# ISSUES AND CHALLENGES IN THE ASSESSMENT OF POPULAR MUSIC PERFORMANCE WITHIN THE UK HE MUSIC SECTOR

**ROSARIO MAWBY**

**MPhil** **2020**

# ISSUES AND CHALLENGES IN THE ASSESSMENT OF POPULAR MUSIC PERFORMANCE WITHIN THE UK HE MUSIC SECTOR

**ROSARIO MAWBY**

School of Arts and Media

Salford Music Research Centre

University of Salford

Salford, UK

Submitted in Partial Fulfilment of the Requirements of the Master's in Philosophy (MPhil), January 2020

# Table of Contents

# List of Tables and Illustrations

# Acknowledgements

I would like to express my deepest gratitude to Dr. Robin Dewhurst for his knowledge and expertise, for agreeing to supervise me when I sought his supervision, and for being the most patient and understanding supervisor any postgraduate student could ask for. The writing of this thesis alongside my career as an academic, lecturer, and freelance musician has been so far the hardest project I have undertaken, and it has not been the easiest journey to complete. Thank you.

I would like to thank my colleague, mentor, and best friend David Bebbington for his continued support and encouragement: I would gladly have given up at several points of my Masters but he believed in me when I could not.

I gratefully acknowledge the funding received from the EWS Educational Trust Fund to undertake my MPhil.

Many thanks to Professor Louise Jackson for the much-sought advice in the final days of submission and the frantic messaging.

To my dearest Tita and Uncle Roy, both wonderful academics, for their continued encouragement in my educational journey and believing in me.

My friends Fiona and Rosie, I would like to thank both of you for the endless study days over the final year of writing. I am genuinely thankful for the constant coffee and cake which has been very much needed.

Finally, I am indebted to my wonderful parents. Without them I would not have considered my second Masters had I not had the continued support and unconditional love that they show me each and everyday. I am proud to have completed my MPhil and I am proud to be your daughter.

## Abstract

This thesis summarises that assessing popular music performance has it issues and challenges. These arise from: the interpretation of language, how we assess in group assessments versus solo performances, creating assessments within the degree framework whilst keeping it industry relevant, the differences in our expectations of what professional standards are, and the lack of training for part-time (and some full-time) members of staff.

It identifies there is a definite lack of research within the field of popular music and the necessity for further research required as well as suggestions for what these areas are. The scope of future research ranges from a new template for an assessment grid that could be used for all popular music performance assessments that cross-reference with the institutions own learning outcomes, to establish what is best practice and why, understanding the student body and their concerns on a national level, as well as wider issues such as the benefits of the passing and failing at undergraduate degree, creating agreed guidelines of what professional standards are, and the best ways in which to train academic lecturers – especially those with a freelance/ portfolio career.

The method of this thesis uses desktop research and analyses the research that surround the areas of contention. It applies principles of other subjects to when there is a lack of research in the area of music. It is important to note that the thesis does not look to provide answers but to open up the dialogue and discussion that is anecdotally spoken about but is not always articulated with findings and research. Further research and evidence is required to see how these scenarios could be improved and resolved.

# INTRODUCTION

## Contextual Background

This thesis looks to address the issues and challenges in the assessment of popular music performance in the UK HE (Higher Education) music sector. The study is based on exploring the key literature rather than running interventions. It has three main concerns: 1. The differentiation of grading criteria application in performance assessments; 2. How this plays out across different institutions within the UK HE sector, and; 3. The role of grade expectation in both student and tutor perceptions of grading. The purpose of this study is to delve into the disparity of the perceived views of marking and to understand why there is an apparent lack of parity between institutions and the level of expectation from both student and tutor. This will be broken down into the following chapters which highlight areas linked to assessment in popular music performance and then investigating relevant wider issues surrounding assessment: 1. Learning Outcomes Linked to Assessment; 2. Group and Solo Assessment; 3. The Interpretation of Language in Assessment; 4. Assessment, Feedback, Teaching and Learning; 5. Professional Standards, and; 6.Training (Gaps in Practice).

I would like to discuss the context of this study and its importance to me as an academic, educator, practitioner, and student. My background as a student in popular music performance in Higher Education has directly inspired the investigation of the issues and challenges in the assessment of popular music performance. This is due to my personal experiences as an undergraduate and postgraduate degree student (University and Conservatoire respectively) which left me confused about the process of how performance was assessed. I was an eager

student that wanted to understand the grading criteria presented to me and to achieve the best mark I could. I was fortunate in my postgraduate learning that my assessment band comprised of tutors from my previous institution. After my performance, an ex-tutor asked me what mark I had received. Upon telling them my result they were surprised by the grade I had achieved. I queried the surprise as I had never questioned academic judgment before (I had also thought my mark seemed rather generous) and through my line of enquiry the response was that had they graded me at my previous institution I would have likely received an entire grade boundary below what I had gained. According to Groothuijsen, Bronkhorst, Prins, and Kuiper (2019) “Practice-oriented educational research is increasingly gaining traction in education research due to its intention to contribute to both educational research and educational practice”. The paper also concludes that “Taking… common quality concerns as a starting point, close collaboration between researchers and teachers could decrease researchers’ challenges concerning legitimacy and relevance of their work and increase teachers’ use of research in educational practice.” Subsequently, my career in academia has also challenged my perceptions of assessment and through practice I have observed the differing ways in which we assess. Due to the personal nature of this thesis some of the observations made as both tutor and student will appear as part of my research in the study. Wesolowski (2012) states that

> For music educators to be successful at music performance assessment, they must be not only prescriptive in evaluating their individual students and ensembles but also willing to assess and improve their methods of teaching and communicating.

Therefore, where situations and circumstances have been anecdotal, I have tried to document this as unpublished work to highlight the issues of my experiences where research has not been able to evidence these particular cases. The thesis is limited in scope to popular music performance in Higher Education in the United Kingdom (predominantly England) as opening up the research to Further Education or even Music Education in High School would make the thesis broad and too wide in scope. Methods alone cannot solely be addressed to improve assessment, every part of the assessment cycle – and any area that impacts upon these specific sections – needs to be considered as part of the design of learning but the focus of this thesis will not look at teaching and learning in the broader sense, but this will need to be addressed if it impacts upon assessment specifically.

## Literature Review

Before reviewing the literature, the definition of Higher Education (HE) must first be addressed. Traditional ideologies present that University was for reading broadly around a subject and choosing a specialism after study. The meaning, and indeed the importance, of HE has since changed (Bryan and Clegg, 2006). University is now more conventionally known as an institution that educates students for their desired career path. Whilst I do not agree with this statement, it is prudent to face the harsh reality in which we work in and to acknowledge that the nature of HE is becoming a transaction between student and institute is increasing in culture in correlation with the increase of tuition fees. Professional statutory and regulatory bodies, the public sector (employers and parents), and the government assert pressures on Higher Education Institutions (HEIs) to perform well. These pressures make it increasingly more important to include transferable skills within our teaching, and carefully to

consider how we use the learning outcomes to demonstrate the skills that are being learned.

There is a vast body of work specifically focusing on assessment in HE. Assessment in HE has received increasing attention in the United Kingdom (UK) over the past few decades in an attempt to justify why we assess and if we assess correctly within the frameworks that exist. The Department for Business Innovation and Skills in their most recent green paper stated that “more providers entered the sector in the last five years than at any time since the last major expansion in 1992” (*Fulfilling our Potential: Teaching Excellence, Social Mobility and Student Choice*, 2015). The main focus of the general wider reading (non-subject specific material) will look at the research specifically after this year (1992) to the present day, unless individual theories need further investigation, or have great ongoing impact on the sector. Research has demonstrated thus far that the literature regarding popular music assessment within HE is minimal in scope. In the *Journal of Popular Music Education* manifesto (2017) they acknowledge that popular music education has roots stretching as far back as the 1930s through curriculum in community college’s in the United States of America. They do however cite that Lucy Green argues that popular music education “came of age” in 2017. Popular Music is therefore considered a relatively new subject within HE in comparison to other areas of study such as English, History, Law, Mathematics and so forth. This could explain the lack of literature at this stage, yet the number of institutions has increased significantly in the last forty to fifty years. Research has generally focused on compulsory education contexts (like Lucy Green’s earlier work) but this has not translated into popular music education in higher education. However, there is more research surrounding music assessment, especially concerning level one,

two and three (all levels refer to the qualifications framework outlined at https://www.gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels). Though not entirely relevant to the study, elements of the research found demonstrates the need for further research in the area of music in general.

The initial areas of investigation are listed below. These were chosen to better my understanding of assessment (and the surrounding context) and where the original hypothesis of there being any potential issues might be:

- Assessment for learning, assessment of learning, assessment as learning
- Formative assessment
- Learning outcomes in relation to teaching and learning
- Relationship between assessment and feedback
- Professional standards and threats to professional standards
- Lecturer's or tutor's ability to assess
- Testing product not process

Approaches to Student Learning

If learning outcomes are central to the assessment process how do students learn? There are many paradigms within learning typologies so the few listed below are the most relevant to the area of study. Four types of learning styles are identified by Entwistle (1993): deep, surface, strategic, and apathetic. The deep approach is characterised by the student who wants to learn and will utilise and apply what they have learned to other areas, including assessment. The student who is a deep learner is described by Entwistle as the "model student". Deep approach learning is

associated with the term deeper learning, a method that tutors aspire to deliver within their course content according to Entwistle. The surface approach is a student who fears failure, so applies the technique of memorising the information they are given. Whilst memorisation is a useful skill it does not embed learning. The results of memorisation, for some forms of assessment, can be successful. Examples of this would be where a music student would memorise previous answers from a mock paper and emulate these in their next exam. However, if information is not retained or not memorised by the student then failure will ensue. The strategic approach is a student who wants to succeed but does so by finding out what 'makes good grades'. Essentially the student aims to 'work the system'. The strategic learner, much like the surface learner, is unable to apply their learning elsewhere as their main focus is attaining the best mark possible. The apathetic approach is a learner who 'drifts' into higher education. They do not engage with their education, but they also have a personality type that indicates a general lack of interest.

Another concept related to the approaches of student learning is discussed by Marton and Säljö (1997). Their understanding of the literature, at that point in time, stressed that assessment could in fact affect the way a student habitually learns. The student can interpret assessment tasks in one of two ways. The first type of student is driven by assessment tasks and produces work they believe is required (encouraging surface approach learning). The second type of student adapts their study to meet the assessment requirements (encouraging deep approach learning). Miller and Parlett (1974) divided the students' driven by assessment into three different psychological profiles: "*cue-conscious*", "*cue-seeking*", and "*cue-deaf*" a paradigm where students need clarification. Cue-conscious refers to the student who responds well to 'cues' delivered by staff. These cues to the cue-conscious

learner might involve trying to impress and being noticed for it, taking tips from staff regarding assessment areas, and/or observing a lecturers' preferences. Evidently, though Miller and Parlett discuss these psychological profiles in 1974 the concept is still discussed in more recent years in 2013 where Aerts outlines how cue-seekers would actively question staff about assessment and/or discover the individuals' interests. Whereas the cue-deaf student worked regardless of assessment and chooses to rely purely "on their own abilities" (Aerts, 2013). This theory of learning suggests that students, dependent on the learner type, demonstrate a differing approach to assessment so how do we address these differences, and does it alter the way in which we assess?

