations). To quote Gynnild (2016), who studied the assessment of vocal performances using analytical assessment:

> Teachers may use the setting of criteria and the marking process as opportunities to develop their knowledge of assessment criteria and their use. The aim is to gain familiarity, understanding and competence with regard to assessment as a social process rather than to simply operating in a documented and regimented system. (p. 226)

The distinction between trained and cultivated gaze has implications for students as well. Barratt and Moore (2005) usefully highlight the deleterious potential of unchecked cultivation in students' formation of perspectives on assessment. Students' adoption of assessment strategies based on local behavioural norms can produce a range of issues, including placing subtle limitations on their capacity to achieve (see Chapter 3, Section 3.3.1). Participants in this study also described issues with students' cultivated approaches to assessment. One member of a jazz department explained:

> What I'm seeing is… students… realising they're having to jump through hoops and leaving it until the last minute… [It] almost defeats the purpose of criteria, because once you know there are criteria… It puts it in a different space in your brain. (Assessor A)

Student participants also described the influence of cultivation on their perspectives on assessment. For example, one classical percussion student explained:

> I think I'm more aware now… maybe when I first… started my degree, I wasn't as aware of, 'this will relate to this in the bigger picture, or this will relate to this in the bigger picture'. It was, 'oh, I have an assessment, I have to do it, I have to get over it and reach that hurdle'. But now I think I have a much more open-ended perception of assessment, in terms of [how] it will relate to something. But it's still marked… I'm still trying to achieve a

mark through what's presented, but I know that it's relating to this bigger picture that I want. Maybe I was a bit blind to that bigger picture when I started.

7.5.2 The social and born gazes

Social and born gazes both reflect a stronger emphasis on membership of particular social categories, and in turn, a stronger overall emphasis on social relations. Examples provided earlier in this chapter highlighted some of these categories, including gender, and instrumental specialism. A key distinction between social and born gazes is that social gazes place less emphasis on interactions with significant others. A social gaze describes a focus more strongly governed by categorical membership, for example, membership of a particular social group or clique, holding a particular role, representing a particular discipline, or identifying as a certain gender (as illustrated in the earlier example). Born gazes, by comparison, also emphasise interactions with significant others. For example, the legitimacy of a famous musician may be construed through the combination of both cultural heritage and their engagement with a particular musical scene.

Contrasting the social gaze with the trained and cultivated gazes highlights subtle points of nuance in the ways that assessment participants view one another. The example provided in Section 7.4.2 highlights that they can co-occur and relate to one another in meaningful ways. Firstly, the "bro club" described by Student B can be analysed in terms stronger interactional relations (IR+, presumably between male students), producing a cultivated gaze. At the same time, a social gaze is present that meaningfully distinguishes students according to gender. These gazes reflect both the given students' ways of viewing the world, as well as bases for their legitimation.

A second useful example is given by continuing the discussion of specialist assessors (see Section 7.3.1), where thinking analytically with the social gaze adds an additional layer of identity into the mix. While expertise in instrumental specialisms can be interpreted in terms of training and/or cultivation, data collected in this research appears to indicate that instrumental specialisms can also be positioned as identities or social categories. To give an example in the form of a question: To what degree is the conceptual

classification of being a guitarist (or violinist, pianist, vocalist, etc.) more or less influential in a given setting than the particular knowledges or values held about the guitar? Research on instrumental stereotyping suggests that this kind of question is worthy of consideration (Butkovic & Modrusan, 2019). Similarly, the studies of musical performance assessment criteria cited in previous chapters (for example, McPherson & Schubert, 2004; Wesolowski, Wind, & Engelhard, 2015) highlight a direct concern with the co-emphasis on social categorisation (assessment of students' work on the basis of a social or born gazes) and disciplinary knowledge, skills, and taste (assessment of students' work on the basis of trained or cultivated gazes).

7.5.3 Implications of gaze

The concept of gaze offers several implications for musical performance assessment practices. Firstly, it highlights the complexity of the range of viewpoints that are available and offers some means of discerning between these to enable meaningful comparison. Comparing assessment participants' attributes in terms of gaze provides one possible means of circumventing the kinds of crosstalk that can make social problem solving difficult. Recognising that different gazes have different bases clarifies one way in which different people can view things in different ways. Distinguishing between a trained gaze (a basis in epistemic relations, knowledge and skills) and a cultivated gaze (stronger interactional relations), for instance, helps to explain why assessment design experts and musical performance teachers may not see eye-to-eye when it comes to the design of assessment tasks for a given course of study.

Secondly, and related to the previous point, gaze can also be used to theorise issues of access in educational settings. The example discussed in Section 7.4.2 offers one indication of how this might occur in practice. Different gaze-building processes may also privilege different *semantic structures*, in the sense that meanings seem likely to be constructed and legitimated in different ways. In this, gazes would seem to co-develop in knowledge- or knower-building situations. In turn, possession of a particular gaze might influence access to other knowledge or cultivation of social attributes due to their positive or negative charging (see Section 7.3.4). Consider, for example,

the perspective of one jazz performance teacher who took a particular stance on what he perceived to be a general trend towards "standardisation" and "codification":

> Everything's trying to be standardised and… you know, codified, when… a lot of what we get from…our music education is through sort of mentorship, and if you think about that… just imagine a situation where the assessment's not codified, all the effort goes into getting the right people to be teaching and the right people to be assessing, and then you let them go… We get so caught up in this criteria and codification… We seem to be less interested in getting the right people, and even hiring of people is to do with their qualifications and it's a codification— rather than maybe getting the right person for the job, in the sense that maybe you can't write down what they are offering you on paper. (Assessor A)

In this example the concepts of standardisation, codification, and qualifications are negatively charged, while notions of "the right people" and mentorship are positively charged. Comments from this participant throughout the interview indicated that his perspective on music education and assessment was likely informed by a cultivated gaze. There were, for example, many mentions of assessor A's personal experiences as a student and personal experiences as an assessor, as well as a re-emphasising of concepts like mentorship. Assessor A was also a casual employee at his institution, and therefore may have had less exposure to training activities that might have influenced his thinking about assessment differently. One of the implications associated with this broad analysis is that in negatively charging concepts such as standardisation, codification, and qualifications, Assessor A may have more limited access to development of this conceptual territory. In turn, this access has implications in situations such as professional dialogues about assessment, or assessment events involving other assessors who may think differently about assessment. Conversely, others in possession of a trained gaze (for example, musical performance assessment scholars) may hold

perspectives that negatively charge concepts such as the master-apprentice approach to learning and teaching, construing them as traditional and out of sync with contemporary educational philosophies and scholarly thinking (see for example, McWilliam, Carey, Draper, & Lebler, 2006). The point of this partially hypothetical and extrapolated example is that gazes provide a means of describing differences in perspectives which move beyond the specific empirical realities of a particular setting—they enable us to see not only that, but also to analyse the bases for these perspectives (cultivation, training, social positioning, born qualities).

Theorising the bases of gazes also enables critical engagement with the processes by which gazes are developed, which in turn provides a means for meaningful engagement with practices. Placing an explicit focus on developing assessment participants' trained gazes, for example, generates some interesting possibilities. For instance, to what extent should students receive explicit instruction in the development of their evaluative capabilities, or on the logic underpinning assessment designs themselves? This idea has some basis in the literature. In a seminal paper, Sadler (1989) argues that students should be provided with "authentic evaluative experience[s]" (p. 135) in order to develop their capabilities in self-appraising the quality of their work. The more recent concept of *evaluative judgement* directly develops this idea (Tai, Ajjawi, Boud, Dawson, & Panadero; 2017), and Tai et al. (2017) consider some strategies for developing students' evaluative capabilities (see also, Boud, Ajjawi, Dawson, & Tai, 2018). The question posed highlights collaborative assessment literacy as one possible strategy to be explored.

A related question is: To what extent should educators receive explicit instruction in enacting such a dialogue with students? This question is particularly relevant given the indications from some interview participants that students and their teachers are actively involved in dialogue about assessment. For example:

> We talk routinely with my horn school about why we do it, why
> we do what we're doing: yes you're being assessed, yes you're
> being forced into this area… but that is—in my opinion—to get

them at a competitive level of an orchestral job. (Assessor B, a classical department leader)

When it comes to the actual assessment things it's just… It's a can of worms, you know?… We talk a lot about… If a guitarist comes in and plays something… You've asked them to do a certain amount of things on a song and they might not do those things but they play with beautiful time, tone, touch, all these hallmarks of music. Do you give them a good mark because of their natural ability or… a bad mark because they didn't do the thing you asked them to do? So that, that's what we've been dealing with lately a bunch. (Assessor E, a jazz performance teacher)

With my own students it's like, 'I'm going to prepare you for this assessment', and we'll talk about all the things that matter to you [the student] and I as performers… These are the kinds of things that I'll be looking for and the other panellists are looking for, so they already know upfront what's expected of them… I would teach to the rubric actually to go look these are the elements that we're going to assess you on. (Assessor O, a jazz performance teacher)

Explicitly foregrounding these types of questions offers a general path for disrupting processes of legitimation that produce undesirable effects. For example, it provides a means by which to address the lack of confidence felt by some music educators about their own abilities as skilled assessors (Section 7.3.2). It might also provide a means by which to address social situations amongst students (for example, the establishment of a "bro club" described previously). To connect with Forbes (2016) and Wrigley (2005) who highlight issues of ideological disconnect in higher music education, gazes provide a language for explicitly engaging with ideological positions. The epistemological constellations (musical constellations and educational assessment constellations) proposed in Section 7.3 may not coalesce for music educators precisely because they are disconnected on an axiological

level. This idea will be explored in further detail in the next chapter which uses the Autonomy dimension of LCT to consider the legitimation of assessment practice itself in higher music education contexts. Ultimately, the usefulness of gaze in practical settings is that it provides a means of making explicit the ways in which assessment participants see one another, and so offers a means for clarifying and interrogating (or changing) the rules of the game in assessment contexts.

Despite the prevalence of resources available for effecting development and change in social settings, actively engaging with participants' gaze in assessment contexts is likely to be a complex process for the reason that it places explicit emphasis on changing both knowledge and values. Part of the problem here is the broad range of possible combinations of knowledge and value structures embodied by people, even within tightly defined contexts (for example, a single instrumental disciplinary area within a conservatoire). Added to this is the entrenched nature of educational practices in higher music education (Parkes, 2010; Partti et al., 2015). While this chapter has proposed that assessment participants' stances and ways of thinking can be productively conceptualised in terms of gaze, developing or effecting change in this space is likely to benefit from a meaningful approach for interpreting the structuring of the social perspectives in question. The following chapter discusses some additional theoretical perspectives that may be useful in this regard. It uses the Autonomy dimension of LCT to develop some ways of interpreting the legitimation of assessment practices (and approaches within these). As part of this Chapter 8 offers some means of theorising the positions and perspectives adopted by assessment participants, in turn providing a means by which these can be made conceptually visible and perhaps more responsive to engagement in professional settings.

7.6 Conclusion

This chapter was the second of three chapters to discuss conceptual approaches to musical performance assessment. It was a response to the second guiding research question: How can bases of legitimation for assessment participants be meaningfully conceptualised? In developing a response to this question, this chapter proposed some ways of

operationalising concepts from LCT to foreground bases of legitimation underlying assessment participants' attributes. This chapter extended the use of the Specialization dimension of LCT from the previous chapter to include some additional theoretical resources. To this end, Section 7.2 introduced the concept of constellations to demarcate conceptual territories, including epistemological constellations of knowledges and skills (underpinned by epistemic relations), and axiological constellations of values, identities, and positions (underpinned by social relations). Constellations are one device for conceptualising the attributes of assessment participants, and were useful in distinguishing epistemic musical constellations from epistemic educational assessment constellations. Section 7.3 discussed the characteristics of assessment participants' epistemic attributes (underpinned by epistemic relations), situating these within the two epistemic constellations, and provided a translation device for realising strengths of epistemic constellations in musical performance assessment contexts. This device can be used to conceptualise ways in which assessment participants are legitimated according to their possession of epistemic attributes. Section 7.4 focused on assessment participants' social attributes (underpinned by social relations), and discussed an approach to conceptualising these attributes using the finer-grained concepts of subjective relations and interactional relations from within the Specialization dimension of LCT. These concepts enabled a more nuanced view of social attributes, highlighting contrasting emphases on social positioning (subjective relations) and interactions with significant others (interactional relations). The final part of this chapter considered the topological space produced through the interaction of these concepts to discuss different types of gaze available to assessment participants in making sense of assessment practices.

Together, Chapters 6 and 7 have broadly focused on legitimation of the product of assessment (students' work), and the people involved. These two chapters have set out several conceptual positions on musical performance assessment that recognise assessment practices as underpinned by contrasting bases of legitimation (epistemic relations and social relations). The following chapter shifts the focus to a third object of legitimation, which is assessment practices themselves. Chapter 8 subsumes the ideas developed

in Chapters 6 and 7, and takes as its conceptual point of departure the Autonomy dimension of LCT which offers some different conceptual tools to those that have been enacted so far.

Chapter 8: That's Not Music!

The Legitimacy of Assessing Musical Performances

8.1 Introduction

This chapter is the third in the thesis to discuss the outcomes of the research. It is a response to the third research question: How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised? This chapter builds upon the previous two chapters, which together describe a range of approaches for conceptualising *languages of legitimation* (Chapter 2) underlying musical performance assessment practices: Chapter 6 explored an approach to conceptualising criteria for musical performance assessment, while Chapter 7 explored bases of legitimation for assessment participants. These chapters set the scene for this final discussion chapter by describing some basic conceptual positions for social structures underlying assessment practices. Building on these chapters, which principally drew on concepts from the Specialization dimension of LCT, this chapter turns to the Autonomy dimension to develop a different set of analytical tools. In focusing on the legitimation of practices themselves, this chapter locates earlier discussion in the broader context of musical performance assessment practice in higher education.

The first part of this chapter (Section 8.2) describes the broader theoretical framing of the chapter and introduces the Autonomy dimension of LCT. The second part of the chapter (Section 8.3) explores approaches for realising the conceptual tools offered by Autonomy in the context of musical performance assessment. Section 8.4 discusses a conceptual space of possibilities characterised by Autonomy codes. These codes propose a range of possible bases for the legitimation of musical performance assessment practices.