Heywood (2000) considers assessment in higher education as a type of judgment. He stresses the importance behind the word assessment, and that the term is not interchangeable with words such as 'exam', 'tests', and 'grades'. Rowntree (1987) describes

> assessment can be thought of as occurring whenever one person, in some kind of interaction, direct or indirect, with another, is conscious of obtaining and interpreting information about the knowledge and understanding, or abilities and attitudes of that other person. To some extent or other it is an attempt to know that person.

The definitions of assessment outlined here provide an overview of the key idea of assessment as a form of communication and passing judgement, but who are we assessing for? Brown and Knight (1994) compiled a list of the following: "the student, other students, tutors, mentors, employers, university

management, financing and other government bodies, funding councils". Although the process involves a two-way relationship, in this case, between student and tutor, assessment impacts more than the immediate parties involved.

## Assessment in Music: Current Debates

Music education and assessment is an active area of discussion in general. Wesolowski (2012) states that "assessment has become one of the most important and pervasive topics in music education." Whilst he refers to the United States, the same sentiments can be applied to music education in the UK. However, the literature regarding popular music assessment in higher education is seriously lacking. With the exception of a Norwegian study (Nørsett, 2015), that looks specifically at popular music performance in HE; a European commissioned study (Cox, 2010); and The Association Européenne des Conservatoires, Académies de Musique et Musikhochschulen (AEC, 2017), the research discusses music in general and music education at secondary and further education level.

Recent music articles agree with the literature that exists around assessment. Scott (2012) discusses the importance of learning outcomes in relation to teaching and learning in music. Whilst the context differs from a higher education setting the role is not too dissimilar from those working in one to one situation with students in HE. She acknowledges the responsibility of music teachers (teachers here refer to compulsory education and peripatetic instrumental and vocal teachers) as being able to provide reliable and valid information for their students' performances. She also argues that assessment should encourage students to reflect upon their work as a way to understanding music from a holistic sense, as well as improving technical skill within their performance. Assessment as

learning is evident here. As mentioned previously, the AEC released their grid of learning outcomes in relation to Higher Music Education (HME) in Europe. They state that the AEC learning outcomes are designed to:

- assist institutions in implementing the requirements of the Bologna Process reforms and, more specifically, in (re-)designing curricula and adopting a student- and competence-oriented approach to curriculum design;

- facilitate the international recognition of students' studies and qualifications, and increase compatibility and transparency within and beyond the HME sector;

- provide current or potential students, employers and other stakeholders with a clear presentation of the main aspects of a HME curriculum and its opportunities;

- • serve as reference point for institutions and relevant stakeholders within quality assurance and accreditation processes in HME;

- help employers and other stakeholders to understand the competences of musicians they hire.

The relationship between learning outcomes, within the HE sector, and beyond the educational sector becomes increasingly more important.

There are several supporting strategies available to classroom teachers in the format of textbooks. Nevertheless, and similar to the literature that already exists, the content focuses on objective testing where there is a clear and definitive answer as opposed to subjective testing where the answer is defined by the judgement of the academic. The nature of music is diverse and usually more than one learning outcome is assessed that is potentially subjective. Music teachers, in general, use assessment on a daily basis. Teachers must be able to understand, and correct, the issues that students face in performance. They would need procedures in place that would allow them to diagnose the problem, as well as having the technical skill to fix this (Wesolowski, 2012). Is it because of our daily ability to assess music and students' practice, that new lecturers feel they already know how to assess? In reality new lecturers feel completely underprepared (Gibbs and Coffey, 2004). We assume that we are masters/ experts of our field, yet we are all at varying levels of professional standing within the music industry. How can we measure the tutor's ability to assess, or how do we train and develop our staff, so that assessing is level and fair?

Traditionally the methods in the assessment of music tests product, not process. For example, students who take examinations towards their 'grades' do so to receive high scores, before starting the next grade (Scott, 2012). First, the disadvantage of this is the students become good at taking tests. It also encourages surface and strategic approaches to learning. Secondly, achievement starts to associate itself with anxiety, stress, and nervousness, and the child, therefore, is unable to play to the best of their ability (Scott, 2012). The learning prior to higher education can be seen to be detrimental if we view it from the position of strategic and surface learning.

Scott (2012) highlights that “assessment for learning and assessment of learning are not mutually exclusive” and goes on to explain that the information provided from one assessment can be utilised by both forms. The teacher uses the assessment to grade the student, yet the student uses the assessment feedback for guiding future practice. Scott refers to secondary level assessment, but the principles can be applied to higher education too. It is an accurate description of how we assess at undergraduate level. Scott does, however, advise that students should still have the opportunity to be assessed formatively.

Arguments around assessment suggest that the attention of assessment should not be brought to the forefront of student’s minds. The challenge here is that in higher education most institutions have some form of assessment period that variously allocates time off to continue finishing work for their assessments without disruption to taught study. Assessment should be at the forefront of higher education if degree certification is the accumulation of learning.

The concept of assessment as learning Scott (2012) applies to music education. The student follows three steps: monitoring their own learning, reflection upon their accomplishments, and using this information to inform future learning to better themselves in order to perform at a higher level. It is evident that some of the general theories about assessment can be applied to music performance. Whilst the literature does not state this, it is worth investigating and researching further. Formative assessment is a useful tool for assessment for learning as “the assessment can be carried out by the teacher, learner, or her peers” (Wiliam, 2010). Mock assessments (formative in nature) allow the students to understand the concept of what the assessment may be like, and the ongoing verbal feedback aids

the student's in their approach to preparing for the final practical/performance. This type of assessment is commonplace in the music sector at HE level as "assessment giving diagnostic insights are likely to lead to better decisions about teaching than those that simply monitor student achievement" (Wiliam, 2010). However, further investigation needs to be sought to see whether or not this impacts the understanding to attaining better marks, and specifically the institutions that do not provide the mock assessment route. The learning function between formative and mock assessment can be different. For example, the formative evaluates where the student is at currently within their learning and can be presented in an informal setting. A mock assessment replicates the actual assessment giving the student experience of environment and setting.

Studies in music performance assessment have looked at the concept of using rubric construction in assessment (DeLuca & Bolden, 2014). This study does not relate to popular music or higher education specifically. The study looked at whether or not mathematical equations could be applied to a nonmathematical subject. Research suggests that rubrics can be used in music performance assessment to assess both the technical aspect of performance and the aesthetic creative production (more subjective than technique). They argue that rubrics allows increased score reliability "by providing consistent grading criteria that are standards-based." The study suggests the problems generic rubrics cause is using subjective descriptions to distinguish between levels. This area of study coincides with Biggs' *qualitative approaches* theory. The language used in rubrics becomes a challenge, as language is a barrier in more than one area of assessment. From the study, they produced three types of performance rubrics: discrete-component rubric, integrated-component rubric, and self-reflection rubric.

Discrete-component Rubric

| Criteria | Performance Component | Level 1 | Level 2 | Level 3 |
|---|---|---|---|---|
| **Descriptive criteria** | Dynamics (i.e., volume and shape) | • Notes are at the same volume | • Notes have varying volumes and melodic lines have overarching shape | • Melodic lines are shaped to work together to provide overarching shape across lines of music |
| **Less-descriptive criteria** | Dynamics (i.e., volume and shape) | • Little dynamics used | • Some dynamics used | • Full dynamics used |

Integrated-component Rubric

| Performance Components | Level 1 | Level 2 | Level 3 |
|---|---|---|---|
| **Breath Dynamics**<br>**Tone quality** | • Inadequate breath support leads to nonpurposeful dynamics in melodic lines<br>• Tone quality suffers throughout melodic line due to lack of breath support | • Breath supports purposeful dynamics (e.g., varying volumes and overarching shapes) throughout melodic lines<br>• Breath supports tone | • Breath consistently supports purposeful dynamics (e.g. varying volumes and overarching shapes) within and across multiple melodic lines<br>• Breath supports varied tone |

Self-Reflection Rubric

| Performance Category | Level 1 | Level 2 | Level 3 |
|---|---|---|---|
| Self-reflection | • Student accurately applies performance criteria to others' performances but does not use performance criteria to self- assess personal performance | • Student engages in teacher- structured self-reflection during practice sessions that leads to changes in performance | • Student engages in self-directed reflection during practice sessions that leads to changes in performance (i.e., progress in practice through self- assessment and planning for learning) |

(DeLuca & Bolden, 2014)

These descriptors are more similar to the General Certificate of Secondary Education (GCSE) and A-Level descriptors in education than higher education.

The only formal study available is the *Assessment in vocal popular music performance in higher music education* (Nørsett, 2015). Private correspondence with Nørsett (2017) indicated that the study was in its preliminary stages and that the information on the poster presentation was her current findings so far. Nørsett states that "several studies have been conducted on a variety of assessment criteria, but little is known on the *assessment in vocal popular music performance* in higher education" (2015). It is important to add here that little is known about assessment in popular music performance, let alone focussing on one specialist area.

The study predominantly looks at the vocabulary used within the assessment process and how students interpret assessment criteria. The method took twenty-six students aged between nineteen and twenty-six from two Higher Education Institutions (HEIs) in Kristiansand, Southern Norway. She asked one question, one

month before the performative assessment, “which assessment criteria do you believe you are assessed by in your exam performance in vocal popular music?” Four assessors were asked “which assessment criteria do you consider in an exam performance in vocal popular music?” at the same time as the students. Answers were required as keywords as opposed to long sentences. Participants were told, as a starting point, that there were four criterions of assessment: technique, interpretation, expression, and communication. The results were then thematically grouped to see if one area was larger in number. Students were split according to the institution that they studied at – to observe any differences. Group A was a group of twenty-one, whilst the five that remained became Group B.
Inadvertently, Nørsett was looking at comparability across institutions in the students' understandings. Interestingly, the students and assessors rated communication as the most important factor of performance assessment. There was little difference between Group A to Group B, other than Group A focussed more on technique and expression, than Group B, who focussed on confidence. She noted that members of Group B were more detailed in their descriptions. The biggest difference was the students answers to that of the assessors, and even the terminology was dramatically different in approach. The words symbolise their understanding of the assessment criteria, more so than the students’ interpretation. The issue is not only applicable to Norway but to UK popular music higher education.

The literature so far highlights a few concerns that arise directly from assessment: how can we use learning outcomes to enhance assessment and not purely in reference to teaching and learning, are staff supported in terms of training to assess, how do we measure professional standards, does the concept of the mock exam aid or inhibit, and is examining product over process enough?

Main Gaps in Assessment

In order to look at the issues and challenges that arise in popular music performance assessment, it is important to understand where the gaps are in the research based on the literature review above. The literature directly related to popular music performance assessment is limited. This part of the introduction will aim to highlight the main gaps that are in assessment, which will inform how I then attempt to break down the study.

Carnegie Mellon University (2020) outlines that:

> University Policies protect the integrity of the university's mission, reputation and operations, and support the management of major institutional risks. They inform our community and third parties about the major risks that the university seeks to address and communicate risk tolerance in key areas. They also promote compliance with laws and regulations, especially those of the federal government.

Presented next are two university policy documents. The purpose of looking at these documents is to view what the Universities are looking to achieve and indeed what they off to the public sector. The University of Salford (UoS) defines assessment as

> the generic term used in this document to cover all forms of formative and summative assessed activity, for example, coursework assignment, presentation, test, portfolio, written examination, unless where otherwise specified. Feedback, as used in this document, refers to all information provided to students that both certifies their level of learning and allows them to improve.
>
> (2016)

The University of Liverpool respectively offers the following definition of assessment

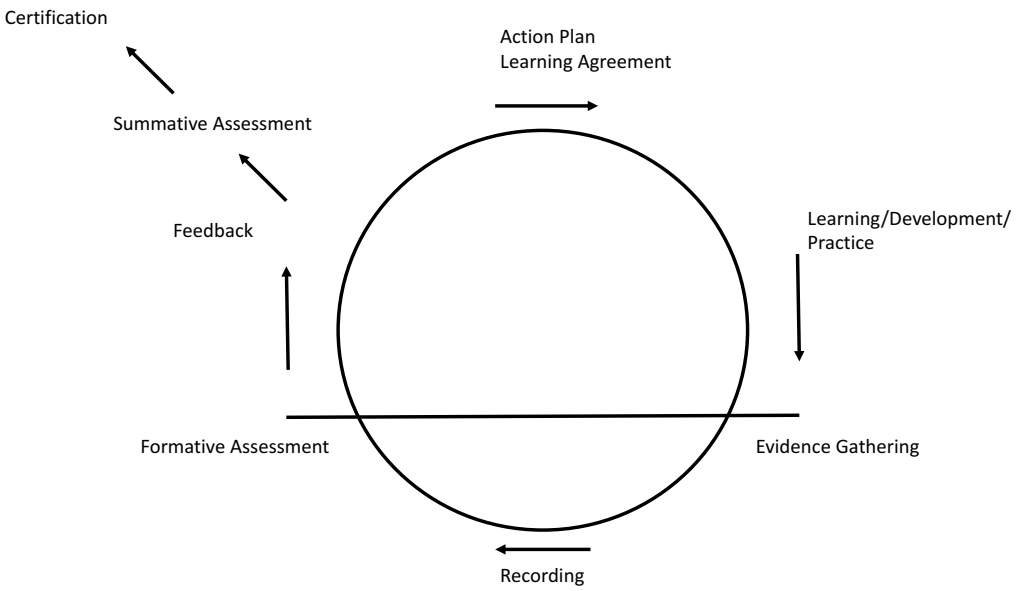
- promoting student learning by providing the student with feedback, normally to help improve his/her performance;
- evaluating student knowledge, understanding, abilities or skills;
- providing a mark or grade that enables a student's performance to be established. The mark or grade may also be used to make progress decisions;
- enabling the public (including employers), and higher education providers, to know that an individual has attained an appropriate level of achievement that reflects the academic standards set by the awarding institution and that agree with UK norms, including the frameworks for higher education qualifications. This may include demonstrating fitness to practise or meeting other professional requirements.

(2016)

Both policies demonstrate that assessment is related to student learning and that there is a strong emphasis on the relationship between assessment and feedback. Salford focus on the form in which assessment can take, whilst Liverpool look at how marking and grading is relevant to the public. Both policies also use the Quality Assurance Agency (QAA) (normal practice for institutions with degree awarding powers) as a foundation to write their assessment policies on validating the agency and its work. Neither discusses the role of the assessor within its definitions.

Moving on from definitions and onto process, the process of assessment Race (1995) suggests that assessment should not be viewed "as a

cycle of consecutive steps" but looked at from an overlapping perspective. The overlapping concept is similar to the Venn diagram (see Chapter 5). The diagram demonstrates that all three factors are equal in size and importance.
Where the ovals overlap, the overlapping areas show how these
three elements impact upon one another. He also suggests that to improve assessment one should imagine the worst-case scenario in order to 'make something better'. Rhodes and Tallantyre (1999) challenge Races' view on the assessment process. They demonstrate in figure 9.1 a diagram that looks at the cycle of assessment.




(Figure 9.1, 1999)

In the cycle, they present four steps: the action plan learning agreement, learning/development/practice, recording, and feedback. The cycle demonstrates how summative and formative assessment is not positioned within the cycle. Summative assessment and certification spur from the feedback section,

whilst formative assessment and evidence gathering is placed near the recording element of the cycle. However, the process is similar for all types of assessment, it could be argued that trying to align summative and formative assessment creates dilemmas (Hounsell & Murray, 1992). Summative assessment is “assessment-for-grading”, yet the outcomes of formative assessment is “assessment-for-learning”. The UoS assessment policy separates and distinguishes between summative and formative assessment. It states that “summative assessment contributes in a defined way to a module mark or grade… whereas formative assessment is an entirely developmental feature not connected to the calculation of module marks or grades”. The concepts overall argue that all assessment is for learning. It is at times difficult for the student to distinguish that summative is not wholly dedicated to the grading system especially when “students are increasing demonstrating customer-like behaviour and are demanding more ‘value’ from institutions” (Woodall, Hiller & Resnick 2014) reiterating the perception that University is no longer about degree certification but what skills it can offer the student beyond the life of education. Yet, what do the students perceive to be good ‘value’? The study concludes that the ‘student-as-consumer’ reference should be reconsidered. It determined that home versus students perceived value to be different. They used a specific framework for reference that highlighted these concepts:

1. Attributes only – product/service features that consumers find to be of benefit, or value

2. Outcomes only – benefits, or value, that consumers derive from their association with an offering

3. Value for money - a readily rationalised balance of benefits and sacrifices, usually based on price and attributes (plus the more obvious outcomes).

4. Net Value - a complex, intuitively balanced combination of all benefits (outcomes and/or attributes) and all sacrifices (monetary and/or non-monetary) perceived to be associated with a particular offering.

5. The cheapest option - a bargain, usually focused on minimum possible sacrifice.
(Woodall, 2003)

Home students (UK citizens) found their value within “price and attributes” and international students found value within “results for the customer (outcomes)”. Knowing this, is this what the students expect, or acknowledge as value, equivalent to what the institutions believe they are providing for them? Early on in the study there are two observations made which are “Higher education is a highly complex service, offering an intense, emergent unstructured, interactional and uncertain environment” and that “Students will inevitably experience both highs and lows, and for universities, of course, satisfaction has now assumed substantial importance, not only in the United Kingdom where the national student survey… puts a premium on satisfaction”.

Brown and Knight (1994) highlight areas that could be improved to speed up the assessment. They suggest “the amount of assessed work they require, its form, when it is required, submission deadlines, feedback to learners, the use of paid markers, [and] assessing with the tutors” are measures that require thought and deeper understanding, but no solutions or suggestions are given here. Problems with summative assessment are identified by Falchikov (2004) and Pelligrino, Chudowsky, and Glaser (2001). Falchikov points to the emphasis of examinations, issues with reliability, and bias in marking. She also addresses the issues surrounding student learning and states that summative assessment encourages surface or strategic learning, contributes stress to the student, or does not promote student motivation. Pelligrino, Chudowsky, and Glaser raise the valid point that the assessment of student learning and understanding is limited. The research suggests that we are no further ahead with resolving these problems. They also discuss that summative assessment measures the students' achievement during the assessment period, but they also raise concerns with aligning assessment. The text here is outdated as aligning assessment is gradually becoming more streamlined in approach, but this is not the case across the country – evidenced by the many external regulatory measures that are in place for institutions to meet.

Brown (1999a) comments on assessment practice. She draws up a comparison of the practical skill ‘hedge-laying’. She argues that setting an essay to demonstrate that the student has the theoretical knowledge to present their skills is inappropriate and that we need to assess the actual skill of ‘hedge-laying’. Brown demonstrates the importance of setting a relevant task at assessment in order to assess the desired learning outcome and the best way in which this can

be achieved. Could this theoretical justification support how and why we assess in popular music performance assessment? It appears that the way in which we assess in popular music performance is taken for granted, and it is not challenged based on the assumption that this must be the way in which we assess, but a lot of which is impossible to replicate with limiting control of external factors.

At a conference held in Hereford College of Arts *Music, Mental Health and Wellbeing* (2019) which I hosted we briefly discussed the literature that existed in popular music performance/ performers and unsurprisingly (and surprisingly) the lack of literature in the area. These chapters in this thesis look to highlight the issues and challenges that arise from assessing popular music performance and to begin to articulate and document the informal discussions that surround the topic.

Why Do We Assess?

One of the advantages of assessing is that the assessment phase allows us to understand several points in the student's educational journey, as well as determine where the teaching can develop. Assessment allows us to monitor students' progress, evaluate their strengths and weaknesses, as well as understanding their understanding of the programme undertaken; however, is this something that can come from a purely formative assessment? Should we potentially adopt to not assess and take a formative assessment and turn it into the summative grade? Does this increase staff workload, or does it reduce it? Does it benefit the student to not be assessed at all? Should we even grade at all at bachelor's level? Could we adopt a pass/fail mark similar to the model of some of the postgraduate degrees adopt (although interestingly postgraduate degrees are opting for the pass/merit/distinction and fail route), simply evidencing that students are able to study at this level? The

value of keeping summative assessment though is that within the music industry it is important to replicate scenarios of a pressured environment. An ongoing emphasis is placed on the importance of preparing our students for the ‘real world’ and one way to achieve this is to replicate this at the assessment point. There is also a level of preparation that this gives the student for what is yet to come. It also addresses other regulatory bodies who specify the need to demonstrate transferable life skills too.

## CHAPTER ONE – LEARNING OUTCOMES LINKED TO ASSESSMENT

The importance of learning outcomes in Higher Education Institutions (HEIs) has been a topic of focus in the last few decades. Aligning learning outcomes against assessment criteria (see Learning Outcomes below) had been an issue of contention for some time. The argument has developed into how much do we assess process, or do we even assess process at all? In popular music performance the majority of process is key to the outcome that is delivered at assessment – this is an area that is worth exploration. Another issues that arise from aligning assessment criteria with learning outcomes across the various institutions is the potential lack of parity of what we are assessing across the sector. An obvious example, performance in music is quite broad. Music educators instinctively know what a performance should look like, for example Lortie (1975) discusses the concept of "apprenticeship of observation" where educators use their own student experience as a basis for their approaches to teaching. There is an expectation of the quality we expect to see and hear, but the information we assess by is inferred from our previous experiences of performance as assessors. Another issue that causes contention is the degree to which we assess technique in comparison to musical style, and the approach we take to assessing subject knowledge in relation to performance when the genres that are presented differ so greatly. These variables affect assessment drastically if institutions differ in their emphasis of one particular level of skill area against another.

### Learning Outcomes

Since the late nineties there has been an increase of literature regarding learning outcomes (Brown & Knight, 1994; Heywood, 2000; Bloxham & Boyd, 2007) but the

key concept in HE concerning learning outcomes is the notion of “constructive alignment” (Biggs, 1996). The alignment model advises course validation, quality enhancement, and staff development. Tutors need to be able to articulate what they want their students to learn. From these learning outcomes the assessment should then be designed to facilitate the students’ learning by creating tasks that demonstrate “the appropriate level” (Biggs, 1996) of learning. Atkins (1995) argued that what we should be assessing is the learning of the students. Atkins also argued that *what* students learn is important too eluding that process is equally as important. Similarly, Bloxham and Boyd (2007) state “assessment practice is judged primarily on whether it effectively measures the intended outcomes of course study in a valid, reliable and transparent way”. Performance tasks in popular music are easily simulated and are replicable in an educational environment so constructive alignment demonstrates value here. The literature indicates that learning outcomes are integral to the assessment process; however, contrary to this, Boud (1995) offered an alternative view that assessment, at the time, undermined learning and, as previously stated, performance relies on the effectiveness of process which is where the student learns the most value and yet we assess the finished product which could have been in-keeping with other educational thought at the time. Performers who suffer from ‘red light syndrome’ (where musicians fear subconsciously when the red light from recording comes on and creates unnecessary pressure) for example do not necessarily perform at their best when they are in a pressured environment which is what assessment can be replicated to be. Joughin (2009) summarises that “*assessment as judgement* therefore seems to be at the core of assessment, and its immediate object is a student’s *work*.” Whilst Joughin does not refer directly to the arts, the definition that “*assessment as*

*judgement"* has added value in the area of performance. We, as examiners, are judging what we see and hear. Although both argue a valid point, fundamentally learning outcomes and constructive alignment, where "the learner constructs his or her own learning through relevant learning activities" (HEA, n.d.), are well-established within HE education in the UK.