8.2 Theoretical Points of Departure

8.2.1 Revisiting the context

The broad rationale for this part of the research is that the legitimation of assessment practices is a very real problem. It is the elephant in the room in the scholarship on assessment: What if those involved simply don't believe in assessment? Why might this situation arise, and what are the implications? This issue is visible in some of the comments provided by research participants. For example, one classical department leader commented: "I don't believe that it [musical performance] should be assessed because that's… the human side of what we do, which is the [emphasised] most important reason we engage in music" (Assessor B). A jazz performance teacher at the same institution provided a similar view, highlighting a distinction between music and science:

> Just as well it's [the interview] anonymous. No it's funny because I talk to my students about this and sort of say to them how the whole world is kind of obsessed with assessment… What does it mean in a musical sense, you know? Most of it's bullshit really, but on the other hand, there are some students who respond to that—it makes them work harder. But if you think about music as a… sort of an opposite human pursuit to science… It's about looking into yourself and finding out about yourself and asking questions about you and, who you are, and what your relationship to the world is. That's what music is. How do you assess all that?… It rhymes with stress almost doesn't it. Assessment, stress. (Assessor A)

The issue with which this chapter is concerned relates broadly to the situation that assessment practitioners—be they staff, students, or others—have feelings and thoughts about assessment practices themselves that lead to value judgments about the legitimacy of those practices and influence actions associated with them. In the music education context, a central issue is the lack of agreement about the appropriate way to assess musical work. These

issues were discussed earlier in the thesis (Chapters 2 & 3). Enmeshed with the broader discussion around the assessment of complex and creative work (see Chapter 3) is the question of disciplinary authenticity. This is a subtle distinction. The debate about strategies for assessing complex work has philosophical, meta-theoretical underpinnings—it has an epistemological flavour, in the sense of grappling with the issue of legitimate representation of knowledge. The disciplinary context adds another dimension, wherein debates about practice quickly proceed to debates about disciplinary authenticity: Thus the distinction between assessment practices as *epistemologically authentic* and assessment practices as *authentic to music*. To quote a classical performance lecturer at one institution on the subject of segmenting musical performances for assessment:

> You know, it's like… It's not playing music anymore. It's like if in science the whole [aim] was to dissect a toad, but instead of a toad you do it virtually on a computer, that's kind of what it is. It's so sterile. It's not music… it's just like a different thing. (Assessor L)

Of course, subscription to a particular strategy for grading students' work is only part of an assessment practice, but it is an example that highlights the kinds of tensions underlying musical performance assessment practices. The issue in question is the disjunct between *musical* ways of doing things, and *non-musical* ways of doing things. The kinds of distinctions drawn in research about musical performance assessment criteria are a different example of the same underlying issue (see Chapters 3 & 6). Musical performance programs in Australian universities are a ripe context for these tensions to emerge, reflecting the convergence of two distinct institutional paradigms underpinned by different sets of values (Chapter 2). This chapter is about making visible some parts of this struggle, and in so doing, offering an approach by which to meaningfully respond to conceptual loggerhead situations in musical performance assessment practices.

8.2.2 The Autonomy dimension of LCT

The previous two chapters focused most strongly on theoretical ideas from the Specialization dimension of LCT to conceptualise aspects of musical performance assessment practices. Chapter 6 used Specialization to conceptualise musical performance assessment criteria, while Chapter 7 introduced constellations, and extended these concepts to conceptualise the attributes of assessment participants. This chapter builds upon the previous two to propose an approach for conceptualising principles underlying the legitimation of assessment practices themselves. To this end, this chapter turns to the Autonomy dimension of LCT, which offers an additional set of theoretical tools for analysing practices. An overview of Autonomy was provided in Chapter 4, however the basic principles are revisited here to preserve the continuity of the discussion.

As with the other dimensions of LCT, Autonomy offers a generic set of concepts that need to be calibrated to context. The basic premise underlying Autonomy is that "any set of practices comprises constituents that are related together in particular ways" (Maton & Howard, 2018, p. 6). Maton and Howard (2018) explain:

> The constituents and the basis of how they are related together may take many forms. Constituents may be actors, ideas, artefacts, institutions, machine elements, body movements, sounds, etc.; how such constituents are related together may be based on explicit procedures, tacit conventions, mechanisms, explicitly stated aims, unstated orthodoxies, formal rules, etc. Autonomy codes explore the boundaries that practices establish around their constituents and the boundaries they establish around how those constituents are related together. (p. 6)

These two sets of boundaries are given analytically by the concepts of

- *positional autonomy* (PA) between constituents positioned within a context or category and those positioned in other contexts or categories; and

- *relational autonomy* (RA) between relations among constituents of a context or category and relations among constituents of other contexts or categories. (Maton & Howard, 2018, p. 6, italics original)

Together, these concepts provide a useful analytical language for teasing apart problematic divisions underlying assessment in higher music education by distinguishing degrees of *insulation* of contexts and categories. The significance of this kind of insulation is emphasised in contrasting related contexts and categories. For example, based on the body of literature locating higher music education in the Australian university context, two distinct constellations are offered in Table 8.1 to distinguish traditional conservatoire education and university education. Studies from which the signifiers included in Table 8.1 are synthesised include Allais (2012a, 2012b), Bridgstock (2009), Burwell, Carey, and Bennett (2017), Forbes (2016), Jørgensen (2002, 2010), Lebler, Carey, and Harrison (2015a), Parkes (2010), Perkins (2013), and Wilson (2018), amongst others cited in Chapters 2 and 3.

Table 8.1 Contrasting constellations of traditional conservatoire education and university education

Traditional conservatoire education	University education
• Master-apprentice	• Research
• Performance-focused	• Learning outcomes
• One-to-one lessons	• High enrolments and large classes
• Professionals artists as teachers	• Online learning
• Western-art music	• Generic transferable skills
• Verbal in-lesson feedback	• Regulation and accountability
• Performance assessment	• Large comprehensive institutions (in Australia)
• Specific knowledge and techniques	• Prevalence of and imperative towards postgraduate qualifications amongst teaching staff
• Intensive practice	
• Small student cohorts	
• High level of skill required to enter training	
• Smaller institutions	

Relationships between the constituents of these constellations have implications for practices arising at their intersection (see for example Forbes, 2016; Wilson, 2018). Hence, the degree to which their constituents (including

assessment practices) are insular has implications for the integrity of practices located within and between them. The next part of this chapter discusses the concepts of positional autonomy and relational autonomy in greater detail to develop an analytical framework for interpreting autonomy in musical performance assessment practices.

8.3 Working with Autonomy

Since Autonomy is concerned with the insulation of particular categories, analytical activities using the concept require a starting point, or primary object of focus. To quote Maton and Howard (2018), "[p]ut simply, to analyse insulation, one must first ascertain what is being insulated" (p. 10). What is important is that the object of focus is treated not as an absolute category—it is a "starting point rather than a conclusion" (Maton & Howard, 2018, p. 10). The previous section offered some general points of departure by highlighting two contrasting constellations—*traditional conservatoire education* and *university education*. The following section operationalises concepts from Autonomy to explore how aspects of assessment practices can be conceptualised as more or less musical, and how this can be linked to possible positions that serve as bases for the legitimation of (and within) assessment practices. The degree to which an aspect is considered to be musical is given by claims observable in a range of data collected through the project. Some examples include distinctions and value judgements in the literature related to systems of categories that discern between musical and non-/extra/less musical (for example, Blom & Encarnacao, 2012; McPherson & Schubert, 2004); the kinds of distinctions and value judgements made by people (see for example the quotes included earlier in this chapter); as well as the presence or absence of disciplinary musical focus in governing documents related to higher music education programs.

8.3.1 Translating the concepts

As in the preceding chapters, a more specific translation device offers a means of illustrating the links between theoretical concepts and their empirical referents. The structure of the device illustrated below differs slightly to those offered in previous chapters, however, to accommodate the focus of Autonomy

on the insulation of a particular object of study. The approach followed here is modelled on research conducted by Maton and Howard (2018), who used Autonomy to study Science and History lessons in Australian pre-tertiary education. Maton and Howard (2018) offer a generic translation device for enacting Autonomy that accounts for some important analytical distinctions within the dimension. Maton and Howard (2018) explain:

> The device poses the question of what makes the context or category being studied a context or category for the actors involved. As befits a device for enacting concepts, this is not a philosophical conundrum but rather an empirical question concerning the object of study at hand. It is to ask: what constituents and what principles of relation (e.g. purposes, aims, ways of working) are considered constitutive of *this* context or category, *here, in this space and time, by these actors?*" The result is a 'target' that provides a starting point for determining autonomy codes… [T]arget constituents embody stronger positional autonomy and all other, non-target constituents embody weaker positional autonomy; target principles for relating constituents embody stronger relational autonomy and all other, non-target principles of relation embody weaker relational autonomy. (Maton & Howard, 2018, p. 10, italics original).

Put another way, the focus on the degree of insulation of *something* allows for categorical distinctions between constituents that are central to that object of study, and those that are progressively less important to the definition of the category. Where finer-grained analysis offers more nuanced distinctions between those constituents (or relations between constituents), compounding levels of distinction are available. Maton and Howard (2018) offer some terminology for these levels, explaining that the basic distinction between target and non-target "can be divided by asking which target constituents and principles are considered *core* and which *ancillary* to the context or category, and which non-target constituents and principles are considered closer to

(*associated*) or further from (*unassociated*) the target" (Maton & Howard, 2018, p. 10, italics original).

8.3.2 Positional autonomy

Positional autonomy reflects the idea that "constituents positioned within a context or category and those positioned in other contexts or categories" (Maton & Howard, 2018, p. 6) may be more or less autonomous from one another. To give a strong example; in the context of music education, course content is likely to be insulated from content relating to the field of chemistry most of the time (higher positional autonomy, PA+). Within music education settings, different fields of music could be more strongly autonomous from one another (PA+), or they might blend together (weaker positional autonomy, PA-), however both of these positions would likely express stronger positional autonomy than one which included concepts from chemistry. There are many possibilities for realising the spectrum of positional autonomy empirically some of which are contemplated here.

As discussed earlier in this thesis, one of the main questions at the heart of assessment practices is the question of what to assess (Broadfoot, 2012; Lebler & Harrison, 2017; see Chapters 3 & 6). Chapter 6 explored possibilities for conceptualising criteria themselves, and highlighted that different assessment participants may value particular criteria differently. Building on Chapter 6, positional autonomy offers one means for depicting which criteria are conceptualised as the most important (the target criteria), and which are considered less important, or are de-valorised (non-target criteria). Of course, different groups may have different targets. For example, those with regulatory interests may be most concerned with the attainment of learning outcomes, including generic graduate attributes—this perspective is conveyed by Lebler & Harrison (2017) who tell us that in Australia "the 'what' that should be assessed has been defined by the regulatory authorities, and thus students' achievement of predetermined learning outcomes is mandated in the Australian Qualification Framework" (p. 96, see also Allais, 2012a). Lebler and Harrison (2017) also tell us, however, that, at the more disciplinary level, "[t]he question of what to assess frequently includes considerations of whether craft or artistry should be the focus of assessment" (p. 97). There are thus, multiple

possible targets and different groups aiming for these targets at the same time. A similar set of distinctions was evident in interviewees' responses to questions about the purposes of assessment. Some highlighted that the main purpose was to enable students to graduate:

> It gives them a grade so they can finish their degree. Sorry, bit simplistic… if you asked me, what's the purpose of learning your instrument, well that's a totally different question. (Assessor N)

> Universities are bureaucratic structures, and students do, 24 units, say, for a degree. You've got to pass those units, and the units have a grading criteria. That's how it all works… That's why we do it [assessment], because… the university really stipulates that… the students have got to pass, either by a percentile pass on a particular graded schema, or a pass-fail… If someone is a 63, or if someone is a pass or fail, I mean, ipso facto, you must have some sort of form of assessment, I think. (Assessor F)

Others felt more strongly that the purpose assessment was to help students to develop their skills:

> I think for me… the reason why I set assessments for my drum students is more due to the process. So that that four weeks becomes a little more disciplined because they're working towards a goal… That's why I assess more. (Assessor E)

> It should be basically to really just motivate people [students] to develop in their areas… I think that's the purpose of [studying] music in general… You're not here to get a qualification… I think music… The reason why you're studying it is because you want to develop your abilities and your expertise in this area. (Student D)

These responses reflect a general theme in literature and discourse around musical performance assessment, which is the distinction between constituents of practice considered to be *more musical*, and constituents of

practice considered to be *less musical.* In terms of positional autonomy, these characterisations are often closely related to claims about legitimacy, where more musical constituents appear often to be imbued with greater legitimacy than less musical constituents. Some examples of this perspective have already been provided in previous chapters: for instance, the view of one Classical department leader and institutional executive that numeric scores work "against the concept of performance" (Assessor B). Similarly, the previous chapter included a quote from Assessor A, who felt that:

> We seem to be less interested in getting the right people, and even hiring of people is to do with their qualifications and it's a codification—rather than maybe getting the right person for the job.

Earlier in the interview, Assessor A made a related comment (also partially printed in Chapter 7):

> Everything's trying to be standardised and… codified, when… a lot of what we get from… our music education is through sort of mentorship… just imagine a situation where the assessment's not codified [and] all the effort goes into getting the right people to be teaching and the right people to be assessing… If you have to assess… a bass student or a guitar student… you can tell very quickly if they're playing—in a jazz sense—if they're playing over changes, whether they know what they're doing, whether they have good time, whether they have ideas, whether they know the chords, without actually playing up and down the scales and all this sort of stuff. You know, you can sense can't you. You don't need to have a list of… criteria that you need to tick off, not really

Both quotes imply legitimation of constituents (marks, qualifications, people, and criteria) on the basis of their conceptual relationships to music. In terms of positional autonomy, musical constituents can be construed in terms of greater importance (embodying stronger positional autonomy, PA+), while less musical constituents can be construed as embodying weaker positional

autonomy (PA-). Similar distinctions are evident in the literature on criteria for musical performance assessment where criteria positioned as more musical are less questionable in terms of their legitimacy (see for example, McPherson & Schubert, 2004; Wesolowski, Wind, & Engelhard, 2015). This set of observations offers a useful starting point for conceptualising musical performance assessment practices in terms of positional autonomy, where more musical constituents of practices reflect stronger positional autonomy, while less musical constituents of practices reflect weaker positional autonomy (Table 8.2). It is worth re-emphasising, however, that targets are not universal, and that multiple groups are aiming for different things. Chapter 6, for example, highlighted a value amongst some participants for progress and effort as indicators of musical achievement, while for others (for example, those with regulatory responsibilities) these criteria are generally de-valorised as academic non-achievements (see Chapter 6, Section 6.5.4).

The range of possibilities for realising positional autonomy in terms of *musical* performance assessment constituents is, of course, very broad. This kind of analytical focus could be directed at virtually any group of constituents, including materials such as documents (for example, course profiles, scores, books, program notes), equipment (for example, instruments, recording devices, computers), and even physical spaces (for example, classrooms, performance venues, practice rooms). The main purpose of this limited illustration is to highlight one way in which categorical insulation (here the insulation of a musical category of assessment constituents) can influence legitimation in musical performance assessment contexts. Positional autonomy is useful for the reason that it highlights which constituents are considered to be core (or perhaps, authentic) targets in musical performance assessment, and in so doing provides a device for making clearer the boundaries between familiar and comfortable constituents of practice, and those that are more foreign. Crossing these boundaries can have substantial effects. The case of learning outcomes is one useful example. The heritage of learning outcomes statements (Sadler, in press) suggests that the case for learning outcomes to be construed as musical in quality is weak, if most simply for the reason that learning outcomes are an interdisciplinary device linked to the qualifications framework model of higher education (Allais, 2012a).

Table 8.2 Conceptualising positional autonomy for musical constituents of assessment practices

Continuum of positional autonomy	Strengths of positional autonomy	Description	Indicators
PA+ ↑ ↓ PA-	PA+	Constituents construed as most musical	Musical performances Musical criteria Musicians Holistic appraisals Effort/progress Extramusical criteria Rubrics for technical exams
		Constituents construed as somewhat musical	
	PA-	Constituents construed as less musical	Formalised criteria Formal qualifications Numeric grades Non-musical criteria Non-musicians Generic learning outcomes
		Constituents construed as least musical	

Allais (2012a) explains one way in which being forced to work with unfamiliar, non-disciplinary tools can have deleterious effects:

> Outcome statements can force the expert or peer evaluators out of their role as expert peers, and into the role of bureaucrats, where they find themselves judging courses on the basis of requirements which are in no way internal to the tradition of their knowledge area or practice. The specificity that different experts bring may be marginalised by the genericism that is inherent in learning outcomes. (Allais, 2012a)

Of course, engaging with constituents beyond the musical position may also be of great benefit—there are, for example, many resources that are useful in music education contexts that are not inherently musical themselves. In the context of assessment, however, positional autonomy is one device for conceptualising how gradations of categorical distinction can influence participation in practices.