One of the contentions that exists between student and staff is how do we assess performance with objective outcomes when the discipline is subjective by nature. Informal conversations with students have eluded to me that they struggle to understand the concept that lecturers can apply an objective approach, so how do we measure performance with valid methods and communicate this to the student body and challenge these preconceived ideas? In the Quality Assurance Agency subject benchmark statements (2016) – designed to help graduates understand what they should expect to know when leaving their studies as a guideline –, practical skills at threshold level for graduating in Music at BA level students should be able to demonstrate the ability to:

- "convey personal expression and imagination in practical music-making through employing appropriate technical and interpretative means
- recognise and use essential components of a musical language (intervals, rhythms, modes, metres, timbre, texture, instrumentation).
- recognise and respond to aspects of musical organisation, whether aurally or by studying a written score
- collaborate in music-making, whether through ensemble performance, co-creation, improvisatory work, or analogous activities".

Potentially one of the ways to resolve parity and keep consistency in the way in which we assess – without losing individuality – is the reflection of these statements within the wording of the learning outcomes in assessment. The wording could be standardised to form the basis of every practical/ performative module across the country but in relation to the individual learning outcomes this could differ in direction yet still employ the same foundations. To communicate this to the student would mean being as open and as transparent as we possibly can but also helping them to understand the terminology: it is the language rather than the concept that can confuse the student. Best practice across the country looks to break down these barriers and students respond well to this. *A Compendium of Assessment Techniques in Higher Education: From Students' Perspectives* (2018) has been created and put together by three different students' perspectives. They look at the different modes of assessment (not specific to music). The assessment is given a summary of what to expect followed by the student's positive and negative aspects. From this study the barriers are being broken down by the student's in order to further deepen our understanding of student perception. As part of the performance modules I have taught on, there have always been briefing sessions where we talk through the grade criteria and the seminar has asked the students to grade either professional performances from various YouTube videos or to grade their peers in workshop/ masterclass environments, or to grade previous videos of ex-student performances. I believe this to be good practice and a study by Hanrahan and Isaacs (2010) would suggest that students are in support of this too, but from the evidence coming out of the aforementioned conference break-out discussions it is clear that this is not being utilised in all performance degrees. The timing of these briefing sessions is also crucial to the development of the students understanding. Through

observations at various institutions these sessions can be delivered after the first assessed performance as part of the feedback process whilst other programmes provide these at level 5 instead of level 4, and some places deliver these sessions prior to the students' first assessment. In my experience the best time to discuss grade criteria with students is in level 4 prior to the first assessment. Steps to clarify the assessment process as early as possible in the programme may be useful in reducing the extent to which the way we assess is misinterpreted.

Reliability Versus Validity

What do we mean by reliability and validity? Elton and Johnston (2002) define reliability within assessment as being able to generate "comparable marks across time, across markers and across methods." In essence, reliability looks to achieve repeatable tests. No performance is alike in music, so whilst the outcome cannot be repeatable (what two performances are ever alike?), the environment and formatting of the assessment can be made repeatable. We can control certain aspects of performance, but these measures vary greatly throughout the sector. Some HEI's allow for the students to organise their own venue for the purposes of their performance assessments. The venue/ environment can be seen as a repeatable measure however having flexibility can aid the student. It would be inappropriate to perform a piano recital in a public house, yet it would be fairly standard to watch a metal band perform in the same venue. Using the same assessors could work as a repeatable measure however each assessor has different areas of specialisms. Do we ensure then that the person second marking is the same throughout all assessments? Academics would argue that this is the responsibility of the external

examiner but parity within a course is vital and staff should have a level of understanding that allows them to carry this out.

Validity is defined by psychologists as a test that measures what it intends to measure (Heywood, 2000). Theoretically then, the design of learning outcomes and constructive alignment should work in tandem with validity. Brown and Knight (1994) discuss validity with a developed view. They offer the opinion that validity is not only achieved by measuring the degree to which learning outcomes have been met but also by measuring the "intended effects". An alternative view expressed by Barnett (2007) looks at the "two different journeys which are meant to work simultaneously together". He describes these journeys as being centred around authenticity versus standards. Whilst this is an accurate description, standards could be confused with the concept of standardisation, therefore, all references made from here on in will be referred to as reliability and validity only. With these theories in mind, what do we believe to be the "intended effects" of performance? There is no research to suggest that there is a definitive answer, which can be applicable for all other forms of assessment (coursework, essay etc.). It would be worth investigating further whether there is a way to create comparability between different programmes without standardising learning outcomes.

According to research, reliability and validity cannot coexist together without a 'trade-off' (Elton & Johnston, 2002). Elton and Johnston's research suggests that curriculum in the United States sways more towards the reliability approach – concentrating on certification, therefore prioritising grades and marking. In the UK the research indicates that the higher education system focuses on validity, implicating that our grading is not as accurate as it could be. Why should this be the case? Music, although subjective, still has quantifiable criterion such as technique,

style and communication, for example. All of these aspects can be measured and set against specific standards if necessary. With this knowledge in mind, it is not unreasonable to assert here that defining “intended outcomes” is valuable in the assessment of performance from assessing using the validity method.

One of the reasons behind the UK preferring validity over reliability is the move towards criterion-referenced assessments, where the assessment criteria are prioritised above all other aspects of the process (Johnson & Blinkhorn, 1992). They specify clearly that the criteria must be valid else “the whole exercise is pointless”. This method is utilised in the assessment of music where the basis of assessment lies within criterion-referenced assessments. It would therefore be worth investigating here whether or not reliability can be redefined within the context of criterion-referenced assessments. It would be worth exploring how these concepts could co-exist in a framework that is negotiated and discussed in the current academic climate. Part of the challenge here is to develop a strategy that supports students who are struggling to understand what they are meant to be assessed on? Contributing factors and circumstances could include the following: an assessment brief which is purposefully vague, including the use of ambiguous terminology, and the lack of clarity as to why they are being assessed. These are common complaints amongst the student body and across varying institutions. One could suggest that the apparent lack of understanding on the part of the student is compounded perhaps by the wider issues of the research in popular music education.

The biggest concern regarding reliability and validity is the potential bias in marking students’ work. A general overview of potential bias looks at the difference between ‘hard’ and ‘soft’ subjects. Hard subjects represent areas such as

mathematics, physics, and engineering, whereas soft subjects' examples would be languages, history, and biology (Elton, 1998). Evidence demonstrates a higher proportion of first- and third-class degrees are distributed to hard subjects than soft subjects. This is a direct result of assessment and not the ability of our students. Moore (1995) suggests reasons for universities not producing many first-class degrees. He suggests the following:

1. The intellectual calibre of the students
2. The setting and marking of examination papers
3. Variation in the range of marks awarded
4. Regression to the mean

Three out of the four suggestions listed concern marking, whereas only one reason is reflected upon the student's abilities. It is detrimental that the marking process is vigorous and great in quality assurance as it impacts certification. Regardless of subject, equality and fairness should be a priority and another reason why parity across the sector is paramount when degree outcomes are based around grades.

There are a few concepts that revolve around marking and the potential for bias and these include the *halo effect* and the *contrast effect*. The halo effect is caused when the students' background effects the assessor (Fleming, 1999). In *Musicians in the Making: Pathways to Creative Performance* (2017) Gaunt also refers to the halo effect specifically in music teaching and how the halo effect can change the power dynamic in the lecturer-student relationship. If the student is favoured by the tutor then the assessor potentially would mark higher than a 'blind' marker (when work is submitted anonymously, or the assessor has had no contact

with the student being marked). The opposite can occur too, where if the student is disliked by an assessor, for whatever reason, the mark or grade given to the student could be affected. How would we monitor this behaviour? The halo effect can be spread if tutors share thoughts about students and inadvertently persuade others to mark differently (Archer & McCarthy, 1988). The contrast effect, a theory referring to written essays, finds the assessor marking the second half of essays against the marking of the first papers, comparing the assessments to one another and not against the relative criteria, which, in effect leads to the examiners setting their own benchmark standards. A study conducted by Daly & Dickinson-Markman (1982) discovered subsequent essays were affected by the essays marked beforehand. Examples of recognising this behaviour in assessors are comments such as “this is not like Jenny”, “this is too good to be Sam’s!” (halo effect), and “how nice to get a good one occasionally” (contrast effect). Combating the halo effect could be achieved by utilising the ‘blind’ marking system but the notion of assessing a musical performance ‘blind’ makes little sense. The assessor needs to watch the performance in order to assess the visual elements against the criteria. A potential way of eliminating bias is to use external parties to assess; however, this causes further problems such as financing those coming from outside and the potential lack of understanding of what is expected from assessing the learning outcomes. Marking comparatively is unavoidable in bigger institutions, unless there was a different panel each time an assessment took place. In order to combat assessors setting their own benchmark standards would be to reiterate those that are set in the QAA. Another theory regarding marking bias is Helson’s theory (Helson, 1964). The lowest and highest marks are not affected but the average marks (middle of the distribution curve) is marked lower. The first paper to be marked sets a prerequisite for the

standard of those that follow, and in comparison, the norm for the other assessments causing issues with criterion-referencing where lecturers become complacent in marking and mark comparatively exacerbating the cycle of auto-pilot marking. There are also what is called ‘marker factors’ to consider (Fleming, 1999). These are simple suggestions such as, not marking when irritable or tired, not marking too late, and taking plenty of breaks. Marking musical ensemble performances raises several issues here. Assessors are required to mark all day (and evening) in some cases. This does not necessarily mean the examiners have the ‘freshest’ minds after a long day. How do we improve these factors given the nature of our discipline and the need to imitate assessment environments of those in the industry?

Type of Learners

As previously mentioned in the introduction, the types of learners is an important area of discussion if learning outcomes are central to the assessment process. How can the assessment process aid our student’s understanding as well as encourage “model student” behaviour? What can we do to help the student to adapt to their study via their understanding of assessment? If more institutions discussed the grading criteria and learning outcomes in greater detail, the level of student understanding would surely increase. As pointed out before, the lack of parity in the way in which this information is disseminated becomes an issue. How and why do we expect our students to know this about the sector when this differs greatly across the sector, especially when in previous routes of education (college, sixth form, high school and so on) the assessment process and what is expected from them is clearly stated in black and white? With these education systems working against each other,

there is a clear divide that potentially leaves the student unprepared for curriculum delivered and assessed at HE.