8.3.3 Relational autonomy

Where positional autonomy is focused on the insulation of categories, relational autonomy is more about the insulation of particular *ways of working*. It reflects the idea that there exists degrees of autonomy "between relations among constituents of a context or category and relations among constituents of other contexts or categories" (Maton & Howard, 2018, p. 6). Here, the notion of relations between constituents of a context could be substantively realised as the relations between the content of music education and kinds of assessment practices or knowledge (this distinction was discussed in Chapter 7). Relational autonomy highlights that the interaction between these two constituents may be more or less autonomous from kinds of interactions—or relations—between constituents in other contexts. For example, the relationship between assessment approaches and musical performance course content might be quite different to the approaches taken in disciplines—or other areas of music education—that make use of multiple choice tests to test students' knowledge (stronger relational autonomy, RA+). Alternatively, the relationship between these constituents might be construed as similar to what occurs in other settings: musical performance competitions, perhaps, or other kinds of assessment involving the appraisal of students' performance of work, such as case-based learning in medicine education. Where the relationships between constituents in one context are less insulated from the relationships occurring in another category, relational autonomy is said to be weaker (RA-).

As with positional autonomy, the categorical focus here is on the musical authenticity of ways of working, manifest in approaches for assessment. Data collected throughout the project reflected strong claims about the legitimacy of particular approaches for musical performance assessment, which can be observed very clearly in debates about analytic and holistic approaches (see Chapter 3 for an introduction). Summarising perspectives on the "quantification of quality" (p. 73) in musical performance assessment, Wrigley (2005) clearly describes a position of strong relational autonomy. In the following passage, the scholars' critical stance towards the conceptual segmentation of musical performance for assessment purposes highlights a clear degree of insulation between a) approaches considered to

be musical, and b) conceptual segmentation as a less musically authentic way of working.

> For some time there has been an ideological resistance musically to the quantification of quality. Within the traditional approach to instrumental music teaching there has been a widely held belief that it is artificial and inappropriate, if not impossible to objectively measure and quantify music performance… Mills (1991) argued that "a performance is much more than a sum of skills and interpretation" (p. 175), and that "...as a segmented assessor, it seems that I must turn the performance into something less coherent than music before I may assess it" (p. 173). Johnson (1997) has contended that the "subjective value-judgements of the examiners" cannot be replaced by the use of criteria (p. 274). He has argued for the reliance on tacit knowledge to continue by contending that "...the examination of performance must always be, as it has always been, dependent upon the experience and good musical sense of the examiners" (p. 275). The results of a recent qualitative study of the perceptions of 15 Australian music conservatorium examiners by Stanley et al. (2002) supported the views of Mills (1991) and Johnson (1997). They found that examiners believed that their use of a limited number of criteria interfered with their desire to holistically assess the performance and they preferred more time to write detailed, personal comments…. [I]n Australian music institutions… there has been an ideological preference to uphold traditional educational models of teaching and learning that support the primacy of the summative purpose of assessment…. (Wrigley, 2005, p. 73).

Some interviewees took similar positions, likening assessment practices—and the approaches of higher education in general—to non-musical disciplines, such as science. Two quotes to this effect were included in Section 8.2.1. Analytically, this is useful for the reason that it paints a clear depiction of how

less musical approaches might otherwise be characterised. Musical authenticity influenced interviewees' perceptions of assessment approaches in other ways as well. For example, several participants commented on the need for recital examinations to reflect real-world performance standards. One classical department leader commented:

> I'm okay with the first year being a jury… But then second year onwards I really think they should play a longer performance… I come from the States and… junior and senior year we all have to give full length recitals of… 70 to 90 minutes of music… [I]f they [students] decide to do the performance stream, and they want to be a performer of some kind, whether they are playing you know, recitals or chamber music or auditions or whatever, they need to have a more proper format of playing such that… they feel the length, they decide the repertoire and then… performing [for] a proper audience. (Assessor K)

Comments provided earlier in this chapter and in other parts of this thesis have illustrated that the formal assessment of musical performance itself was construed by a number of research participants as relatively unmusical. The practical implications associated with the relative insulation of musical approaches for assessing musical performances are similar to those discussed in the last section, however they are perhaps more clearly evident in the amount of debate devoted to them. Table 8.3 provides substantive indicators for various strengths of relational Autonomy. Having developed interpretations of both positional autonomy and relational autonomy, the following section considers the topological space produced through their interaction.

Table 8.3 Conceptualising relational autonomy for musical ways of working in assessment practices

Continuum of relational autonomy	Strengths of relational autonomy	Description	Indicators
RA+ ↑↓ RA-	RA+	Approaches construed as most musical	Real-world performances Mentorship One-to-one lessons Recording Workshops Holistic appraisal Longer recitals
		Approaches construed as somewhat musical	
	RA-	Approaches construed as less musical	Group lessons Grading Shorter recitals Numeric marking No lessons Segmental approaches Scientistic approaches
		Approaches construed as least musical	

8.4 Using Autonomy to Interpret the Legitimation of Assessment Practices

The focus in the previous sections was on realising positional autonomy and relational autonomy in the context of musical performance assessment. The approach taken focused on the insulation of musical constituents of assessment practices (positional autonomy) and musical ways of working (relational autonomy). Together, these sections provide the basic positions from which the Autonomy codes discussed below are generated (Figure 8.1). These codes offer a conceptual framework for interpreting practice, and the following sections describe the conceptual features of these codes to discuss how they relate to musical performance practices. The basic Autonomy codes as defined by Maton and Howard (2018) are that:

- For *sovereign codes* (PA+, RA+) status is accorded to strongly insulated positions and autonomous principles. What is valued emanates from within the context or category and acts according to its specific ways of working: internal constituents for internal purposes.

- For *exotic codes* (PA–, RA–) legitimacy accrues to weakly insulated positions and heteronomous principles. What is valued are constituents associated with other contexts or categories and ways of working from other contexts or categories: external constituents for external purposes.

- For *introjected codes* (PA–, RA+) legitimacy resides with weakly insulated positions and autonomous principles. What is valued are constituents associated with other contexts or categories but oriented towards ways of working emanating from within: external constituents turned to internal purposes.

- For *projected codes* (PA+, RA–) status resides with strongly insulated positions and heteronomous principles. What is valued are constituents from within that are oriented towards ways of working from elsewhere: internal constituents turned to external purposes. (Maton & Howard, 2018, p. 7, italics original)
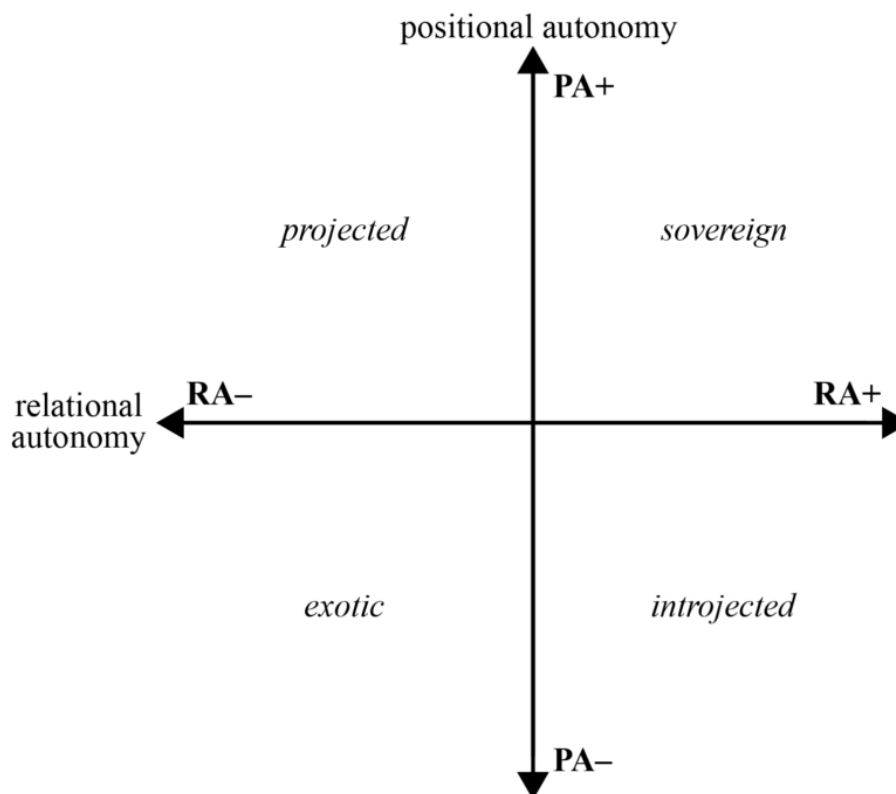

Figure 8.1 The Autonomy plane (Maton & Howard, 2018, p. 6), reprinted with permission

8.4.1 Describing the Autonomy codes

The descriptions provided by Maton and Hoawrd (2018) combined with the translations of positional autonomy and relational autonomy provided above offer some points of departure for a set of Autonomy codes. The codes described here reflect the focus of this chapter on the musicality of assessment—selecting a different object of focus would produce different Autonomy codes. for musical performance assessment.

Firstly, the sovereign code describes situations wherein both the constituents of categories and their associated ways of working are relatively insulated (PA+, RA+). Interpreted in terms of a focus on musical constituents and approaches, this region could be construed broadly as an area comprising musically authentic practice. The emphasis in a *musical sovereign code* is on musical constituents, which includes music itself, as well as closely associated products, people and concepts. For example, professional performers and mentors, instruments, real-world performance venues. At the same time, it reflects an emphasis on musical ways of working, including performance practices, educational approaches, and more holistic assessment strategies. In some ways this seems the simplest code to recognise, for the reason that claims about what is and isn't considered to be authentically musical are often quite clear.

Secondly, the introjected code describes situations where the constituents of the musical performance assessment practice are less musical (PA-), while the associated ways of working remain relatively musical (RA+). Indicators of introjected codes include emphases on factors such as progress and effort, which were valuable to a number of interview participants in the context of musical education, although these constituents are themselves inherently less musical than core constituents. Other examples that embody introjected codes might include the *soft* and/or *non-musical* (which nonetheless relate to musical performance activities) categories described by Blom and Encarnacao (2012), which include factors such as reliability, punctuality, social factors, and leadership. Further regions of the introjected code might also include even less musical constituents again—for example, generic learning outcomes—where the associated ways of working remain relatively musical (for example, in a holistic assessment scenario).

Thirdly, the projected code reflects situations where the emphasis remains on musical constituents of practices (PA+), however their associated ways of working are relatively unmusical (RA-). A pertinent example here is the use of analytic and/or segmented approaches for assessing musical performances (Chapter 3). A projected code can also be theorised as underlying the *imperative* towards acquisition of postgraduate qualifications for institutional staff who are already musical performance industry professionals.

Finally, the exotic code describes situations wherein both the constituents of musical performance practices as well as the approaches and ways of working involved are more weakly insulated (PA-, RA-). At the more extreme edge of this region are aspects of practice that have little to do with products, people, concepts, and processes considered to be musically authentic. At the more general level this might include things like course enrolment processes, and non-musical coursework. In the context of musical performance assessment, some possibilities are given by McPherson and Schubert (2004) who determine non-musical criteria which are further removed from actual musical performance events than those provided by Blom and Encarnacao (2012).

8.4.2 Implications

The concepts described in this chapter are an attempt to provide some approaches for meaningfully visualising processes underlying the legitimation of musical performance assessment practices. Autonomy codes provide one means for interpreting why some assessment participants may struggle to embrace or effectively adopt particular assessment strategies. For example, the use of assessment strategies construed as less musical (for many participants and scholars, analytic assessment) combined with a focus on more authentically musical constituents (for example, musical performance recitals, assessors who are musical performers) produces a projected code wherein a less musical approach is being projected onto a musical set of constituents. For some, this kind of projected code may be construed as a foreign imposition, or a mandate to conform to other ways of working that are associated with stereotypically non-musical fields such as science (a clash with a sovereign code). For others, the projected code may signal a

strengthening of practice through the use of foreign approaches that offer something more inherently musical approaches do not (for example, measurement concepts). Similarly, a focus on less musical constituents (for example, generic learning outcomes statements) amidst more musically authentic assessment designs may be construed on the one hand as a subversion of an otherwise authentic process, or on the other as the effective purposing of musical approaches to the ends of higher-order development. Thinking in these terms is useful for unpacking sentiments such as those expressed by Assessor A at the beginning of this chapter, who described assessment as "bullshit". Reviewing the interview as a whole, it is clear that Assessor A feels that there is an imperative at play which valorises both non-musical constituents (for example, qualifications, specific criteria, inauthentic performance lengths), as well as non-musical ways of working (for example, standardisation and codification). Juxtaposed with Assessor A's clear value for musical authenticity (a sovereign code) is an exotic code, where the emphasis is on neither musically authentic constituents or musically authentic ways of working. Although others might see value in this exotic code, for Assessor A the exotic code is the nub of the problem. The possible differences in perspectives matter, for the reason that they explain why different groups may not see eye to eye on musical performance assessment matters: they are not operating within the same code.

One of the useful applications of the set of concepts developed here is to provide a counterpoint to the concept of authentic assessment (see Chapter 3, Section 3.4.3.3). The examples described above highlight that authenticity is more complex than the binary authentic/inauthentic, or even a single continuum of more/less authenticity. It highlights the importance of discerning which underlying principles assessment practices are authentic to, and prompts the question of whether fuller authenticity (for example, in the form of a sovereign code) need always be the goal.

8.5 Summary

This chapter was the last of three chapters to discuss conceptual approaches for musical performance assessment. It was a response to the third guiding research question: How can bases of legitimation for assessment practices in

higher music education be meaningfully conceptualised? The focus in this chapter was on using concepts from the Autonomy dimension of LCT to conceptualise the insulation of *musical* performance assessment practices. Section 8.2 revisited the basic context of the research, highlighting the apparent issue of musical authenticity as a determinant of the legitimacy of assessment practices. Section 8.2 reintroduced the Autonomy dimension of LCT, and Section 8.3 focused on the concepts of positional autonomy and relational autonomy to realise the insulation of constituents of musical performance assessment practices (positional autonomy) and ways of working within musical performance assessment practices (relational autonomy). Section 8.3 provided some simple devices for realising these concepts, which in turn provided points of departure for theorising the characteristics of Autonomy codes in the context of musical performance assessment (Section 8.4). Autonomy codes provide one means of distinguishing why some assessment participants may have cohesive relationships with musical performance assessment while others may not.

Together, Chapters 6, 7, and 8 described a range of tools for conceptualising different aspects of musical performance assessment practices. The discussions developed in these chapters explored intersections between concepts from LCT and various elements from the substantive context of musical performance assessment. Each of these discussions aimed to respond to the guiding research questions by translating between the worlds of theory and musical performance assessment to propose some ways of viewing practices that relate to real contexts. The following chapter summarises these discussions and concludes the thesis.

Chapter 9: Conclusions

9.1 Introduction

This is the concluding chapter of the thesis. Following the previous three chapters which responded to the guiding research questions, this chapter summarises the research and discusses the implications of the work. Section 9.2 reviews the chapters comprising the thesis. Section 9.3 reflects on the the responses to the research questions developed in the later chapters of the thesis and considers the implications of the work. Section 9.4 discusses the contribution of the thesis, its limitations, and directions for future research. Section 9.5 is the end note for the thesis.