A doctorate study on generic assessment practices, undertaken in Hong Kong (Tang, 1991), questioned students about their “preparation and strategies” for written examinations. The ensuing responses were then separated into four categories: *surface, deep, achieving* and *unclassified* (Entwistle, 1993) (previously referred to in the introduction). Surface strategies included “rote learning, surface memorization, [and] question spotting”. Deep strategies comprised of “deep memorization, relating information, visualizing patients’ conditions, [and] discussion”. Achieving strategies revealed “organization of study time and materials”. Unclassified strategies included processes such as “going through test papers, [and] underlining”. Tang based the categories on the Biggs 1978 model. Entwistle’s’ model (1993) would compare the ‘achieving’ strategy to the strategic learning approach. If our students work with achievement in mind, working to a validity model does not assist the students’ understanding about the importance of learning; however, if learning is what is most important then do we look to assess at all?

A study looking into how popular music students prepare for performance would be interesting and insightful, but is it possible to categorise this in the same way as how Tang categorised her study? For example, memorisation could potentially be a key factor in the way a musician learns repertoire and the level of memorisation could be recorded in the same way. The four categories however would naturally vary greatly to the Tang study (1991) though as what would be included as part of the preparation strategies for performance in comparison to written examinations; ‘unclassified’ strategies such as underlining and going through previous test-papers are clearly not applicable in the context of a performance

assessment. Potential ‘unclassified’ strategies would encompass learning strategies not traditionally used in music such as reading and listening.

Prior Experience

It is important to recognise that students will have experienced different teaching styles throughout their journey of education prior to university and these have the potential to impact on the way they learn coming into HE (Brown, 2018). The level of divergence in students’ prior experience is an area over which we, as academics, have no control. Students are not ‘equals’ when entering degree level courses (Brown, Rust, & Gibbs, 1994). Given that the students’ background and experiences vary greatly, part of what is needed at level 4 is to develop the fundamental skills required for performing music, but what do we value as the basic standard of performers? There are debates about the key skills required in music and the benefit of different theoretical approaches and an argument about ‘dumbing down’. A number of professional musicians are unable to read but still have a livelihood in performing as they rely on other skills such as their listening skills and ‘notating’ music in a way that is helpful to them. Though this may be the case in the industry, there is a strong emphasis on the ability to read in some institutions where it is part of the core curriculum (Orlando and Speelman, 2012). Knowledge of music theory enables musicians easily to interpret other’s compositions and a lack of understanding in this area can slow down the rate of learning. Yet, if this is not approached in a unified manner, how do we achieve reliable results when the learning outcomes differ greatly? These issues though arise much earlier than HE where in some secondary schools’ music is not offered a part of core curriculum, whilst others offer music but do not teach theory and the ability to read notation.

Some schools benefit from peripatetic teaching where others do not and even the standard of teaching delivered at this age varies too where the skill of reading is sometimes lacking in the tutor's abilities – perpetuating the existing problem.

At a recent workshop that I undertook as a freelancer I found that the sixth formers (Level 3) we had been working with were unable to follow their own stave on the piece provided (I mention this due to the importance of understanding transition pedagogy and the impact prior experience has on success later on (Kift, 2009)). These students aspiring to move into Performing Arts in HE were unable to demonstrate a skill which is a basic requirement to those wanting to study Musical Theatre at most institutions in the country. These students come from the more traditional route studying Advanced Subsidiary Level (AS Levels) and Advanced Levels (A Levels) (FHEQ, 2014) whereas others come from the non-traditional routes of Business and Technology Education Council (BTECs), Diplomas, and National Vocational Qualifications (NVQs). There is no defined standard at entry level, and students will have a diverse understanding of the subject. Institutions can make it a requirement to have these soft skills in place at the point of entry, but these students are suffering because the system at secondary level has arguably set them up to fail. If level 4 is in effect a year to 'level-out the playing field' then should we assess at all? Or should we simply encourage the learning and focus on engendering the processes needed to develop as musicians first, without the pressure of assessment? Understandably, marks at level 4 do not contribute to the end result of the degree, so this at least provides the students with a relatively unpressured practice run at the assessment period if nothing else.

Whilst learning outcomes and alignment exist the importance of the value of a degree should also be taken into consideration. What is also clear in this initial level

of research is that more investigation towards looking at reliability and validity coinciding and not merely ‘trading-off’ is needed, and that this study is not just applicable to the study of music considering the stressed importance of transferable skills and employment determining these factors too. Adjustments could usefully be made to accommodate not only the different type of learners, but also specifically targeting the students that have a less fortunate background in music education with the aim of raising the basic standard overall. The threshold of this standard and particular area should also be raised at the point of secondary education and earlier. The impact of change from policy level determining whether or not music is practiced as a core subject is a significant impacting factor to affect teaching and learning at HE which in turn impacts how we assess as well as determining how this is then communicated to the student to understand the way in which we assess. Continued lobbying (Incorporated Society for Musicians, 2020) to ensure that music is not marginalised within the curriculum at primary and secondary levels is therefore an important area to consider.

## CHAPTER TWO – GROUP AND SOLO ASSESSMENT

The literature explored within this study up until this point is focussed mainly on traditional methods of assessment, for example, essay, examinations, coursework and so on; however, what about non-traditional methods of assessment? Brown and Knight (1994) acknowledge that particular subjects will need to be assessed via different forms of assessment. They specifically mention the creative arts (music, dance, painting, sculpture, and theatre studies) and that creative production is a necessity. These subjects see students working as a whole team, in small groups, and/or individual performances (AEC, 2017). Rhodes and Tallantyre (1999) suggest that observing an individual within a real-life situation (for example, a work placement) may be the best way to assess skills in this context. Brown and Knight (1994) also recognise that these kinds of activities are complex in nature, requiring aesthetic judgement as well as academic reasoning. The difficulty in these subjects is defining and assessing the learning outcomes. Brown and Knight explain that the subjects' "course team… needs to rely heavily on explicit and overt criteria in order that his judgements are valid and reliable". Assessment here requires professional judgement which relies on the extent of the experience of the examiners as practitioners within their specialisms. How can this measure be controlled? Learning outcomes need to be agreed in order to judge "students' abilities against a set of standards which have to be flexible but objective". There appears to be little information regarding assessment and the creative arts and considering the growth popular music has seen in the last forty to fifty years demonstrates that this is an area that needs further research.

Assessing Groups

The performing arts sector within HE presents different issues regarding summative assessment than other disciplines due to the nature of the way our assessments differ – especially when assessing in groups. The concept of assessing groups has been discussed in more depth in relation to creative production (Brown & Knight, 1994; Brown et al., 1995); however, group assessment defined in the literature refers to presentations. Group assessment allows students to do the following things: develop a range of important skills, achieve more in a group than as an individual, and there is less marking and therefore less demand on resources (Brown et al., 1995). If applied to performance (music, dance, theatre studies) these type of assessments become increasingly more challenging and although requiring less marking could be more intense. The expectations from the assessors would be they would need to be alert for a longer period of time as usually group performances take place in the evening after a day's teaching. Assessing in this way presents other issues. Courses that have a large cohort would need to spread assessments over a longer period of time in comparison to smaller cohorts that can be assessed in one day. The disadvantage to a larger cohort would mean students who performed last would have theoretically longer to work on their assignment. However, the assessors at this point could grow tired from the process and start marking comparatively. Concerns regarding group assessment is not only the issues around marking comparatively but the "fair allocation of grade for group-work" (Heathfield, 1999). Group assessment proposes evident issues around marking consistently, comparatively, and fairly.

Polifonia Project

In the European Commission funded Polifonia Project the handbook *Admissions and Assessment in Higher Music Education* (Cox, 2010) published its findings to look at the "issues in assessment in Higher Education". Cox opens the discussion of assessment regarding individuals and the overview is of music in general but not specific to popular music. Also, it should be taken into consideration that the focus in assessment here is how assessment is used from admissions to final award, and only in part discusses and investigates the summative assessment process, whereas this study is predominantly looking at the relationship between summative assessment and the factors that cause challenges for conducting this process. There are some relevant issues however that arise in the handbook that can be applied to our sector and are important to acknowledge as part of the challenges that are arising from assessing music in HE as well as a specific case study that looks at how grade criteria design could possibly develop to improve assessment.

Cox explains the following:

> Higher Music Education, like any other discipline, requires its assessors to be consistent in their judgements and to benchmark these against widely acknowledged standards. Consistency is important because it underpins the principle of fairness in how applicants and students are treated. Fairness is a laudable principle in itself, but a fairness that can be clearly demonstrated in practice, and which is therefore based upon clearly accountable procedures, is also becoming increasingly important.

In the previous chapter, benchmark statements and the level of standard this represents was briefly discussed and it would appear that across Europe these

standards are adhered to, but Cox also acknowledges that the systems across Europe vary widely and is one of the reasons behind the Bologna Process. More information about the Bologna Process can be found at https://ec.europa.eu/education/policy/higher-education/bologna-process_en as well as the Dublin Descriptors http://ecahe.eu/w/index.php/Dublin_Descriptors which are part of the Learning Outcomes work. As this study specifically addresses the issues and challenges that arise in the UK sector then to a certain extent the details of this process are irrelevant however the argument for why it is needed is relevant to this study. The website front page describes that the process is needed because "widely differing education and training systems in Europe have traditionally made it hard for Europeans to use qualifications from one country to apply for a job or a course in another. Increased compatibility between education systems makes it easier for students and job seekers to move within Europe." (2018). One of the fundamental reasonings behind the thesis (alluded to in the preface) is the lack of comparability in performance assessment across the sector. It is encouraging to know that these sentiments are felt in Europe too. The statement assures us that the training we receive should indeed have parity in order to move around in the industry freely and not be judged by the training we received – performance is performance.

Consistency is an apt description of what we are trying to achieve in higher education, but the problem remains that with the flexibility of learning outcomes the consistency to apply fairness can be vague. He also terms procedures to be 'clearly-accountable' and, again, it is difficult to have a clear account of a performance when no two performances are alike.

Case study 8 in the Polifonia Project looks specifically at a conservatoire in the UK. The study that took place looked at possibly adopting a new style of grading. Aspects of the study helped to adapt current criteria but according to Cox it “did not go beyond a design exercise”. At the time it did not duplicate any current systems in place. One of the reasons for looking to develop the current system in place were listed as the following:

- Students’ recitals are not always consistently of one quality throughout; a student’s physical stamina and their mental discipline to sustain a performance across anything up to an hour are relevant factors in judging them as performing musicians. Should examiners base their mark on the best that the student has shown themselves to be capable of at some point in the recital, or should they mark to some imaginary mean standard of the recital as whole?

Out of the options given here, it would be reasonable to assume that most examiners mark to an ‘imaginary mean standard’. A mark should not be awarded based on what the highest (and lowest) point of a recital is. It contradicts the idea that marking should be consistent, and if the student cannot perform reasonably consistently then this should be taken into consideration. For example, during any public performance an artist may not perform consistently well. We are our own worst critics. We all leave however feeling an overall sensation of the experience and reflect – be that good, very good, or simply not performing at our best. Practitioners rely on their ‘gut-feeling’, mainly when professional standards have been obtained, but you cannot measure ‘gut-feeling’ yet what it does demonstrate is that the alignment of grades to

an ‘imaginary mean standard’ does indeed take place. How do we then identify the mean? Could there be a way to calculate this more specifically?

Cox (2010) raises the following points:

- Even where the recital itself is broadly of one quality throughout, some students are very proficient technically but relatively uninteresting artistically, while others can be enthralling to listen to despite being prone to occasional technical lapses. Individual cases vary but, in general, there is almost certainly a level of technical perfection that a live audience would be prepared to sacrifice for the sake of an inspiring performance; the question is how to reflect this in examination judgements.

- Then again, a performance may be of a high standard both technically and interpretatively, but somehow lacking in terms of the student’s engagement with the audience – their stage manner, dress, posture, projection and, in the case of spoken introductions being provided, their articulateness. These matters are perhaps secondary to core issues of how the music itself is played, but they do affect audiences and are relevant to a music student’s competence to enter the profession. As such, shouldn’t they also be reflected in a judgement upon a student’s performance as a recitalist - especially if competences related to communication skills have been linked with this assessment?

The last two points made are an accurate description for most when assessing performance – regardless of style or genre. How do we measure this? Are the

elements given equal weighting? Do we weight one more than the other, and, if so, how do we justify or validate our reasoning for doing this? The study identified three criterions: technique, interpretation, and presentation. For simple calculations (and to make the assessing process easier) the ratio was applied 2:2:1 respectively. The examiners felt that technique and interpretation were roughly worth the same value. Whilst this conservatoire remains anonymous, the conclusions drawn here is that the department involved in the study is likely to either be classical or jazz, where this opinion is reflective of the genres. The popular music sector would find the statement to be a little more controversial. We have styles that traditionally are not technical in nature, – the strongest example of this being punk – so applying equal weighting in genres such as this would not necessarily work.

The second reason for aiming to develop a more sophisticated design was based on the fact that examiners found difficulty in marking against 'musical response' in comparison to where the students was in their studies. The following examples were given:

- Say that two students, one in the third year of their 1st-Cycle study, the other coming to the end of their second year of 2nd-Cycle study, each perform a recital which feels of a broadly similar level. It seems clear that the student still at lower stage deserves some kind of additional reward for their achievement but, if so, how big should this be?

- It is true that the fact that the two cycles should have different learning outcomes, including in those areas relevant to recital performance, ought to help examining panels in their decision-making. In reality, though, examiners relying on these learning outcomes need to weigh up the

> nuances that lie in the slightly different wordings of 1st- and 2nd-Cycle learning outcomes and then put a numerical value on these. Not all panels feel equally confident about this and it is clear that, without specific guidance, there is some scope for inconsistency to creep into the ways in which different panels make such adjustments.
>
> (Cox, 2010)

"Specific guidance" and "scope for inconsistency" are highlighted here as main issues and it is difficult to disagree that these issues arise in popular music performance assessment frequently. Lack of training is something that will be discussed in more detail later on in the thesis, but it should be acknowledged here that this is one of the reasons for specifically looking to change grading criteria, alongside the subtlety of language. If we, as assessors, struggle to see the difference how can we expect our students to have a clear understanding of what is expected from them? Could these problems be resolved by just simply delivering more training around assessment as opposed to changing the way we assess to make it easier for marking? Arguably if we improve the way in which we grade we create more consistency for the student, but does this then have a detrimental effect on the student for simplifying the process as opposed to stream-lining the process? The final example given by Cox (2010) discusses that

- On a more practical level, some examiners spoke of finding it difficult when, as can often happen for timetabling reasons, they were obliged to assess in quick succession different students at different stages in a cycle or even different cycles. The sense of having to rapidly recalibrate one's own internal value-systems between each recital can be distracting. This

> raised the question of whether any necessary recalibrating could be made according to an external, pre-determined system. Such a system would enable the initial judgements to be purely musical, and only subsequently require examiners to make an alignment with a particular level in the student's ladder of academic progress.

Again, the conservatoire raises valid points that cannot be ignored. There are many assessors who report, through anecdotal conversations, that they feel there is not an issue with assessment and that our approach to assessment across the sector is adequate. The excuse is that we do not want to standardise our programmes and therefore we have a right to assess how we wish (Cox, 2017). This attitude does not put our students first, and leaves them, once graduated, in an incomparable system with little vote of confidence in the work they are applying for.

Moving on, the study also proposes the creation of a new scoring system that was unlike the conventional UK system of a mark awarded out of one hundred percent. Interestingly, the new scoring system is awarded out of sixty. Technique (24), interpretation (24), and presentation (12). Whilst the study does not confirm this, most conservatoire systems do not believe in marking in the top percentile. It is believed among many scholars that this is unachievable as the student always has more to learn, and there is always room for improvement. If this is the case, why does the top percentile exist if this is unachievable, and why can one institution apply this unsaid rule where others happily award one hundred as they feel that is the best that student can achieve – most importantly at the time? Again, the 'imaginary mean' evidently exists here. An award out of sixty implies that this is achievable and brings the best of both arguments together. In appendix D are the two tables which

demonstrate the redesign of the grade criteria. Technique, interpretation, and presentation is broken down into twos. Each category has a descriptor (Cox, 2010):

- **technique** *should be* **secure**
- **interpretation** *should be* **persuasive or, at the higher levels, compelling**
- **presentation** *should be* **effective**

The second table then distributes these marks across the HE grades system of 1st class honours, 2:1, 2:2 etc.. It keeps the terminology the same, but the marks are greater to achieve higher boundaries. Is this a concept that could work for popular music performance assessments and could this work in a group context too? If we were to adopt this form of grading how would this work for modules in which technique is valued at a higher ratio than maybe a group final performance where presentation is possibly equally as important?

The research suggests that the essence of comparability is definitely something that is sought for. The Polifonia Project investigates if there is a common process in which this can be done. It supports the idea that even with comparability we can avoid standardisation. Aligning learning outcomes with a common framework would bring a degree of parity. In addition to this, potentially building a list of popular music repertoire would also be beneficial. The list of popular music works would not be a definitive resource and should indeed be a flexible document; it would represent the standard for each level of study and it ought to be agreed upon by the vast HE lecturers of popular music in order to cover the large breadth of our field. It could be considered by some a syllabus instead of curriculum, but the flexibility of the document is for guidelines only, similar to the guidance in the validation documents

of how we may assess a module component. I predict that just the act of this process would bring debate and raise some contention of what the expectation of the standard of our students should be. It is evident that more research is required around group assessment as well as the management of the intensity of practice of performative assessment.

## CHAPTER THREE – THE INTERPRETATION OF LANGUAGE IN ASSESSMENT

This chapter seeks to analyse at how students can misinterpret language in assessment as well as investigate how teachers and assessors might break down the barrier of language and understanding of our practice. It looks to address, and delve into further, the interpretation of language by staff. It also analyses our use of language in assessment.

Perhaps one of the most problematic areas of the assessment process for students is the language used in assessment. The use of “final vocabulary” (Rorty, 1989) in feedback can have a negative impact on students. The concept of final vocabulary looks to, in this case, retrospectively tell a student what they have done but not necessarily why. Authors like Boud go on to articulate this particular problem (1995, 2000) and examine Rorty's’ terms and proposes that the language is “abusive” and “judgemental”. The word judgemental implies that the response is not justified or validated. It also alludes to opinion and that in itself is not a viable method to assess. Final vocabulary also includes words such as ‘good’, ‘right’, and ‘professional standards’ however Boud argues that they communicate “nothing of substance”. Interestingly, one of the repeated phrases used in music performance is the notion of professional standards, but what does it mean to be of a professional level, or how do we articulate what it is to be a professional? The use of language in feedback is used to highlight the problems that the student is already aware of, but they find that the feedback lacks guidance and direction. Boud also raises the question about what could happen if we did not have an assessment period (where all feedback is collated and distributed in one go)? Could we use physical interaction (verbal feedback), and would this improve the quality of the learning experience for

the student? One way of assessing in performance, whilst adding intrinsic value to the learning, is collating formative assessments and selecting the best performances to be used at summative. The advantage of assessing this way is that the students understand the reasoning behind the function of the work. Assessing performance through formative processes would alleviate the pressure around marks and potentially reduce stress around the assessment period. Admittedly though, performance should have an essence of pressure, most performers feel something prior to going on stage – be that nerves, anxiety, excitement, etc. but the additional pressure of the need to achieve in order to receive a grade is not reflective of the music industry. The disadvantage of this is that the work essentially is assessed twice, and feedback is given back twice (conforming to the model of giving back feedback in the assessment period).

If verbal feedback (Duke & Henniger 1998; Duke & Henniger, 2002) became common practice then maybe the language used would be less factual and less 'statement' like. The practice of verbal feedback is practised within the popular music degree programme at the University of Salford, where verbal feedback is given after performances. Other institutions follow a similar principle where a group session is taken after the assessment to discuss the performance in an informal manner. Sessions like these are essentially formative feedback but students remain unconscious to the fact that this is indeed feedback that they are receiving. It is important to note that since Boud raised these relevant and probing questions of the assessment process regarding feedback some external factors should be highlighted here. As part of the National Student Survey (NSS) feedback is one of the questions featured in the annual questionnaire. The NSS is also in itself a student feedback mechanism, and the impact this has had means that universities' now have their own

policies regarding how quickly they must return feedback after assessment work has been submitted. Interestingly, in a publication produced by the Higher Education Academy (2012) a comment regarding the "scores for assessment and feedback remain low in the National Student Survey as students express concerns about the reliability of assessment criteria, challenge the fairness of their experience and say they are dissatisfied with the nature and timing of feedback", but the HEA's reasoning for the dissatisfaction (as well as (William & Kane, 2008; Surridge, 2008, Beaumont, O'Doherty & Shannon, 2011)) comes from the lack of time staff have to devote to this element of assessment. This could be a factor but could the lack of understanding regarding formative feedback be an issue here? If students are unaware that they are receiving feedback then the question is null and void.

The questions in the NSS survey are also cause for concern (Williams & Cappuccini-Ansfield, 2007; Gibbons, 2012; Bennet & Kane, 2014). For example, the wording of the NSS questions almost looks to challenge the student and invites behaviour that leads to blame culture. Questions from the core questionnaire asked the following under "Assessment and feedback":

8. The criteria used in marking have been clear in advance.

9. Marking and assessment have been fair.

10. Feedback on my work has been timely.

11. I have received helpful comments on my work.

These do not feel like questions: they are merely statements that leave no room for flexibility or interpretation. Question 10 asks whether or not feedback was given back

in a timely manner. Does the student interpret this within the University/ Course policies or is it purely the perception of the student as to whether or not they *feel* like they received their work in a timely manner? Scenario 1: work is assessed in the last teaching week before Christmas week. For most places, there is a 4-week turnaround (working weeks). Meaning the student may not receive their feedback for another 6-7 weeks dependent on when that particular institution returns after the Christmas break – taking into consideration that staff may have taken holiday in this period. The adverse effect or not receiving the feedback more imminently is the amount of time that has occurred between performance and feedback which could lead to feedback being misinterpreted as the event has already passed (and potentially partially forgotten about). When writing an essay the work is there in black and white for the student to view with the feedback alongside especially with the development of the Turnitin software which allows for the feedback to be disseminated in both varying ways of commenting and with inputting aural feedback too. With performance, there is a lack of connection made when the student cannot physically draw upon the performance visually. The other challenge we face as lecturers are relaying tacit information. The nature of our discipline discusses sometimes the intangible. With an essay, it is clear and evident where the feedback is and what it directly refers to. In performance when we give written feedback there is no reference point. Most performers are not holistically aware of the performance or have the ability to recall their delivery with the skill to reflect. Observations undertaken by me at the University of Salford demonstrated giving verbal feedback after individual showcases in Level 6. The concept was good, and the students felt that they were able to discuss their performance and feel the value of the feedback given as they felt they could express their thoughts too. The students felt they were

able to contribute to the analysis of their performance and could reflect much more easily because it had just happened. Similarly, in sessions that I have delivered at various levels, I have taken classes that involved using previous footage of their performances and used this to feedback as well as to peer review. When asked in module surveys and informal conversations the student feels added value to have their feedback given alongside something tangible that they can see and witness.