9.2 Summary of the Thesis

This thesis has included nine chapters, laid out in five main parts. The first part of the thesis comprised the introductory chapter, which provided an overview of the problem-situation to which the work is a response and defined the guiding research questions for the research. At the broader level this work is a response to the need for alternative ways of viewing assessment that account for its socially-situated nature, and in so doing, to develop more meaningful ways of understanding assessment practices. Assessment in music education at the broader level is characterised by competing sets of values, and traditional approaches and philosophies remain contested. In responding to this situation, the focus in this research was not on developing better techniques or approaches for musical performance assessment. Rather, this research focused on some of the underlying mechanisms of musical performance assessment practices that influence the ways in which practices unfold. In this work these are characterised by languages of legitimation—actions, sayings, thoughts, and feelings that make claims for the legitimacy of one thing over another. The research was guided by three questions which were developed to focus the study:

1. How can bases for achievement in musical performance work be meaningfully conceptualised?

2. How can bases of legitimation for assessment participants be meaningfully conceptualised?

3. How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised?

Following the introductory chapter, the second part of the thesis was the conceptual framework, which was comprised of Chapters 2, 3, and 4. The conceptual framework of the thesis describes the main points of departure according to which the research was oriented and developed. Crucially, the conceptual framework was an integrated part of the research, evolving and informing the research over the duration of the study: it is comprised of a range of resources, or modules, that include theoretical and philosophical positions, literature, personal perspectives (Appendix A), and contextual interpretations. Although the conceptual framework was structured across three chapters it should be thought of as a unified aspect of the research. As well as informing the research, the framework is also an outcome of the research—it defines a collective point of departure for studying assessment practices that will hopefully be of use to other scholars and in future work. In this, the framework is one of the starting points offered by this thesis—it is hoped that it will become a resource for assessment research that others will build upon and extend in their own work.

Chapter 2 introduced the conceptual framework by describing its qualities, including its integrated nature and iterative development. The main inspiration for the approach that was taken is the work of Maxwell (2009, 2013) who describes the architectural properties of conceptual frameworks and how they differ to other research products such as literature reviews. Accordingly, this thesis did not include a review *of* literature in the traditional sense, but rather addressed literature *for* the research. This is an important distinction, since a review of assessment literature alone would not have adequately contextualised the study. The main reason for this is that the issues in which the thesis is interested (Chapter 3) are generally not obvious in much of the literature on assessment. Chapter 2 provided an overview of the broader theoretical points of departure for the research including a description of a realist stance, and a view of assessment as a social practice characterised by

languages of legitimation. Chapter 2 also introduced in a simple way the first concept from Legitimation Code Theory (LCT) mentioned in the thesis, which was semantic gravity. The suggestion in that chapter was that semantic gravity offers a useful way of conceptualising the contextual specificity and relatedness of research constituents (literature, perspectives, concepts and theories, and so forth), which is useful for recognising ways in which concepts typified by a greater semantic distance from assessment in more specific music education contexts (for example, studies of assessment concepts at a non-disciplinary level) nonetheless apply and relate to these contexts (or not). Space precluded a deeper exploration of this application of semantic gravity which provided a useful thinking tool throughout the research—this can be explored more specifically in future work. The final part of Chapter 2 described the contextual framing of the research by explaining the relevant features of the Australian higher music education context in which the work was conducted, and how musical performance assessment was located within this.

Chapter 3 continued the conceptual framework by focusing more closely on situating the research questions that were developed to guide the research. In situating these questions, this chapter described a series of gaps to which the questions are oriented. These gaps included:

- The need for alternative means of conceptualising bases of achievement in musical performance assessment that move beyond typological depictions of criteria
- The need to focus on assessment participants and their attributes in and of themselves, rather than focussing on assessment techniques, approaches, and trustworthiness
- The need for meaningful ways of interpreting the relationships between positions on the legitimation of assessment practices (including their finer aspects), to challenge oppositional thinking by making the underlying principles of positioning more explicit

Chapter 4 was the final chapter to describe the conceptual framework and explained the main features of LCT that were adopted in the study. It provided an overview of the general architecture of LCT highlighting its multidimensional

structure. The main aspects of LCT that informed this research were the Specialization, Semantics, and Autonomy dimensions, the 4-K Model (within the Specialization dimension), and the concepts of cosmologies and constellations which are not directly within any dimension of LCT. As a part of the conceptual framework, the different concepts from LCT described in Chapter 4 were used in different ways. Some were more explicitly operationalised in the later chapters of the thesis, and these included the Specialization dimension, the Autonomy dimension, the concept of gaze produced through the interaction of subjective relations and interactional relations from within the 4-K model, and the concept of constellations. The Semantics dimension is used less explicitly, although the concept of semantic gravity is introduced in Chapter 2. The entire suite of concepts is provided in Chapter 4 for the reason that the framework as a whole had a notable influence on the development of the research as part of the conceptual framework—this influence is discussed further in Appendix A.

Chapter 5 described the qualitative methodology that informed the research, and explained the main perspectives according to which data in the project were construed. Chapter 5 also explained the analytical approach taken in the research, and distinguished between multiple complementary strategies for enacting qualitative analysis. These strategies included both categorising and connecting strategies, which together provided means for the development of translation devices which describe how concepts from LCT are related to data. Chapter 5 also discussed the trustworthiness of the research.

Chapters 6, 7, and 8 responded directly to the research questions that guided the study—as noted in Chapter 5, these chapters can be likened to long-form translation devices which discuss in detail connections between theoretical concepts and data which act as referents for real-world practices. These chapters are reflected upon more fully in the following section. The remaining parts of the thesis are this chapter, which summarises and concludes the thesis, and Appendix A. Appendix A includes an extended researcher memo, which describes in personal language the experience of conducting the research. To this end, Appendix A provides additional detail to complement what is included in the body of the thesis and includes extra detail

describing the way in which the research unfolded, as well as the influential challenges, successes, and realisations that occurred throughout the project.

9.3 Responding to the Research Questions

This research focused on conceptual approaches for several aspects of musical performance assessment practices. In this, the aims of the research were generative rather than reductionist—the focus was on exploring possibilities rather than tightly defined probabilities. As noted in the previous section, the basic theoretical point of departure was an interest in assessment practices as socially-situated activities characterised by languages of legitimation. This was emphasised in different ways by each of the research questions developed to guide the research. The following sections revisit each of these questions to offer some final thoughts on their exploration in the thesis.

9.3.1 Conceptualising assessment criteria

The first research question asked: How can bases for achievement in musical performance work be meaningfully conceptualised? A response to this question was explored in Chapter 6 which used concepts from the Specialization dimension of LCT to theorise criteria for musical performance assessment. The main aim in exploring this question was to explore an alternative to typological representations of criteria, which are a fundamental regulator of students' achievement in educational assessment (Chapter 3). The Specialization dimension of LCT offers some possibilities for doing this by highlighting kinds of relations that can be seen to underlie substantive criteria—epistemic relations underlying techniques and knowledge, and social relations underlying values and attributes. To this end, these concepts provide one means of reconceptualising some of the typical problem-categories that are commonly discussed in the appraisal of musical performance. Some examples of problem-categories include artistry, flair, originality, and expression. In contrast with typological conceptions (for example, technique-artistry typologies), the view here is of these characterised by the relative co-occurrence of emphasis on epistemic relations and social relations. The strength of the approach outlined in Chapter 6 is that the languages of

explanation explored are not tethered to particular substantive instances—they accommodate contrasting interpretations of substantive concepts (for example, flair, originality) across contexts. In this, it avoids the fixed relegation of concepts to categories by moving to a topological approach, rather than a typological one. Chapter 6 developed some starting points for interpreting the characteristics of this topological space (characterised by Specialization codes) which can be adapted to suit more specific contexts.

9.3.1.1 So what?

One possible application of this work is that it might usefully inform assessment participants' assessment literacy. Rather than providing predetermined criteria, thinking explicitly about different kinds of specialisation provides a way to intuit similarities, differences, and preferences, which may be a particularly useful professional awareness in both holistic and analytic assessment situations. Specialization codes offer useful terms for discussing differences between criteria for the reason that they provide a way to avoid meticulous specification of terms that remain unavoidably subjective. While it may seem overly simplistic to suggest that assessment practices could benefit from explicit discussion of criteria—in the sense of being explicit about their underlying characteristics—the act of making these characteristics transparent could have powerful effects. The main asset here is the smaller number of theoretical concepts (the internal language of description) which enable simple communication without reducing the range of substantive possibilities for realising them. This is a good example of the advantage of a topological view—it may be a useful device for promoting dialogue and for discussing different ways of seeing what is important in a musical performance. For example, arguments about why a student's performance was or was not of a certain quality in highly descriptive terms can be difficult to manage for the reason that different perspectives may inadvertently speak past one another. The recognition of clashes between knowledge and knower codes provides a basis (and a language) for engaging with the situation to explore ways in which alternative approaches might be possible that would produce more cohesive outcomes. This is facilitated by translation between the theoretical concepts and more specific ideas about criteria. Enabling a dialogue that moves

between theory and description is a powerful way to envisage alternatives, or at the least to develop empathy with alternative perspectives. It is a means of more clearly articulating the rules of the game *underlying* practice rather than assuming all of the players involved are equally equipped to understand those rules. The argument here, insofar as criteria are concerned, is that seeing criteria is not enough—actively considering *how we see criteria* can have a substantial influence on how they are understood and enacted. By working in terms of degrees of emphasis upon sets of relations the materials developed in Chapter 6 offer a means of getting beneath the complexities subsumed by types to provide a more nuanced view of what matters in a particular situation.

One of the main aims in Chapter 6 was to provide a starting point for realising theoretical concepts in practice. The devices described are heuristic, and the data to which they are calibrated will not reflect the particularities of other contexts. Rather, they are resources for interpretation, and offer an additional language for discussing bases of achievement in musical performance assessment practices.

9.3.2 Conceptualising bases of success for participation in assessment

The second research question asked: How can bases of legitimation for assessment participants be meaningfully conceptualised? The main response to this question was developed in Chapter 7, which also predominantly used concepts from the Specialization dimension of LCT. In exploring this question, Chapter 7 is a response to the deficit of assessment research that foregrounds assessment participants as the primary interest. As explained in Chapter 3, this position is not predicated on a notion that assessment participants are absent in research about assessment (which would be quite untrue), but is rather a recognition of the dominance of approaches, techniques, practices, and trustworthiness over *people* (see Chapter 3). The focus in Chapter 7, as such, is on conceptualising the attributes of assessment participants to recognise bases upon which successful participation in assessment is determined.

In Chapter 7, the concepts of epistemic relations and social relations were used to theorise epistemic attributes and social attributes. An important learning described in this chapter was that the structural properties of these

sets of attributes are not necessarily alike. Where epistemic attributes were more easily classifiable in terms of distinct constellations of knowledge, social attributes appeared not to be comprised of such readily available structure. To this end, additional concepts from the 4-K Model were used to illustrate a different way of conceptualising social attributes.

9.3.2.1 So what?

This analysis highlighted that differences in the underlying bases of aspects of practices (here, the attributes of assessment participants) could necessitate different ways of conceptualising their structural properties. To use an analogy: one would not use a ruler to weigh a block of butter. This in turn has implications for the ways in which assessment participants are equipped to participate in assessment practices. Although assessment participants may possess sound attitudes towards assessment, a lack of knowledge (epistemic attributes) may prevent them from accessing discourse about it. Similarly, a privileging of technical knowledge of assessment is not by definition a recipe for effective communication with those whose relationships to practice are experiential rather than trained. This in turn has implications for concepts such as assessment literacy which sometimes appear to privilege knowledge (a basis in epistemic relations) over knowing (a basis in social relations). It leads us to consider closely the relationship between formalised knowledge about assessment concepts and theories, and the attributes of those held responsible for knowing about them or putting them into practice. An emphasis on the former paints a picture of technical knowledge to be employed by those who are able to access it, while an emphasis on the latter paints a picture of something closer to personal theories of assessment that involve the coupling of assessment concepts with the personal attributes of those involved. Selecting the appropriate mix of these approaches invariably means taking into account the broader context of a given educational situation, however in order to make this selection it is helpful to be aware of what the underlying bases of the possible approaches are, and therefore what kinds of emphasis or de-emphasis are being applied to knowledge and/or knowing.

One conceptual resource for making visible the basis upon which possession (or lack thereof) of particular epistemic attributes or social

attributes can influence the legitimation of assessment participants is the concept of gaze, which was discussed in the latter part of Chapter 7. The various modalities of gaze described in Chapter 7 offer means of realising how who you are matters. For those enacting legitimation, gazes describe the standpoint from which this takes place in terms of what kinds of attributes are deemed important. The possession of a particular gaze is, at the same time, a basis upon which assessment participants may be legitimated. As in the case of criteria above, the usefulness of this device is in its capacity to make explicit bases for the legitimation of assessment participants in ways that depart from meticulous empirical description. In addition to highlighting the ways in which assessment participants view one another, the various modalities of gaze also offer a resource for stepping outside of personal perspectives.

9.3.3 Conceptualising the legitimation of assessment practices

The third research question asked: How can bases of legitimation for assessment practices in higher music education be meaningfully conceptualised? The main response to this question was developed in Chapter 8, in which the Autonomy dimension of LCT was used to provide a theoretical language for the legitimation of practices. In responding to this research question, Chapter 8 addresses a need for ways of recognising how positions on assessment are marked out and related to one another. Autonomy provides one means of doing this by focusing on the insulation of practices in terms of both their constituents (positional autonomy) and their ways of working (relational autonomy). The discussion developed in Chapter 8 focuses on the insulation of one particular characterisation of assessment practices which reflects the musical authenticity of assessment practices. Although the the discussion developed in Chapter 8 was focused on the relative *musicality* of aspects of assessment practices, different analyses could be produced that focused on other categories as the core object of focus: for example, regulatory accountability, or student learning.

In translating the concepts of positional autonomy and relational autonomy, interpretations of the characteristics of Autonomy codes are made possible, and some examples are developed in Chapter 8. These codes illustrate relationships between different aspects of musical performance

assessment practices and the concept of musical authenticity. These codes are:

- Sovereign codes, which describe assessment practices and their constituents considered to be musically authentic
- Introjected codes, which reflect a more flexible bounding of legitimate constituents, although the legitimacy of practices remains contingent upon subscription to ways of working considered to be musically authentic
- Projected codes, which reflect an emphasis on musically authentic constituents and more weakly bounded ways of working
- Exotic codes, which reflect less musically authentic practices, both in terms of their constituents and their ways of working

9.3.3.1 So what?

The reason for focusing in Chapter 8 on the insulation of musical authenticity in assessment practices is that musical authenticity is of clear importance to assessment practitioners in a range of ways. These include the authenticity of criteria, the authenticity of task designs, the authenticity of assessment participants themselves, and so forth. The Autonomy codes provide a topological space for positioning these different aspects of practices relative to core musically authentic constituents and approaches. These codes therefore provide a useful means for interpreting the flavour of practices, which could be construed as more or less musical (and perhaps more or less legitimate), but in different ways. Examining how participants in practices relate to different codes is one way make more transparent the legitimation of practices: for example, a projected code wherein non-musical assessment strategies are coupled with musical constituents (e.g. the performance, the venue, professional musicians as assessors) may clash with a sovereign code wherein alternative assessment strategies considered to be more musical are more highly valued.

9.3.4 Additional thoughts

9.3.4.1 On the value of being explicit

The main reason for which the conceptual approaches explored and developed in this thesis could be conceived of as *meaningful* (Chapter 1) is that they provide resource for making implicit aspects of practices explicit by providing a language of description. The implicit aspects in question are the organising, underlying principles of practices, not their substantive features. For example, the idea that musical development should involve different emphases on developing technical skills and interpretive or expressive skills at different points in students' programs is a simple enough statement, however what is actually meant by technical, interpretive, or expressive often turns out to be difficult to pin down. While the language itself is explicit (we may state that different aspects of educational programs emphasise different kinds of skills), what is actually meant may not be transparent to all involved: One person's technique may be another person's interpretation. This is not an argument for more meticulous description, but rather, for more effective communication. Translation back and forth between specific substantive concepts and more theoretical concepts is one way to develop a deeper sense of what matters in an assessment practice without being required to precisely define it in some absolute way. For assessment in music, precise language is often unavailable, and the argument here is that by focusing on underlying characteristics of practice we retain the possibility of diversity without being overwhelmed by the diverse range of possibilities.

9.3.4.2 On forward motion within the field of assessment

The relativist code from the Specialization dimension of LCT was introduced in Chapter 4 and briefly discussed in Chapter 6. It is typified by a lack of emphasis on either epistemic relations (knowledge and skills) or social relations (values and attributes). Although the relativist code is not inherently bad, the de-emphasis on both kinds of relations has implications for knowledge-building and knower-building within the code. Put simply, how can powerful knowledge, skills, or ways of understanding and viewing the world be productively constructed if both are de-emphasised? At the field level, one way

in which this kind of lack of emphasis can be produced is through the establishment of multiple competing specializations within the field. Maton (2014), in his research on the field of British cultural studies, observed that the emergence of competing positions within the field precluded cumulative forward progress. It does not seem unreasonable to propose that the field of educational assessment faces similar risks. Chapter 3 described different paradigms for assessment (measurement and learning) as well as a range of different approaches for assessment that do not necessarily co-exist. On a personal level, distinctions between these paradigms were keenly felt throughout the project as I travelled to various parts of the world to engage with fellow assessment researchers and to attend various conferences and events. The differences were striking, and are more fully described in Appendix A. Individually I observed coherent identities in the various communities that I visited: what I came to perceive as a knowledge code where measurement was the dominant paradigm (the emphasis being on statistically valid and reliable specification) and a knower code where feedback and students' learning were the main focus (the emphasis being on the characteristics and experiences of students). This is a broad personal provocation, and one which assuredly warrants more meticulous investigation. What I intend to highlight by this provocation, however, is the risk inherent in proceeding in different directions for the field of assessment as a whole. This thesis has described some resources for engaging with this issue, and comparative case studies are one way in which future research may be able to reveal the presence and character of these divisions in more specific contexts.

9.4 Contributions, Limitations, and Directions for Future Research

9.4.1 Contribution of the research

The principal contribution of this thesis is to develop the extant conceptual language available to meaningfully describe and interpret assessment practices in higher music education contexts. To this end, the research adopted LCT as an existing explanatory framework on the basis that it offered some useful points of departure for developing conceptual language. Although

LCT has been employed in studies of assessment and music, the two had not previously been brought together with a focus on musical performance. This work constitutes a bridging of that gap and is offered as a detailed point of departure for further research at the intersection of these two fields. To this end, the main practical contribution of this thesis is to offer starting points— both for further research and for practice—by developing languages of description (Chapters 6, 7, and 8). These languages provide means of meaningfully conceptualising assessment practices in ways that foreground the systems of legitimation that underlie social practices. To this end, a broader contribution of this thesis is to respond to calls for alternative ways of viewing assessment practices in higher education (Chapter 1), and in turn, to re-emphasise the need for meaningful conceptual language and devices in the field (Chapter 3). In addition to the work developed in the later chapters of the thesis, the conceptual framework (Chapters 2, 3, & 4) is offered as a resource for other studies that take social practice approaches in studies of assessment.

9.4.2 Limitations of the research

This research has several main limitations, some of which have been mentioned or alluded to earlier in the thesis. From a methodological standpoint, the research is limited in terms of the kinds of insights it provides. As a theoretical qualitative study, it does not provide any kind of statistically robust findings in the sense that quantitative studies seek to do. Indeed, even within the qualitative paradigm this research is not oriented to a particular case or site to the extent that deep insights about practice in particular contexts are reliably provided. This has implications for the kinds of generalisability that are available to the research (Chapter 5, Section 5.5). While this is a deliberate feature of the research design it does mean that an anchor to a more specific substantive context is absent in this research. This is a limitation particularly in view of the quantity of abstract conceptual material dealt with in the work which ultimately may result in the work being less accessible. To use theoretical terminology, this work expresses weaker semantic gravity and relatively high semantic density. To this end, future work could explore alternative methodological approaches in more tightly defined settings.

Related to the previous point, a second limitation relates to the conceptual density of the work. To paraphrase Maton (2019), LCT occasionally causes headaches. The explicit conceptual focus in this thesis means that despite its interest in practice it remains at a distance from actual practices, and so applications of the concepts developed in this thesis remain to be explored in closer encounters with actual practices in defined contexts. This limitation is related to the directions for future research discussed below. Added to this is the adoption of LCT as the solitary set of more specific theoretical tools that inform the work. The problem-situation outlined in Chapter 1 does not preclude the use of other theoretical devices in responding to it, and future work may be able to make productive use of the conceptual framework by exploring other theoretical approaches—some possibilities are provided below.

A third limitation is that the work was conducted by a single researcher, as distinct from a research team. While access to expert perspectives from doctoral supervisors and colleagues (see Appendix A) were an important resource in counteracting the risk of an unbalanced, solitary perspective, the work ultimately remains that of an individual. This limitation was to a large extent unavoidable given the doctoral study arrangement, and future work will benefit from the involvement of collaborators.

9.4.3 Directions for future research

Further to the directions proposed so far in this chapter there are several key directions available for further study to build upon this research. The main possibilities can be thought of in terms of two related kinds, research directions and practical directions. Related to the discussion of limitations above, the first main research direction that warrants attention is the application of the conceptual framework in studies of practice in more specific contexts—for example, within the scope of a single department at a university, or across several disciplinary areas within an institution. This thesis has sought to make available a range of resources to this end, and therefore provides a range of possible avenues for studies of substantive practices. In this, future work should not seek to adopt the entire range of concepts discussed in this thesis, but rather to adopt only what is necessary for the intended study. This direction

of study is likely to be useful firstly for the reason that it will provide opportunities for the refinement of conceptual approaches, and secondly for the relationship such studies may have to actual practices. These kinds of studies may provide concrete avenues for better understanding and engaging with substantive practices in specific contexts.

The second main research direction that can be followed is to focus more deeply on applying and refining individual concepts discussed in this thesis in studies of educational assessment. This thesis offers an array of starting points, and more specific programs of research might focus on building upon these. One possibility, for example, is to build upon the use of semantic gravity discussed in Chapter 2 (Section 2.4) to examine more closely the relationships between disciplinary assessment contexts and assessment concepts and practices that are less tethered to disciplines. Related to this is a third possible direction, which is to explore intersections between the concepts discussed in this work and other theories. Chapter 4 highlighted some possibilities, including Actor Network Theory (see Bearman & Ajjawi, 2018, for an accessible introduction in an assessment context) and Systemic Functional Linguistics, which is already used in conjunction with LCT (see for example Martin, Maton, & Doran, 2020). Other possibilities include intersections with frameworks, such as Biggs (1993) 3P model of teaching and learning, or other kinds of practice theory (for some suggestions see Boud et al., 2016; Schatzki, 2012).

On the practical side, there is potential for the conceptual approaches developed in this thesis to inform actual practical activities. One possibility is to incorporate them into tools for professional development and training to provide staff (and perhaps students) with means of making sense of practice. To this end, concepts can be simplified as necessary—for instance, epistemic relations and social relations can form the basis of productive conversations about the focus in assessment designs, to facilitate conversations about students' developmental trajectories, to help assessors intuit their personal perspectives relative to those of others, and so forth. This last possibility highlights possibilities for moderation, or in holistic assessment contexts, what Sadler (2013) refers to as *calibration*. It is not necessary to use technical terminology in contexts where this is inappropriate—for example, with staff or

students who have no background in theory of the kind discussed in this work. Wolff (2015) provides a useful example of how theory can be adapted for practice in this way—in the context of engineering education she adapted the epistemic plane (Chapter 4, Section 4.3.4), modifying the terminology to facilitate ease of understanding for her participants.

9.5 End Note

This thesis studied musical performance assessment practices, with the aim of exploring possibilities for the meaningful theoretical interpretation of various aspects of these practices. Ultimately, the resources and directions proposed in this research are a starting point. There are many possibilities for application and adaptation of the concepts, and the means of this will vary depending upon more specific contexts. Hopefully this thesis has provided the reader with some fresh ways of thinking about assessment, both at the general level as well as at the level of more specific music performance education contexts. This work has consciously avoided a trope of much educational assessment research, which is advocacy for one approach or another—in this it is for all assessment practitioners, regardless of paradigm.

References

Alasuutari, P. (1995). *Researching culture: Qualitative method and cultural studies*. London, England: Sage.

Allais, S. (2012a). Claims vs. practicalities: Lessons about using learning outcomes. *Journal of Education and Work, 25*(3), 331–354. https://doi.org/10.1080/13639080.2012.687570

Allais, S. (2012b). 'Economics imperialism', education policy and educational theory. *Journal of Education Policy, 27*(2), 253–274. https://doi.org/10.1080/02680939.2011.602428

Annett, J. (2002). Subjective rating scales: Science or art? *Ergonomics, 45*(14), 966–987. https://doi.org/10.1080/00140130210166951

Antia, B., & Kamai, R. (2017). Writing biology, assessing biology: The nature and effects of variation in terminology. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication, 22*(2), 201–222. https://doi.org/10.1075/term.22.2.03ant

Archbald, D. A., & Newmann, F. M. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school*. Reston, VA: National Association of Secondary School Principals Assessment.

Archer, M. (1995). *Realist social theory*. Cambridge, England: Cambridge University Press.

Atkinson, P. A., & Coffey, A. (2004). Analysing documentary realities. In D. Silverman (Ed.), *Qualitative research: Theory, method, practice* (2nd ed., pp. 45–62). London, England: Sage.

Australian Government. (2015). *Higher education standards framework*. Retrieved from https://www.legislation.gov.au/Details/F2015L01639

Australian Qualifications Framework Council. (2013). *Australian qualifications framework* (2nd ed.). Retrieved from www.aqf.edu.au

Baird, J. A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, Policy & Practice, 24*(3), 317–350. https://doi.org/10.1080/0969594X.2017.1319337

Barratt, E., & Moore, H. (2005). Researching group assessment: Jazz in the conservatoire. *British Journal of Music Education, 22*(3), 299–314. http://doi.org/10.1017/S0265051705006467

Bearman, M., & Ajjawi, R. (2018). Actor-network theory and the OSCE: Formulating a new research agenda for a post-psychometric era. *Advances in Health Sciences Education, 23*(5), 1037–1049. https://doi.org/10.1007/s10459-017-9797-7

Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories* (2nd ed.). Buckingham, England: The Society for Research into Higher Education & Open University Press.

Becker, H. S. (1986). *Writing for social scientists: How to start and finish your thesis, book, or article*. Chicago, IL: University of Chicago Press.

Becker, H. S. (2007). *Writing for social scientists: How to start and finish your thesis, book, or article* (2nd ed.). Chicago, IL: University of Chicago Press.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment*

in Education: Principles, Policy & Practice, 18(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Berg, D. N., & Smith, K. K. (Eds.). (1988). *The self in social inquiry: Research methods*. Newbury Park, CA: Sage.

Bergee, M. J. (2003). Faculty interjudge reliability of music performance evaluation. *Journal of Research in Music Education, 51*(2), 137–150. https://doi.org/10.2307%2F3345847

Bergee, M. J., & Cecconi-Roberts, L. (2002). Effects of small-group peer interaction on self-evaluation of music performance. *Journal of Research in Music Education, 50*(3), 256–268. https://doi.org/10.2307/3345802

Bernstein, B. (1975). *Class, codes and control, volume III: Towards a theory of educational transmissions*. London, England: Routledge & Kegan Paul.

Bernstein, B. (2000). *Pedagogy, symbolic control, and identity: Theory, research, critique* (2nd ed.). Lanham, ML: Rowman & Littlefield.

Biglan, A. (1973a). Relationships between subject matter characteristics and the structure and output of university departments. J*ournal of Applied Psychology, 57*(3), 204–213. https://doi.org/10.1037/h0034699

Biglan, A. (1973b). The characteristics of subject matter in different academic areas. *Journal of applied Psychology, 57*(3), 195–203. https://doi.org/10.1037/h0034701

Biggs, J. B. (1993). From theory to practice: A cognitive systems approach. *Higher Education Research and Development, 12*(1), 73–85. https://doi.org/10.1080/0729436930120107

Biggs, J. (1998). Assessment and classroom learning: A role for summative assessment? *Assessment in Education: Principles, Policy & Practice, 5*(1), 103–110. https://doi.org/10.1080/0969595980050106

Blackie, M. A. (2014). Creating semantic waves: Using Legitimation Code Theory as a tool to aid the teaching of chemistry. *Chemistry Education Research and Practice, 15*(4), 462–469. https://doi.org/10.1039/C4RP00147H

Blom, D., & Encarnacao, J. (2012). Student-chosen criteria for peer assessment of tertiary rock groups in rehearsal and performance: What's important? *British Journal of Music Education, 29*(1), 25–43. https://doi.org/10.1017/S0265051711000362

Blom, D., & Poole, K. (2004). Peer assessment of tertiary music performance: Opportunities for understanding performance assessment and performing through experience and self-reflection. *British Journal of Music Education, 21*(1), 111–125. https://doi.org/10.1017/S0265051703005539

Blom, D., Stevenson., I, & Encarnacao, J. (2015). Assessing music performance process and outcome through a rubric: Ways and means. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 125–140). Cham, Switzerland: Springer.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning.* New York, NY: McGraw-Hill.

Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher*

education: A practical guide. Maidenhead, England: Open University Press.

Bourdieu, P. (1998). *Practical reason: On the theory of action* (P. Press, Trans.). Stanford, CA: Stanford University Press.

Boud, D. (1995). Assessment and learning: Contradictory or complementary? In P. Knight (Ed.), *Assessment for learning in higher education* (pp. 35–48). London, England: Kogan Page.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education, 22*(2), 151–167. https://doi.org/10.1080/713695728

Boud, D. (2006). Foreword. In D. Carless, G. Joughin, N-F. Liu, & Associates (Eds.), *How assessment supports learning: Learning-oriented assessment in action* (pp. ix–x). Aberdeen, Hong Kong: Hong Kong University Press.

Boud, D. (2010). Assessment for developing practice. In J. Higgs, D. Fish, I. Goulter, S. Loftus, J-A. Reid, & F. Trede (Eds.), *Education for future practice* (pp. 251–262). Rotterdam, The Netherlands: Sense Publishers.

Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (Eds.). (2018). *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work*. Abingdon, England: Routledge.

Boud, D., Dawson, P., Bearman, M., Bennett, S., Joughin, G., & Molloy, E. (2016). Reframing assessment research: Through a practice perspective. *Studies in Higher Education, 43*(7), 1107–1118. https://doi.org/10.1080/03075079.2016.1202913

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399–413. https://doi.org/10.1080/02602930600679050

Boudon, R. (1980). *The crisis in sociology*. New York, NY: Columbia University Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706QP063OA

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners* [Kindle version]. London, England: Sage.

Bridgstock, R. (2009). The graduate attributes we've overlooked: Enhancing graduate employability through career management skills. *Higher Education Research & Development, 28*(1), 31–44. https://doi.org/10.1080/07294360802444347

Broadfoot, P. (1996). *Education, assessment and society: A sociological analysis*. Buckingham, England: Open University Press.