Price and Rust (1999) and Price and O'Donovan (2006) observe language from a different perspective. They explore how tacit information presents its own challenges to the learner; however, the biggest challenge within music is the transfer of tacit knowledge, especially when this is presented in areas to develop as part of the summative feedback for performative assessment. Most practitioners/ assessors use their own personal experiences to inform their judgement. The difficulty for assessing voice, for example, is the instrument is a part of the singer's anatomy. How it could feel for one person might feel different for another so the way we describe how technically something could be improved can become mechanical in nature. This also creates problems when the student is not interested in the technical aspect alone and wants to know how to improve stylistically and you then comment on 'feel' and 'groove' words that sometimes do not articulate to the best of the word's ability.

The language in assessment criteria can be confusing to students too, especially when they may interpret the criteria differently from each other and then discuss what is expected. Assessment is a mystery to some students, and knowledge of assessment criteria is limiting for the student – unless shown by a member of staff to explain assessment within context (Brown, 1999c; Crook, Gross,

& Dymott, 2006); however, Price and Rust argue that even with guidelines in place there was no/little change to the work submitted by the student. If this is the case, what can we do as educators to change this? Guidelines could be too prescriptive and take responsibility away from the student to a certain degree. A solution to this particular issue may be that it becomes common practice to advise students on how to use the grading criteria. In my own professional practice where I previously lectured at Hereford College of Arts, I would run sessions in performance/ principal study masterclasses and allow the students to use the grading criteria to assess either their peers or examples found on YouTube. Feedback from these sessions have been positive and students feel that they understand what they are being assessed on. Whilst this is not a new concept, it does demonstrate that students would benefit from this practice in other institutions. Another way of potentially assisting and clarifying this process further for students is the practice where we link learning outcomes with the relevant grading criteria categories. It is also debatable as to whether or not institutions should think about standardising grading criteria and use the same template to associate their own learning outcomes against. This would allow some control over parity whilst allowing for individuality across the board.

One of the main areas of the discrepancy is the terminology used regularly in popular music. Terms that cause contention within the industry itself, in-turn causing internal debate in education does not improve the assessment process. Loaded terms such as *original* and *authentic* cause the most problems in popular music as these can be interpreted many ways, so why do these words appear in grading criteria if they are under such contention?

The biggest challenge in assessment is what constitutes something to be outstanding, excellent etc., and the impact this then has on the reliability of marks. Should we look to standardise performance regardless of form and ethos? Whilst this does not feel much like a solution, elements of what constitutes a great performance should be outlined. Is this something that could be added to the QAA, where a benchmark for the first-class standard is devised to keep assessment in-line across the UK? Undoubtedly, this is not just specific to music but again the idea of achieving something that is not tangible is hard to assess into the using the whole spectrum of grades available wherein other disciplines (or even other areas of music such as theory) there could be a finite answer. The interpretation of language does not just lie within the understanding of the language but the parity in which this language is used across the UK. What constitutes as outstanding in one place could be classed as excellent in another and this is a challenge for assessors who may mark in more than one place. It also does not help students comparing themselves to their peers in other institutions across the country. Why is this not standardised and part of the system when ideally our degrees should be worth the same value, regardless of where you study? Yes, there is, of course, an argument that musicians strive to study at conservatoires or universities that are deemed to be held in high esteem and reputation precedes but that should not mean that these grades are harder/easily obtained in other places.

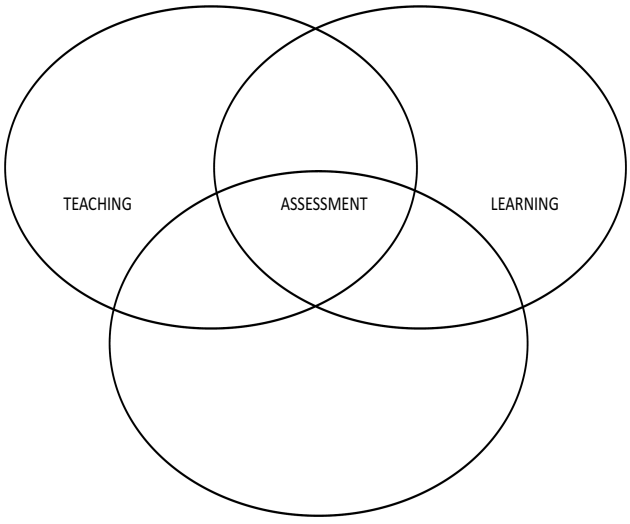
Whilst there is clearly a wider issue here in general, the challenge that comes in music is there is no clear set of vocabulary across Popular Music. If we look to standardise the terminology used across the sector then a better understanding would prevent confusion for staff, external examiners, and students. The language

that causes such discourse in the theories around authenticity and originality naturally causes conflict when it appears in the criterion.

## CHAPTER FOUR – ASSESSMENT, FEEDBACK, TEACHING, AND LEARNING

Much of the literature so far looks at either the following combinations: ‘Teaching and Learning’, ‘Assessment and Feedback’, ‘Teaching, Learning and Feedback’, ‘Learning and Assessment’ or any of the above as individual concepts. However, there is (as far as the research undertaken thus far indicates) no reference to frameworks that address all four as part of an integrated process. Taking this into consideration this chapter looks at the research around these areas and looks to understand why assessment features much less (and what is missing) than its counterparts in studies in HE.

Learning is essential to the assessment process, but teaching is equally as important. In *Assessment for Learners in Higher Education* (Brown & Knight, 1994) a Venn diagram depicts three ovals of equal measure.




(Brown & Knight, 1994)

The diagram demonstrates that all three factors are equal in size and importance. Where the ovals overlap, the crossover of areas show how these

three elements impact upon one another; however, what does not feature in the illustration is one other key factor of the assessment process: feedback. Feedback is becoming one of the most contested areas of assessment. Jo Johnson when Minister of State for Universities and Science (MP) stated

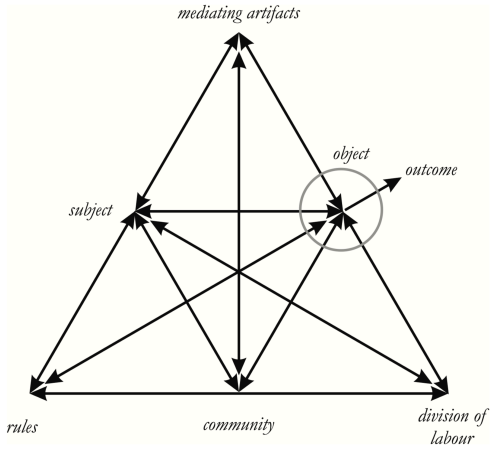
> the NSS records scores for 'assessment and feedback' and these have traditionally been the area of the student experience with the lower satisfaction levels. This area has seen a focused effort by providers and is now at 74% (from 73% in 2014 and 64% in 2008).
>
> (*Fulfilling our Potential: Teaching Excellence, Social Mobility and Student Choice*, 2015)

The importance of feedback is getting students to understand and learn from their assessment. Race (1995) raises a valid point and recommends that we frame our feedback, so it is constructive, useful, and meaningful to our students. Students want feedback for several reasons (Irons, 2008). They wish to understand their feedback – these reasons vary from not understanding assessment, misinterpreting the feedback given, and their perceptions of feedback being different to that of the staff. For these reasons it is argued that the way assessors feedback needs to be improved. These improvements range from quick and helpful feedback, the quality of feedback, and the importance of peer feedback (Falchikov, 1995). Quick and helpful feedback has been addressed by institutions where agreements have been created with students promising that they will return feedback within a certain amount of working days – whether this is achieved is another thing entirely. Boud

(1995) argues we should feedback from a subjective point of view, with description, and avoid writing in a style that is compiled of final vocabulary: similar to his thoughts concerning verbal feedback but in written form. Providing effective feedback is important for two reasons, according to Bloxham and Boyd (2007), the potential to enhance pedagogy (from formative assessment – and consequently feedback) and in raising academic achievement. Yet, tutors face issues in students not reading their feedback, or understanding their feedback and consequently this part of the cycle is broken. Bloxham and Boyd also refer to the "spiral" which demonstrates the progression of a programme. The visual concept integrates "guidance, feedback and target setting" into the course structure. Unlike summative feedback, the issue with formative feedback is that the student does not always realise they are receiving feedback at all in turn affecting their response to the NSS questions related to feedback. In performance, it is hard to believe that in a vocational subject where feedback is constantly provided in class that the idea of feedback being an issue within our discipline seems absurd, but it still suffers badly with satisfaction levels at NSS no matter how hard we try to focus our efforts. This is a noticeable issue within popular music performance, and it is widely recognised anecdotally that our students do not perceive feedback within a session to be feedback. How do we combat scenarios where this occurs? Explaining to students the many ways they can receive feedback is crucial when in secondary school and college they recognise feedback as weekly reports that they give to their parents, which is usually provided with final vocabulary: much to the contradiction of what is advised at degree level.

The tutor can improve students' learning and the assessment process by training them in assessment techniques (Race, 1995) or feedback literacy (Carless &

Boud, 2018). By doing so, the student is able to be “better managers of their own learning” as well as owning the ability and skills to assess self and peers. Some would argue that training in assessment techniques is like training the student for the assessment itself and helping them to only pass, but essentially this is a life-skill that every student would value in knowing and evidences good practice. It also allows the student to have “ownership of the process” (Race, 1995). This would work for self-assessment but less so for summative assessment. The more the student understands about their learning the more they are made accountable. Observing human activity theory (Engeström, 2001) the diagram below highlights the impact and contributing factors there are on an individual.




If we read this diagram with the subject as the learner and the object the tutor in reflection with Music (Henley, 2019) we can see how the outcome (of what we desire as a tutor or for the learner) has many other areas to consider. Much of the teaching in popular music performance is an ongoing collaboration between lecturer and student. We work closely together on their individual craft and these boundaries of responsibility can become blurred. The expectation, and indeed level of the division of labour, can become unclear for both sides. The subject, in this case, is the learner and the contributing factors that surround their environment in which they learn. In

addition to this, another challenge here is how do we expect students to understand how to assess subjectivity in an objective manner when, at times, tutors struggle to do the same? Therefore, guiding students in how to assess performance is a challenge within itself.