Broadfoot, P. (2012). *Assessment, schools and society*. Abingdon, England: Routledge.

Brophy, T. S. (2019). Assessment in music education: The state of the art. In T. S. Brophy (Ed.), *The Oxford handbook of assessment policy and practice in music education* (Vol. 2, pp. 902–932).

Brown, G. T. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment*

Matters, 3(1), 45–70. Retrieved from https://search.informit.com.au/documentSummary;dn=5157559398508 62;res=IELHSS

Brown, G. T., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education, 27*(1), 210–220. https://doi.org/10.1016/j.tate.2010.08.003

Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London, England: Kogan Page.

Burke, P. J. (2012). *The right to higher education: Beyond widening participation*. Abingdon, England: Routledge.

Burnard, P., & Fautley, M. (2015). Assessing diverse creativities in music: A spectrum of challenges, possibilities and practices. In M. Fleming, L. Bresler, & J. O'Toole (Eds.), *The Routledge international handbook of the arts and education* (pp. 254–267). Abingdon, England: Routledge.

Burnard, P., Trulsson, Y. H., & Söderman, J. (Eds.). (2015). *Bourdieu and the sociology of music education*. Surrey, England: Ashgate.

Burwell, K., Carey, G., & Bennett, D. (2019). Isolation in studio music teaching: The secret garden. *Arts and Humanities in Higher Education, 18*(4), 372–394. https://doi.org/10.1177/1474022217736581

Butkovic, A., & Modrusan, I. (2019). Personality differences among musicians: Real differences or stereotypes? *Psychology of Music,* 1– 11. https://doi.org/10.1177%2F0305735619849625

Cain, M. (2015). Participants' perceptions of fair and valid assessment in tertiary music education. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 87–106). Cham, Switzerland: Springer.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research, 42*(3), 294–320. https://doi.org/10.1177%2F0049124113500475

Carey, G., & Lebler, D. (2012). Reforming a Bachelor of Music programme: A case study. *International Journal of Music Education, 30*(4), 312–327. https://doi.org/10.1177%2F0255761412459160

Carey, G., & Grant, C. (2015). Teacher and student perspectives on one-to-one pedagogy: Practices and possibilities. *British Journal of Music Education, 32*(1), 5–22. https://doi.org/10.1017/S0265051714000084

Carey, G., Grant, C., McWilliam, E., & Taylor, P. (2013). One-to-one pedagogy: Developing a protocol for illuminating the nature of teaching in the conservatoire. *International Journal of Music Education, 31*(2), 148–159. https://doi.org/10.1177%2F0255761413483077

Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education, 69*(6), 963–976. http://doi.org/10.1007/s10734-014-9816-z

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325.

https://doi.org/10.1080/02602938.2018.1463354

Carroll, C. L. (2017). *Playing the field: An Australian case study of student popular musicians' informal learning in senior secondary classroom music education* (Doctoral dissertation, University of Sydney, Sydney, Australia). Retrieved from https://ses.library.usyd.edu.au/handle/2123/17232

Carroll, C. L. (2019). Seeing the invisible: Theorising connections between informal and formal musical knowledge. *Research Studies in Music Education*, 1–19. https://doi.org/10.1177/1321103X18824641

Chen, T-H. (2010). *Knowledge and knowers in online learning: Investigating the effects of online flexible learning on student sojourners* (Doctoral dissertation, University of Wollongong, Wollongong, Australia). Retrieved from https://ro.uow.edu.au/theses/3099/

Clarence, S. (2016). Exploring the nature of disciplinary teaching and learning using Legitimation Code Theory Semantics. *Teaching in Higher Education, 21*(2), 123–137. https://doi.org/10.1080/13562517.2015.1115972

Coll, R. K., & Zegwaard, K. E. (2006). Perceptions of desirable graduate competencies for science and technology new graduates. *Research in Science & Technological Education, 24*(1), 29–58. https://doi.org/10.1080/02635140500485340

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Daniel, R. (2001). Self-assessment in performance. *British Journal of Music Education, 18*(3), 215–226. https://doi.org/10.1017/S0265051701000316

Daniel, R. (2004). Peer assessment in musical performance: The development, trial and evaluation of a methodology for the Australian tertiary environment. *British Journal of Music Education, 21*(1), 89–110. https://doi.org/10.1017/S0265051703005515

Daniel, R. J., & Parkes, K. A. (2015). Assessment and critical feedback in the master-apprentice relationship: Rethinking approaches to the learning of a music instrument. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 107–124). Cham, Switzerland: Springer.

Dann, R. (2014). Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice, 21*(2), 149–166. https://doi.org/10.1080/0969594X.2014.898128

Davidson, J. W., & Da Costa Coimbra, D. (2001). Investigating performance evaluation by assessors of singers in a music college setting. *Musicae Scientiae, 5*(1), 33–53. https://doi.org/10.1177%2F102986490100500103

Davidson, J. W. & King, E. C. (2004). Strategies for ensemble practice. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 105–122). Oxford, England: Oxford University Press.

Dawson, P., Bearman, M., Boud, D. J., Hall, M., Molloy, E. K., Bennett, S., & Joughin, G. (2013). Assessment might dictate the curriculum, but what dictates assessment? *Teaching and Learning Inquiry, 1*(1), 107–111. https://doi.org/10.2979/teachlearninqu.1.1.107

Deeley, S. J., & Bovill, C. (2017). Staff student partnership in assessment: Enhancing assessment literacy through democratic practices. *Assessment & Evaluation in Higher Education, 42*(3), 463–477. https://doi.org/10.1080/02602938.2015.1126551

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2000). *The SAGE handbook of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Department of Education and Training. (2015). *Higher education in Australia: A review of reviews from Dawkins to today*. Canberra, Australia: Australian government.

Department of Employment, Education and Training. (1988). *Higher education: A policy statement*. Canberra, Australia: Australian government.

Doran, Y. J. (2019). *Knower-building: Building axiological constellations in the humanities*. Keynote presented at the Third International Legitimation Code Theory Conference, Johannesburg, South Africa.

Doran, Y. J. (In press). Cultivating values: Knower-building in the humanities. *Estudios Lingüistica Aplicada*. Retrieved from https://www.researchgate.net/profile/Yaegan_Doran/publication/33529 2157_Cultivating_values_Knower-building_in_the_humanities/links/5d5cf371299bf1b97cfa32f8/Cultivating-values-Knower-building-in-the-humanities.pdf

Elliott, C. A. (1995/1996). Race and gender as factors in judgments of musical performance. *Bulletin of the Council for Research in Music Education, 127*, 50–56. Retrieved from https://www.jstor.org/stable/40318766

Elton, L., & Johnston, B. (2002). *Assessment in universities: A critical review of research*. Retrieved from https://eprints.soton.ac.uk/59244/1/59244.pdf

Fautley, M. (2010). *Assessment in music education*. Oxford, England: Oxford University Press.

Fautley, M., & Murphy, R. (2015). Editorial. *British Journal of Music Education, (32)*1, 1–4. https://doi.org/10.1017/S0265051714000400

Filer, A. (Ed.). (2000) *Assessment: Social practice and social product*. London, England: Routledge.

Fiske Jr, H. E. (1975). Judge-group differences in the rating of secondary school trumpet performances. *Journal of Research in Music Education, 23*(3), 186–196. https://doi.org/10.2307%2F3344643

Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher education, 64*(1), 119–133. https://doi.org/10.1007/s10734-011-9484-1

Forbes, M. (2016). If the shoe doesn't fit: A case and a place for collaborative learning for music practice in higher education. *Australian Journal of Music Education, 50*(1), 53–62. Retrieved from https://search.informit.com.au/documentSummary;dn=9031012947023 73;res=IELHSS

Garraway, J., & Reddy, L. (2016). Analyzing work-integrated learning assessment practices through the lens of autonomy principles. *Alternation, 23*(1), 285–308. Retrieved from http://digitalknowledge.cput.ac.za/handle/11189/5902

Gellner, E. (1959). *Words and things*. London, England: Routledge.

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education,* (1), 3—31. Retrieved from http://eprints.glos.ac.uk/3609/

Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education, 24*, 355–392. https://doi.org/10.2307/1167274

Ginsborg, J., & Wistreich, R. (2010). *Promoting excellence in small group music performance: Teaching, learning and assessment*. Retrieved from https://www.advance-he.ac.uk/knowledge-hub/promoting-excellence-small-group-music-performance-teaching-learning-and-assessment

Glaser, B. G. (2001). *The grounded theory perspective: Conceptualization contrasted with description*. Mill Valley, CA: Sociology Press.

Goodman, E. (2002). Ensemble performance. In J. Rink (Ed.), *Musical performance: A guide to understanding* (pp. 152–167). Cambridge, England: Cambridge University Press.

Grenfell, M. (Ed.). (2008). *Pierre Bourdieu: Key concepts*. Durham, England: Acumen.

Griffith University. (2016a). *Research ethics manual: Beneficence: Risks and benefits in human research*. Brisbane, Australia: Griffith University.

Griffith University. (2016b). *The Griffith graduate*. Retrieved from https://policies.griffith.edu.au/pdf/The%20Griffith%20Graduate.pdf

Griffith University. (2020a). 1360QCM_Y1 Performance Study 1 (Part 1) course profile: Aims, outcomes & graduate attributes. Brisbane, Australia: Griffith University. Retrieved from https://courseprofile.secure.griffith.edu.au/student_section_loader.php?section=2&profileId=113923

Griffith University. (2020b). 1360QCM_Y1 Performance Study 1 (Part 1) course profile: Assessment plan—Y2 recital. Brisbane, Australia: Griffith University. Retrieved from https://courseprofile.secure.griffith.edu.au/student_section_loader.php?section=5&profileId=113923

Gynnild, V. (2016). Assessing vocal performances using analytical assessment: A case study. *Music Education Research, 18*(2), 224–238. https://doi.org/10.1080/14613808.2015.1025732

Halliday, M. A. K. (1979). Modes of meaning and modes of expression: Types of grammatical structure, and their determination by different semantic functions. In D. J. Allerton, E. Carney, & D. Holcroft (Eds.), *Function and context in linguistics analysis: Essays offered to William Haas* (pp. 57–79). Cambridge, England: Cambridge University Press.

Harrison, S. (2014). *Pro-active music higher degrees: Final report*. Brisbane, Australia: Griffith University.

Harrison, S. D., Lebler, D., Carey, G., Hitchcock, M., & O'Bryan, J. (2013). Making music or gaining grades? Assessment practices in tertiary music ensembles. *British Journal of Music Education*, *30*(1), 27–42. https://doi.org/10.1017/S0265051712000253

Heron, J. (1988). Assessment revisited. In D. Boud (Ed.), *Developing student autonomy in learning* (pp. 77–90). London, England: Kogan Page.

Holliday, A. (2007). *Doing and writing qualitative research* (2nd ed.). London, England: Sage.

Hyry-Beihammer, E. K. (2010). Master-apprentice relation in music teaching: From a secret garden to transparent modelling. *Nordic Research in Music Education Yearbook 12*, 161–178. Retrieved from https://nmh.brage.unit.no/nmh-xmlui/handle/11250/172285

Iusca, D. (2014). The effect of evaluation strategy and music performance presentation format on score variability of music students' performance assessment. *Procedia - Social and Behavioral Sciences, 127*, 119–123. https://doi.org/10.1016/j.sbspro.2014.03.224

James, D., & Diment, K. (2003). Going underground? Learning and assessment in an ambiguous space. *Journal of Vocational Education and Training, 55*(4), 407–422. https://doi.org/10.1080/13636820300200242

Jansen, G., & Peshkin, A. (1992). Subjectivity in qualitative research. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 681–725). San Diego, CA: Academic Press.

Johnson, P. (1997). Performance as experience: The problem of assessment criteria. *British Journal of Music Education, 14*(3), 271–282. https://doi.org/10.1017/S0265051700001248

Jørgensen, H. (2002). Instrumental performance expertise and amount of practice among instrumental students in a conservatoire. *Music Education Research, 4*(1), 105–119. https://doi.org/10.1080/14613800220119804

Jørgensen, H. (2010). Higher music education institutions: A neglected arena for research? *Bulletin of the Council for Research in Music Education, 186,* 67–80. Retrieved from www.jstor.org/stable/41110435

Kingsbury, H. (2001). *Music, talent, & performance: A conservatory cultural system.* Philadelphia, PA: Temple University Press.

Knight, P., & Yorke, M. (2003). *Assessment, learning and employability.* Maidenhead, England: Open University Press.

Kolb, D. A. (1981). Learning styles and disciplinary differences. In A. Chickering (Ed.), *The modern American college* (pp. 232–255). San Francisco, CA: Jossey-Bass.

Lamont, A., & Maton, K. (2008). Choosing music: Exploratory studies into the low uptake of music GCSE. *British Journal of Music Education, 25*(3), 267–282. https://doi.org/10.1017/S0265051708008103

Lamont, A., & Maton, K. (2010). Unpopular music: Beliefs and behaviours towards music in education. In R. Wright (Ed.), *Sociology and music education* (pp. 63–80). Basingstoke, England: Ashgate.

Lau, A. M. S. (2016). 'Formative good, summative bad?'–A review of the dichotomy in assessment literature. *Journal of Further and Higher Education, 40*(4), 509–525. https://doi.org/10.1080/0309877X.2014.984600

Lebler, D. (2006, July). *3D assessment: Looking through a learning lens.*

Paper presented at the International Seminar of the Commission for the Education of the Professional Musician, Hanoi, Vietnam.

Lebler, D. (2008). Popular music pedagogy: Peer learning in practice. *Music Education Research, 10*(2), 193–213. https://doi.org/10.1080/14613800802079056

Lebler, D. (2015). Future directions for assessment in music. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 1–7). Cham, Switzerland: Springer.

Lebler, D. (2019). Music assessment in Australian higher education. In T. S. Brophy (Ed.), *The Oxford handbook of assessment policy and practice in music education* (Vol. 1, pp. 197–216).

Lebler, D., Carey, G., & Harrison, S. (Eds.). (2015a). *Assessment in music education: From policy to practice*. Cham, Switzerland: Springer.

Lebler, D., Carey, G., & Harrison, S. (2015b). Glossary. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 295–296). Cham, Switzerland: Springer.

Lebler, D., & Harrison, S. (2017). Evaluating progress and setting directions. In J. Rink, H. Gaunt, & A. Williamon (Eds.), *Musicians in the making: Pathways to creative performance* (pp. 93–107). New York, NY: Oxford University Press.

Lebler, D., Harrison, S., Carey, G., & Cain, M. (2015). *An approach to aligning assessment with threshold learning outcomes in the creative and performing arts*. Brisbane, Australia: Griffith University.

Locke, L., Spirduso, W. W., & Silverman, S. J. (1999). *Proposals that work* (4th ed.). Thousand Oaks, CA: Sage Publications.

Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy & Practice, 25*(5), 442–467. https://doi.org/10.1080/0969594X.2016.1268090

Manns, J. W. (1998). *Aesthetics*. New York, NY: M.E. Sharpe.

Martin, J. L. (2013). *On notes and knowers: The representation, evaluation and legitimation of jazz* (Doctoral dissertation, University of Adelaide, Adelaide, Australia). Retrieved from http://hdl.handle.net/2440/84700

Martin, J. L. (2016). Musicality and musicianship: Specialization in jazz studies. In K. Maton, S. Hood, & S. Shay (Eds.), *Knowledge-building: Educational studies in Legitimation Code Theory* (pp. 193–213). Abingdon, England: Routledge.

Martin, J. L. (2018). Writing about music: The selection and arrangement of notation in jazz students' written texts. *British Journal of Music Education, 35*(1), 73– 89. https://doi.org/10.1017/S0265051717000171

Martin, J. R. (1992). *English text: System and structure*. Philadelphia, PA: John Benjamins.

Martin, J. R., Maton, K., & Y. J. Doran. (Eds.). (2020). *Accessing academic*

discourse: Systemic Functional Linguistics and Legitimation Code Theory. Abingdon, England: Routledge.

Mason, J. (1996). *Qualitative researching*. London, England: Sage.

Maton, K. (2004). *The field of higher education: A sociology of reproduction, transformation, change and the conditions of emergence for cultural studies* (Doctoral dissertation, University of Cambridge, Cambridge, England). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.718.2864&rep=rep1&type=pdf

Maton, K. (2007). Knowledge-knower structures in intellectual and educational fields. In F. Christie, & J. Martin (Eds.), *Language, knowledge and pedagogy: Functional linguistic and sociological perspectives* (pp. 87–108). London, England: Continuum.

Maton, K. (2014). *Knowledge and knowers: Towards a realist sociology of education*. Abingdon, England: Routledge.

Maton, K. (2016a). Legitimation Code Theory: Building knowledge about knowledge-building. In K. Maton, S. Hood, & S. Shay (Eds.), *Knowledge-building: Educational studies in Legitimation Code Theory* (pp. 1–23). Abingdon, England: Routledge.

Maton, K. (2016b). Starting points: Resources and architectural glossary. In K. Maton, S. Hood, & S. Shay (Eds.), *Knowledge-building: Educational studies in Legitimation Code Theory* (pp. 233–243). Abingdon, England: Routledge.

Maton, K. (2019). *The rules of the game: Why what we do matters*. Keynote presented at the Third International Legitimation Code Theory Conference, Johannesburg, South Africa.

Maton, K. (2020). Semantic waves: Context, complexity and academic discourse. In J. R. Martin, K. Maton, & Y. J. Doran (Eds.), *Accessing academic discourse: Systemic Functional Linguistics and Legitimation Code Theory* (pp. 59–86). Abingdon, England: Routledge.

Maton, K., & Chen, R. T-H. (2016). LCT in qualitative research: Creating a translation device for studying constructivist pedagogy. In K. Maton, S. Hood, & S. Shay (Eds.), *Knowledge-building: Educational studies in Legitimation Code Theory* (pp. 27–48). Abingdon, England: Routledge.

Maton, K., & Chen, R. T-H. (in press). Specialization from Legitimation Code Theory: How the basis of achievement shapes student success. In J. R. Martin, K. Maton, P. Wang, & Z. Wang (Eds.), *Understanding academic discourse*. Shanghai, China: Shanghai Jiao Tong University.

Maton, K., Hood, S., & Shay, S. (Eds.). (2016). *Knowledge-building: Educational studies in Legitimation Code Theory*. Abingdon, England: Routledge.

Maton, K., & Howard, S. K. (2018). Taking autonomy tours: A key to integrative knowledge-building, *LCT Centre Occasional Paper 1* (June): 1–35. Retrieved from

http://legitimationcodetheory.com/home/publications/lct-centre-occasional-papers/

Maton, K., & Moore, R. (2010a). Introduction: Coalitions of the mind. In K. Maton & R. Moore (Eds.), *Social realism, knowledge and the sociology of education: Coalitions of the mind* (pp. 1–13). London, England: Continuum.

Maton, K., & Moore, R. (Eds.). (2010b). *Social realism, knowledge and the sociology of education: Coalitions of the mind*. London, England: Continuum.

Matruglio, E., Maton, K., & Martin, J. R. (2013). Time travel: The role of temporality in enabling semantic waves in secondary school teaching. *Linguistics and Education, 24*(1), 38–49. https://doi.org/10.1016/j.linged.2012.11.007

Maxwell, J. A. (2006). Literature reviews of, and for, educational research: A commentary on Boote and Beile's "scholars before researchers". *Educational Researcher, 35*(9), 28–31. https://doi.org/10.3102/0013189X035009028

Maxwell, J. A. (2009). Designing a qualitative study. In L. Bickman & D. J. Rog (Eds.), *The SAGE handbook of applied social research methods* (pp. 214–253). Thousand Oaks, CA: Sage.

Maxwell, J. A. (2012). *A realist approach for qualitative research*. Thousand Oaks, CA: Sage.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: Sage.

McLean, M., Abbas, A., & Ashwin, P. (2013). The use and value of Bernstein's work in studying (in)equalities in undergraduate social science education. *British Journal of Sociology of Education, 34*(2), 262–280. https://doi.org/10.1080/01425692.2012.710007

McPherson, G. E., & Schubert, C. (2004). Measuring performance enhancement in music. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 61–82). Oxford, England: Oxford University Press.

McPherson, G. E., & Thompson, W. F. (1998). Assessing music performance: Issues and influences. *Research Studies in Music Education, 10*(1), 12–24. https://doi.org/10.1177%2F1321103X9801000102

McWilliam, E., Carey, G., Draper, P., & Lebler, D. (2006, February). Learning and unlearning: New challenges for teaching in conservatoires. Paper presented at The Reflective Conservatoire Conference: Apprentices and Sorcerers, London, England. Retrieved from http://hdl.handle.net/10072/13298

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Mills, J. (1991). Assessing musical performance musically. *Educational*

Studies, 17(2), 173–181. https://doi.org/10.1080/0305569910170206

Monkhouse, H. (2010). Using assessment effectively: Learning environments that work for tertiary music performance students and staff. Retrieved from https://altf.org/wp-content/uploads/2016/08/Monkhouse_H_TF_Final-report-2010.pdf

Monkhouse, H. (2015). Bachelor of music: Purpose, desires and requirements. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 62–73). Cham, Switzerland: Springer.

Moore, R. (2008). Capital. In M. Grenfell (Ed.), *Pierre Bourdieu: Key concepts* (pp. 101–118). Durham, England: Acumen.

Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research, 62*(3), 229–258. https://doi.org/10.3102%2F00346543062003229

Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher, 23*(2), 5–12. https://doi.org/10.3102%2F0013189X023002005

Moss, P. A. (1996). Enlarging the dialogue in educational measurement: Voices from interpretive research traditions. *Educational Researcher, 25*(1), 20–29. https://doi.org/10.3102%2F0013189X025001020

Muckler, F. A., & Seven, S. A. (1992). Selecting performance measures: "Objective" versus "subjective" measurement. *Human Factors, 34*(4), 441–455. https://doi.org/10.1177%2F001872089203400406

Newsome, E. (2015). A search for balance: The development of a performance assessment form for classical instrumental music in the tertiary context. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 153–170). Cham, Switzerland: Springer.

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education, 14*(2), 149–170. https://doi.org/10.1080/09695940701478321

Nisbet, I., & Shaw, S. D. (2019). Fair assessment viewed through the lenses of measurement theory. *Assessment in Education: Principles, Policy & Practice, 26*(5), 612–629. https://doi.org/10.1080/0969594X.2019.1586643

Norton, A., & Cherastidtham, I. (2018). *Mapping Australian higher education (Grattan Institute Report No. 2018-11)*. Retrieved from https://grattan.edu.au/report/mapping-australian-higher-education-2018/

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*(1), 115–138. https://doi.org/10.1146/annurev.psych.48.1.115

Parkes, K. A. (2010). Performance assessment: Lessons from performers. *International Journal of Teaching and Learning in Higher Education, 22*(1), 98–106. Retrieved from https://eric.ed.gov/?id=EJ913534

Parkinson, T., & Smith, G. D. (2015). Towards an epistemology of authenticity in higher popular music education. *Action, Criticism and Theory for Music Education, 14*(1), 93–127. Retrieved from https://kar.kent.ac.uk/69350/

Partti, H., Westerlund, H., & Lebler, D. (2015). Participatory assessment and the construction of professional identity in folk and popular music programs in Finnish and Australian music universities. *International Journal of Music Education, 33*(4), 476–490. https://doi.org/10.1177/0255761415584299

Perkins, R. (2013). Learning cultures and the conservatoire: An ethnographically-informed case study. *Music Education Research, 15*(2), 196–213. https://doi.org/10.1080/14613808.2012.759551

Pike, K. L. (1982). *Linguistic concepts: An introduction to tagmemics*. Lincoln, NE: University of Nebraska Press.

Polifonia Working Group. (2014). *Final report on assessment and standards*. Retrieved from https://www.aec-music.eu/polifonia/working-groups/assessment--standards

Popper, K. (2005). *Realism and the aim of science* (W. W. Bartley III, Ed.). Abingdon, England: Routledge.

Price, M., Carroll, J., O'Donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: A critical commentary on current assessment practice. *Assessment & Evaluation in Higher Education, 36*(4), 479–492. https://doi.org/10.1080/02602930903512883

Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education, 34*(1), 1–20. https://doi.org/10.1080/03054980701476386

Pulman, M. (2009). Seeing yourself as others see you: developing personal attributes in the group rehearsal. *British Journal of Music Education, 26*(2), 117–135. https://doi.org/10.1017/S0265051709008390

Reason, P. (1988). Introduction. In P. Reason (Ed.), *Human inquiry in action: Developments in new paradigm research* (pp. 1–17). Newbury Park, CA: Sage.

Robson, C. (2011). *Real world research* (3rd ed.). Chichester, England: John Wiley & Sons Ltd.

Rootman-le Grange, I., & Blackie, M. A. (2018). Assessing assessment: In pursuit of meaningful learning. *Chemistry Education Research and Practice, 19*(2), 484–490. https://doi.org/10.1039/C7RP00191F

Rowntree, D. (1987). *Assessing students: How shall we know them?* (2nd ed.). London, England: Kogan Page.

Russell, B. E. (2010). *The empirical testing of a musical performance assessment paradigm* (Doctoral dissertation, University of Miami, Coral Gables, Florida). Retrieved from https://scholarlyrepository.miami.edu/oa_dissertations/387/

Russell, B. E. (2015). An empirical study of a solo performance assessment model. *International Journal of Music Education, 33*(3), 359–371.

https://doi.org/10.1177%2F0255761415581282

Rust, C. (2002). Purposes and principles of assessment. *Learning and Teaching Briefing Papers Series*, 3–5. Retrieved from http://ltac.emu.edu.tr/DOCUMENTS/p_p_assessment.pdf

Rust, C., O'Donovan, B., & Price, M. (2005). A social constructivist assessment process model: How the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education, 30*(3), 231–240. https://doi.org/10.1080/02602930500063819

Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education, 13*(2), 191–209. https://doi.org/10.1080/0305498870130207

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education, 30*(2), 175–194. https://doi.org/10.1080/0260293042000264262

Sadler, D. R. (2009a). 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, D. R. (2009b). Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 45–64). Dordrecht, Netherlands: Springer.

Sadler, D. R. (2010). Fidelity as a precondition for integrity in grading academic achievement. *Assessment & Evaluation in Higher Education, 35*(6), 727–743. https://doi.org/10.1080/02602930902977756

Sadler, D. R. (2013). Assuring academic achievement standards: From moderation to calibration. *Assessment in Education: Principles, Policy & Practice, 20*(1), 5–19. https://doi.org/10.1080/0969594X.2012.714742

Sadler, D. R. (2014). The futility of attempting to codify academic achievement standards. *Higher Education, 67*(3), 273–288. https://doi.org/10.1007/s10734-013-9649-1

Sadler, D. R. (2015) Backwards assessment explanations: Implications for teaching and assessment practice. In D. Lebler, G. Carey & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 9–19). Cham, Switzerland: Springer.

Sadler, D. R. (In press). Assessment tasks as curriculum statements: A turn to attained outcomes. *Japanese Journal of Curriculum Studies.* Retrieved from https://www.researchgate.net/publication/338669993_Assessment_tasks_as_curriculum_statements_A_turn_to_attained_outcomes

Sayer, A. (2000). *Realism and social science*. London, England: Sage.

Schatzki, T. (2012). A primer on practices. In J. Higgs, R. Barnett, S. Billet, M. Hutchings, & F. Trede (Eds.), *Practice-based education: perspectives and strategies* (pp. 13–26). Rotterdam, The Netherlands: Sense Publishers.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation*

(Vol. 1, pp. 39–83). Chicago, IL: Rand McNally.

Shay, S. B. (2004). The assessment of complex performance: A socially situated interpretive act. *Harvard Educational Review, 74*(3), 307–329. http://doi.org/10.17763/haer.74.3.wq16l67103324520

Shay, S. (2005). The assessment of complex tasks: A double reading. *Studies in Higher Education, 30*(6), 663–679. http://doi.org/10.1080/03075070500339988

Shay, S. (2008a). Beyond social constructivist perspectives on assessment: The centring of knowledge. *Teaching in Higher Education, 13*(5), 595–605. https://doi.org/10.1080/13562510802334970

Shay, S. (2008b). Researching assessment as social practice: Implications for research methodology. *International Journal of Educational Research, 47*(3), 159–164. http://doi.org/10.1016/j.ijer.2008.01.003

Silverman, D. (2011) *Interpreting Qualitative Data* (4th ed.). London, England: Sage Publications.

Stanley, M., Brooker, R., & Gilbert, R. (2002). Examiner perceptions of using criteria in music performance assessment. *Research Studies in Music Education, 18*(1), 46–56. https://doi.org/10.1177%2F1321103X020180010601

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan, 72*(7), 534–539.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. New York, NY: Cambridge University Press.

Swanwick, K. (1998). The perils and possibilities of assessment. *Research Studies in Music Education, 10*(1), 1–11. https://doi.org/10.1177/1321103X9801000101

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education, 76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3

Tertiary Education Quality and Standards Agency. (2017). *Glossary of terms*. Retrieved from https://www.teqsa.gov.au/glossary-terms

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York, NY: MacMillan.

Thompson, S., & Williamon, A. (2003). Evaluating evaluation: Musical performance assessment as a research tool. Music Perception: An Interdisciplinary Journal, 21(1), 21–41. https://doi.org/10.1525/mp.2003.21.1.21

Thompson, W. F., Diamond, C. T. P., & Balkwill, L. (1998). The adjudication of six performances of a Chopin etude: A study of expert knowledge. *Psychology of Music, 26*(2), 154–174. https://doi.org/10.1177%2F0305735698262004

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education, 14*(3), 281–294. https://doi.org/10.1080/09695940701591867

Torrance, H. (2017). Blaming the victim: Assessment, examinations, and the

responsibilisation of students and teachers in neo-liberal governance. *Discourse: Studies in the Cultural Politics of Education, 38*(1), 83–96. https://doi.org/10.1080/01596306.2015.1104854

Tregear, P., Johansen, G., Jørgensen, H., Sloboda, J., Tulve, H., & Wistreich, R. (2016). Conservatoires in society: Institutional challenges and possibilities for change. *Arts and Humanities in Higher Education, 15*(3–4), 276–292. https://doi.org/10.1177/1474022216647379

Tuck, R. (2007). *An introductory guide to national qualifications frameworks: Conceptual and practical issues for policy makers.* Geneva, Austria: International Labour Office.

Vaira, M. (2004). Globalization and higher education organizational change: A framework for analysis. *Higher Education, 48*(4), 483–510. http://doi.org/10.1023/B:HIGH.0000046711.31908.e5

van Heerden, M., Clarence, S., & Bharuthram, S. (2017). What lies beneath: Exploring the deeper purposes of feedback on student writing through considering disciplinary knowledge and knowers. *Assessment & Evaluation in Higher Education, 42*(6), 967–977. https://doi.org/10.1080/02602938.2016.1212985

Vella, R., & English, H. (2015). Embedding creative and critical thinking in performance studies—the challenge. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 141–151). Cham, Switzerland: Springer.

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education, 43*(5), 840–854. https://doi.org/10.1080/02602938.2017.1412396

Vu, T. T., & Dall'Alba, G. (2014). Authentic assessment for student learning: An ontological conceptualisation. *Educational Philosophy and Theory, 46*(7), 778–791. https://doi.org/10.1080/00131857.2013.795110

Wald, N., & Harland, T. (2017). A framework for authenticity in designing a research-based curriculum. *Teaching in Higher Education, 22*(7), 751–765. https://doi.org/10.1080/13562517.2017.1289509

Wanner, T., & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. *Assessment & Evaluation in Higher Education, 43*(7), 1032–1047. https://doi.org/10.1080/02602938.2018.1427698

Wesolowski, B. C. (2012). Understanding and developing rubrics for music performance assessment. *Music Educators Journal, 98*(3), 36–42. https://doi.org/10.1177%2F0027432111432524

Wesolowski, B. C., Wind, S. A., & Engelhard Jr, G. (2015). Rater fairness in music performance assessment: Evaluating model-data fit and differential rater functioning. *Musicae Scientiae, 19*(2), 147–170.

https://doi.org/10.1177%2F1029864915589014

Westerlund, H. (2006). Garage rock bands: A future model for developing musical expertise? *International Journal of Music Education, 24*(2), 119–125. https://doi.org/10.1177/0255761406065472

Wiliam, D. (2017). Learning and assessment: a long and winding road? *Assessment in Education: Principles, Policy & Practice, 24*(3), 309–316. https://doi.org/10.1080/0969594X.2017.1338520

Williamon, A. (1999). The value of performing from memory. *Psychology of Music, 27*(1), 84–95. https://doi.org/10.1177%2F0305735699271008

Wilmot, K. D. (2019). *Enacting knowledge in dissertations: An exploratory analysis of doctoral writing using Legitimation Code Theory.* (Doctoral dissertation, University of Sydney, Sydney, Australia). Retrieved from https://ses.library.usyd.edu.au/handle/2123/20498

Wilson, J. (2018). *Artists in the university: Positioning artistic research in higher education.* Singapore: Springer.

Wolff, K., & Hoffman, F. (2014). 'Knowledge and knowers' in engineering assessment. *Critical studies in Teaching and Learning, 2*(1), 74–95. http://doi.org/10.14426/cristal.v2i1.19

Wolff, K. (2015). Insights into conceptual and contextual engineering problem-solving practices. In *Proceedings of the 3rd Biennial Conference of the South African Society for Engineering Education* (pp. 189–198). Durban, South Africa: SASEE.

Wright, R. (2010). *Sociology and music education.* Farnham, England: Ashgate.

Wrigley, W. J. (2005). *Improving music performance assessment* (Doctoral dissertation, Griffith University, Brisbane, Australia). Retrieved from https://research-repository.griffith.edu.au/bitstream/handle/10072/367747/02Whole.pdf?sequence=1

Zhukov, K. (2015). Challenging approaches to assessment of instrumental learning. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in music education: From policy to practice* (pp. 55–70). Cham, Switzerland: Springer.

Appendix A: Researcher Memo

This memo supports the thesis by providing extra information about how the study unfolded. In part this is to support the transparency of the research. Although most qualitative research methodologists emphasise the importance of this, it is my personal observation that personal accounts are nonetheless fairly uncommon in actual publications. Another part of the intention behind including this memo is to provide those readers who might also be undertaking doctoral studies (especially at the beginning) with a human account of the process. While I found other published theses to be very useful resources for cultivating my personal understanding of what it means to conduct a doctoral research project, these documents also rarely explain the twists and turns that (if conversations with peers are to be believed) are frequently the reality of doctoral study. To that end, and reflecting one of the underlying themes of this thesis, this memo is also in part an attempt to help others see the rules of the game more explicitly by making visible my own experiences. Accordingly, this memo is framed as a personal account of the research, and I have attempted to document as many of the influential moments (a.k.a.: twists and turns) in the project as possible.

Personal Background

The decision to undertake the project followed a 5-year period of undergraduate study, which included four years (2011–2014) of study in a jazz performance program and an additional year of study (2015) devoted to an Honours level research project. In Australia, Honours is built-in to some undergraduate programs, while it is an optional extra path of study in others. In general, the purpose of Honours level study is to provide undergraduates with an intensive research education which will equip them to pursue further study at Masters or Doctoral levels. I used my Honours study as an opportunity to explore a process that had caused me a good deal of stress and frustration as an undergraduate student, which was musical performance assessment.

The underlying reasons for choosing to study this are linked to the broader identity I had formed over my lifetime in relation to learning. As a school student in a rural Australian town I had always been a high-achiever in

the academic sense. I developed a meaningful relationship with music as a teenager (at the age of 12 to be exact), and at the age of 13 I began to learn the bass guitar. Most of my musical forays during this part of my life were in garage bands, wherein we played mostly rock and metal music (incidentally, I now spend most of my time playing acoustic folk music). My decision to enrol in a jazz performance program was heavily driven by the fact that other kinds of music programs (generally labelled as popular music or creative music making) did not provide one-to-one lessons as a part of their curricula. As an undergraduate, I was both highly motivated and highly limited—many of my peers were already equipped with jazz knowledge and skills that extended far beyond my own. In this environment, assessment became a high-pressure focal point of my studies, and I tended to treat it as a proving ground for myself. During the entire four years of my undergraduate program, I was awarded the same grade at the end of every semester—a 6 out of a possible 7, which is generally synonymous in Australia with a distinction (4 is generally regarded a pass, 5 a credit, and 7 a high-distinction). Framed as a percentage, institutions generally award distinctions to students who achieve in the 75–85% region in coursework (or 70–80% at some institutions).

I left my undergraduate degree relatively frustrated at my inability to break past the distinction, and I began to wonder at that time what the cause might be. Some questions I asked myself at that time were a) what if I was simply unable to execute the skills necessary, b) what if I had developed a reputation early in my degree as a distinction-level student—might this be inhibiting my ability to achieve more highly, and c) what if simply following the task description was not actually the basis for succeeding? Underlying all of this was a strong but fairly naive concern with whether we actually really knew what we were doing when it came to assessment. Although I now believe this concern to have been well-placed, at that time it was instinctive, intuitive, and very 21-year-old. This basic interest inspired the trajectory for my Honours research project, during which time I met the people who would become the principle supervisors for my doctoral work.

Beginning the Research

By the close of my Honours studies I felt assured that my interest in assessment was well-placed, and my earlier free-form concerns had been replaced by a more focused question: How can we possibly assess musical performance? It is *so* subjective. Incidentally, this is the comment that most people make at parties when the topic of my research arises: "Ah, that sounds interesting, music is so subjective". My question was rendered problematic as I began to discover the work that already attended to this issue. There seemed to be a range of possible solutions, and people seemed to disagree fairly resolutely about the best way assess both music and other kinds of creative work. Influential studies at the beginning of the research included the suite of papers by Sadler included in the reference list of this thesis, as well as Wrigley's (2005) doctoral thesis, which seemed to ask the very questions I had been concerned with, and which also seemed to offer a very concrete solution. The main problem was that there were other fairly concrete solutions too, and they often seemed to contradict one another (see for a contrasting example Sadler, 2015). My earlier question was replaced by another: What is this thing (musical performance assessment)? The evaporation of my initial question started a process that I have called *unriddling* in conversations about my research. My earlier question was replaced by something that was more similar to a sentiment or a sense that needed to be unriddled to reveal a coherent question. In this sense, the entire doctorate can be thought of as a process of unriddling during which the focal point of the research crystallised little by little.

Coming to Theory

This research is heavily oriented towards theory, and this is mainly (I think) the result of searching for a way to *understand* assessment as part of my unriddling process. I became consistently frustrated with the literature on educational assessment in this respect, which seemed to offer very many concepts that seemed like nice ideas with limited practical value beyond characterising the general flavour of one type of assessment or another. Further, scholars seemed to take firm stances on the benefits and limitations

of groups of concepts (for example, analytic and holistic assessment) without acknowledging the presence of what was clearly a difference in underlying philosophies about assessment. As part of what was intended to be a thorough literature review process I studied and developed well-referenced summaries of many assessment concepts, however I consistently lacked the means to meaningfully link these into the research project. In part, it was this experience that confirmed the need to look at assessment in a way that diverged from a focus on kinds of assessment, and why some might be better or worse.

An influential gateway concept within the project was the notion of practice (in the sense of assessment practice), which seemed to routinely crop up as a means for simply describing what people do. Delving deeper into the idea of practice led to my first encounter with the work of Pierre Bourdieu later in the first year of the project. Bourdieu's concepts of habitus, capital, and field provided an appealing alternative to the plethora of assessment concepts, in that they felt like helpful ideas for thinking about assessment—I later learned that Bourdieu's ideas have been described by others as "good to think with". Discovering an edited book focusing on Bourdieu and music education (Burnard, Trulsson, & Söderman, 2015) provided helpful illustrations of his concepts in action, and led to the discovery of a second edited book (Wright, 2010) which incorporated other sociological ideas. This book introduced me to the ideas of Bernstein and a focus on knowledge in practices which seemed to complement my existing interest in Bourdieu. There is comparatively less interest in Bernstein's ideas in relation to music education (although there is some), however the inclusion of Lamont and Maton (2010) in Wright's book (Wright, 2010) provided another gateway. In seeking to learn more about Bernstein's ideas (which were difficult to interpret without prior exposure), Maton's publications seemed to offer useful snippets of explanation. A characteristic limitation of both Bourdieu's and Bernstein's ideas is that means of actually putting them to work (beyond thinking with them) are not overly clear. Seeking to better understand how to apply these ideas in research led to the discovery of Maton (2014) which seemed to provide some solutions, albeit in language that seemed very difficult to understand at that earlier stage of the process. After a significant amount of reading and re-reading Maton's text I became interested in exploring applications of LCT in my project more

deeply. In addition to describing LCT, Maton (2014) also provides insightful discussion of peripheral topics (for example, theory) which expanded my perspectives information, knowledge, and meaning significantly. As I began to work with the ideas of Bourdieu, Bernstein, and Maton to try to better understand assessment (a process of years) the theoretical focus of the work seemed to emerge quite organically. The focus of the research gradually shifted from an emphasis on the broad question of how to assess, to a focus on viewing assessment. Confirmation of the value of this direction was provided when I encountered the works of Shay (2004, 2008a, 2008b) and Boud et al. (2016), who seemed to agree that conceptualising assessment was a focus worth pursuing.

Conferences, Colleagues, and the Research Trajectory

Attending conferences and having the opportunity to catch up with or meet new colleagues had a major influence on the trajectory of the project. The conferences that I attended over the course of the project were:

- Transforming one-to-one teaching and learning symposium (May, 2016, Brisbane, Australia)
- Higher degree research symposium (2016, Brisbane, Australia)
- The Sixth International Symposium on Assessment in Music Education (ISAME6, April, 2017, Birmingham, England)
- The Seventh International Symposium on Assessment in Music Education (ISAME7, March, 2019, Gainesville, Florida, USA)
- The Artistic Research Symposium (November, 2018, Brisbane, Australia)
- Assessment in Higher Education (June, 2019, Manchester, England)
- The Third International Legitimation Code Theory Conference (July, 2019, Johannesburg, South Africa)

In addition to these conferences several other opportunities were sought that had an important influence on the project. These were 1) a week spent in Sydney (Australia) in 2017 visiting the Legitimation Code Theory Centre, and

2) a week spent in Melbourne (Australia) in 2019 visiting the Centre for Research in Assessment and Digital Learning.

All of these experiences provided valuable insights through exposure to different fields, experts within those fields, and different ways of thinking. This was a very important aspect in terms of developing a sense about where the project fit into the research landscape. Experiences at the ISAME conferences helpfully provided a context for discussing assessment in music education, however they also highlighted some important differences between my work and other assessment work being conducted within the field of music education. In particular, these conferences featured a strong emphasis on a) measurement, and b) pre-tertiary music education and music teacher education—there seemed to be little discussion of musical performance assessment in higher education, and less sociological and/or socio-cultural work (particularly at ISAME7). Since most of my Australian assessment colleagues and the scholars whose work I was most familiar generally fit into the learning paradigm of assessment (as distinct from the measurement paradigm—see Chapter 3 of the thesis), the strong focus on measurement at these conferences was disconcerting, and it was difficult to see how my work related to the general interests of the community. Attending the Assessment in Higher Education (AHE) Conference was similarly disconcerting— this community was in many ways the opposite of the ISAME community, and was much more strongly aligned with the learning paradigm. An important realisation that came after these experiences was that both communities seemed to be broadly advocating for particular ways of making sense of assessment. In LCT terms (see Chapter 4 of the thesis), the ISAME community (characterised by a much stronger American presence) seemed to strongly embody a knowledge code (being focussed on measurement), while the AHE community (characterised by a much stronger British and European presence) seemed to strongly embody a knower code, where the focus was much more strongly on the students' learning than on theoretical perspectives. Exposure to both communities provided a valuable opportunity to calibrate my own understanding of the broader field, however it was also in some ways an isolating experience, in that it seemed to validate a personal suspicion that the broader field of assessment was fractured into inward-facing groups.

Attending the Legitimation Code Theory Conference (LCT3) directly after AHE was a revelatory experience, for the reason that the theoretical focus of my work was a much more natural fit. This community included a small number of scholars working in music and assessment, though not at the same time. Interactions with the LCT community were a crucial resource in learning about the framework, and it was during LCT3 that the broader aim of my research became much clearer. A good deal of LCT research could be broadly described as case studies. A likely reason for this is that LCT, being an analytic explanatory framework, operates very effectively in case study research. The aims of my research however, were slightly different. In bringing together the fields of assessment in higher education, assessment in music, and LCT, my research could more accurately be thought of as a translational work that connected these regions. This insight was a very important milestone in the unriddling process, and helped to galvanise the research at a point where my personal motivation was suffering from a lack of being able to identify strongly with a research community. This led to several further insights which were 1) a reaffirmation that my primary target audience was the music assessment community, and 2) that I would need to spend much more time exploring the framework of LCT as a part of my study, rather than incorporating it as a distinct theoretical module for conducting analysis (which is perhaps the usual application of the framework). As this research trajectory crystallised I became much more cogent of the fact that the value of the study was not in explaining assessment practices in a particular specific situation, but rather, that it was about trying to find a way to look at assessment differently, and to figure out how to translate this for my audience. To this end, colleagues in the various research communities were immensely helpful in enabling me to test and calibrate my ideas. My supervisors deserve much acknowledgement for persevering with and supporting my interest in a theoretical direction that was largely unfamiliar to them at the outset.

On Data

The role of data in the research took some time to properly understand, and this understanding formed alongside the development of the research trajectory. Since the aims of this research were primarily theoretical and

translational, the role of data was not ultimately to provide information about assessment practices in a particular setting. Rather, data became a resource for developing connections between theoretical ideas and the field of musical performance assessment. This was not immediately apparent at the outset, and coming to this perspective was an important part of the process of unriddling the research. This eventually led to the conclusion that the literature in which I had been immersed since the beginning of the project were in fact a valuable source of data. My early frustrated encounters with the cottage industry of assessment research (an analogy from Knight & Yorke, 2003) were put to better use as I began to see the various positions and ideas in the literature as information. I became more convinced during this process that the range of concepts located within the field of educational assessment are limited in their capacity to actually describe the field of assessment—they seemed to be stuck within it.