The Graduate Standards Programme (GSP) issued a report in 1997 stating that “the separation of teaching from summative assessment” was to be rejected as an option for recommendation, so the Quality Assurance Agency (QAA) (created to carry out the recommendations of the GSP) used constructive alignment “at the heart” (Bloxham & Boyd, 2007) of its approach to monitoring standards in HE. Erwin (1995) stated two years before the report that “deciding what to teach and assess is one issue, not two.” The Biggs model, which was explored before, is integral to teaching, learning, and assessment as we know. These theories underpin the research prior to the millennium and are valid today, but is this simply integral to our system in 2019 as this is so firmly embedded into the university sector? On the contrary, Brown (1999b) discusses how some academics think of assessment as separate to teaching and learning, and that assessment is difficult for them as they do not understand the reasons behind why we assess. There are, however, flaws in the system where separation is central to the assessment process. The design of assessment in this manner can cause certain challenges, mostly with students not understanding the relevance of what they are being assessed in. In striving to be innovative and creative, assessments are redesigned but when separated from the learning outcomes in mind, creating a division between teaching from the summative, the reason for why we assess then becomes meaningless. Even though Biggs’ model is the key to how we validate most programmes in the UK the mentality of redesigning assessments without learning outcomes in mind is most certainly

counterproductive. Brown argues that assessment should not be viewed as an option or a “bolt-on”, and this argument is agreeable to most academics. The most concerning factor though here is that assessment is still, more than twenty years later, a highly contested issue. It begs the question if assessment is such a contentious area of discussion is the Biggs’ model as perfect as it seems or is time, with the continuing change in HE that another model of assessment is considered?

Part of the argument is now no longer why do we assess at all, but do we assess too often (Santelises, 2015)? Whilst there is an issue surrounding whether or not we assess too often, I would argue that there is also another rising factor of assessment practice and that is do we now assess too many disparate areas within the assessment itself. For example, how many learning outcomes should we assess at any given point? The number of learning outcomes varies between institutions and the choice of whether we assess all learning outcomes at any given one point or not between components is another element to consider on top of this. Degree programmes vary widely with their validations, and the expectations each validator has as well as the impact they have on those that are validated through another body. Having worked through a revalidation three years ago (one where we had two differing remits to address), one of the key changes made to the popular music programme was to reduce the amount the student was assessed. The observations made from the previous validation was that there were too many components within the performance assessment and that the student’s grade was therefore diluted by too many additional assessed skillsets. We believed that is was not important for a student to write a reflection alongside a performance when the learning outcome did not, in fact, address the written element. It was essentially assessing for assessing’s sake and it added no value to the student or the assessment, it merely allowed

scope for the students' mark to suffer and ultimately detract from their ability to perform – which surely is the main focus of any performative assessment? This particular case is reflective of the not so robust system of the validation that went through before. It begs the question of how the course was initially considered for validation and what regulatory factors accepted this in the first place. Similarly, observations from my experience in other institutions is the noticeable changes that still need to be made in order to realign learning outcomes with assessment. Aligning learning outcomes is not exactly new knowledge yet somehow there are validated degrees that allow too much flexibility where learning outcomes do not align with the assessed outcome. In the *Higher Education and Research* Bill (2017) it mentions the Office for Students (OfS) role in the regulation of the validation process:

- "encourage choice and competition where this is in the interests of students and employers, the OfS will be able to take concrete steps aimed to improve validation services, and address some of the barriers new providers can face when seeking a validating partner."

- "address the lack of transparency and opportunity for providers to compare various offers. This will enable a more diverse sector and greater choice for students."

- "actively encouraging providers to develop validation services, and setting out exemplary validation arrangements to help informed negotiation between validators and providers who seek validation."

It is worth investigating further why disparity still exists if the role of the OfS is to address the lack of transparency, which undoubtedly is causing some of the disparity?

The Teaching Excellence Framework

The idea of the teaching excellence framework (TEF) was introduced by the government in November 2015. The TEF was instigated to assess the level of quality across higher education at provider level. There are six core-metrics in total that assess the following:

- The Teaching On My Course
- Assessment and Feedback
- Academic Support
- Non Continuation
- Employment Or Further Study
- Highly Skilled Employment Or Further Study

There are several flaws in the TEFs system which determines status from the students understanding of the NSS which is vague in nature, and ultimately encourages a complaint culture if misunderstood. The TEF has demonstrated over the last four years that it is difficult to create parity, assess, and grade across the institutions when comparing Universities to alternative providers. Does it challenge the argument that if this is the case then can this be done across department level? The move to pilot TEF at a subject level commenced in 2017-18, with the intention of phasing the TEF at subject-level in two years' time. With awards lasting up to three

years, universities' and colleges will look to reapply (or apply) for TEF status, and with due consideration to subject-level TEF it is vital that programme leaders and lecturers know and understand the TEF in order to maintain or attain status. With providers now needing to abide by external regulators and meeting their requirements, it is surprising that the earlier issues raised continue to exist; however, the increase of institutions and the increasing battle for students will cause this knock-on effect of standards slipping, and institutions cutting corners. In the Main Panel Chair's report, Professor Janice Kay states that "one of the clear findings arising from the pilot is that neither model was fully fit for purpose for generating ratings". It would appear that the TEF has some way to go before becoming a reliable source for regulating and assessing HEIs and alternative providers. In the Arts Panel report the subject-specific comments that were the most useful to this study are highlighted below:

> 50. Arts is the only subject group that only has one CAH2 subject linked to it. This means that there is no differentiation between subject group (arts) and subject (creative arts and design). The subject unit of creative arts and design covers all aspects of visual arts, design, performing arts, photography, film, TV production, music and dance. The panel liked the way that CAH2 brought together studio-based teaching and learning practices. In terms of student share of higher education population, this is the largest subject (much larger than Celtic studies for example). This means that Model A and Model B play out in ways that are different from those found in subject groups that have between two and seven subjects within their grouping.

51. The key advantage of the current CAH2 approach to arts is that by drawing together these diverse disciplines into one subject the TEF panel is more likely to have reportable data. This increases the robustness of the decision-making and ensures that TEF is rooted in student outcome data. The key disadvantage of CAH2 for arts is that it brings together and merges delivery from across conservatoire, specialist and mainstream further education and higher education art, design and drama. These disciplines have very different approaches to teaching and contact hours so the usefulness of teaching intensity metrics will be very limited. The challenge associated with the breadth of CAH2 creative arts and design surfaced where drama and art and design were located within one provider and it appeared to be the case that what was being reported on was two quite distinct subjects. This also pointed to a concern that arts' breadth and size had the potential to allow some of its sub-disciplines to hide within larger metric workbooks.

52. Whilst the panel concluded that it was able to secure ratings across creative arts and design if changes are going to be made to CAH2 the panel proposes that performing arts, music and dance are disaggregated into a new subject area. Additionally, the panel proposes that architecture and communication and media are brought into the Arts Panel.

The lack of differentiation between the subject groups is an issue and it would make sense that the panel judged the bringing together of studio-based practices to be a beneficial move. It highlights that there is no single 'one-size fits all' solution for a department as large as this within larger universities. In specialised institutions, there

is maybe an argument for grading the school the same way. Music specialist providers, for example, would work well under these models and there would be no issue with these places hiding sub-disciplines within the data as there is nowhere to hide within the larger metrics as highlighted above. The key disadvantage of merging the different types of teaching delivery does not only affect art, design, and drama but music too. The panel proposed that performing arts, music and dance created another category to be assessed too which is certainly an advisable suggestion that should be seriously taken into consideration. It would appear that this is reflective of most of the research that exists in tandem with education. The arts are grouped into one category where this thesis argues that music should be researched further in line with educational processes as our pedagogy is also not a 'one-size fits all process'.

Course Design Changes in Revalidation

Much of the research in music is lacking around assessment and the way in which we assess. This is frustrating as our discipline genuinely has differing processes and approaches due to the nature of the discipline. It would also be beneficial to the practice if this was as documented as heavily as the research around assessment in the early 1990s (and earlier) with writing and tests in general. It is not to say that there is no research out there but simply to say that there is not enough. Do we as practitioners believe that as a vocational subject our research predominantly revolves around our/the practice and the pedagogy only? The examples I present next are areas of my work as a lecturer and the changes that I have made, with my experienced colleague David Bebbington to document some of the work that is being practised on assessment. I am sure there are others, but the music community do

not discuss their practices enough (recent discussions at NAMHE have stated this too) and that our methods are discussed less than other contemporary practices. At my previous place of work, we were commended highly by both our external examiner and our external panel of advisors for revalidation on the flexibility with which our students were assessed. The flexibility of these assessments has come from the course design where there is flexibility built within the teaching. The modules are shaped so that students can access information in different ways. For example, where students can be assessed in performance, composition/ arrangement, or both could be due to the module learning outcomes focusing on creativity – which can be demonstrated in both performance and composition. The streamlining of these modules, by reducing the number of elements contributing to an assessment, and the flexibility of the outcomes naturally allowed for better student engagement as the perception was they could tailor the course to suit their individual, creative journey. We removed the written element that existed in this module to remove natural grade deflation. Some students felt that they were assessed unfairly as their perception of their grade of 'performance' was distorted by the reality of them having to be assessed on their writing and analytical skills. Student engagement will inevitably improve the performance of the individual. They can focus on their interests by developing their own professional practice in a unique way thus offering a better experience. Performance is the main contributing factor as to why students opt to do performance degrees: obvious. More of the performance element then surely should not be an issue in our field. The level of student engagement will naturally be higher and therefore the assessment should not be divided from the process.

Feedback and Grading

Feedback design for performance, is important. Examples of how these are designed differ across the country and they identify different aspects. When writing feedback directly to each learning outcome (sometimes presented in individual boxes) one separates the learning outcomes which for all intent and purpose should be intrinsic at all times. This method is encouraged to guide the assessor in how much they should or should write. This is mainly aimed towards part-time members of staff or professionals that do not have a background of academia. Good practice conducted at Hereford College of Arts demonstrates a visual aspect that echoes the grading criteria in order to show students where they belong in the grade boundary. Typically, this would be an asterisk in the box, but this has been developed further still by placing this in the lower, middle, or upper section of the box to indicate where against the grading boundary that particular learning outcome falls. This development came from the music department of the college.

Again, a disparity in marking in music (but possibly exists outside the field, for this thesis I have focused my research around popular music) looks at marking in the middle of the range. For example, using only a 2, 5, and 8 in each grading boundary (42, 45, 48, 52, 55, 58 etc.): ignoring all other marks. Though it must be added that the full range of marks is utilised at the fail boundary. The institution(s) must remain anonymous. The argument is that there is not much difference between 42 and 43. The system that this encourages is that the value of the percentage is not worth having and is similar to the reasoning behind not using the top 20% of the grading boundaries (a tacit system used within music). Other assessors do not like using the cusp boundary (that of which that ends in a 9) due to the students complaining (again not a problem just within music). In my experience, and discussions with

colleagues, the issues that around the cusp is how do we validate to the student this mark. What was that ‘one mark’ we could not find to put them there. It could be argued that if the assessor tried hard enough they would be able to ‘find this mark’. Others would say a mark ending in 8 communicates the student is close but not close enough. If you are able to write efficient feedback the cusp mark can always be justified, and the visual feedback design allows the student to see how this works. Avoiding using the full range is a benefit to the member of staff and there is no benefit to the student where you essentially eschew marks.

## CHAPTER FIVE – PROFESSIONAL STANDARDS

In popular music we refer to ‘professional standards’ frequently, and the statement itself is held in high regard, but what is to represent the equivalent of those at professional standard? Who accounts for what this is, what the threshold may be, and who fundamentally has the expertise to deem the student equivalent at the point of assessment? This chapter seeks to break down what professional standards are to investigate whether or not professional standards are on the decline, and to see how we can maintain professional standards if they are indeed being maintained.

It is unsurprising that the reputation of a university is important to students (Purcell et al., 2008). The UK system places emphasis on league tables such as The Guardian or the Times Higher Education. However, comments made by Johnson MP (2015) stated that

> League tables are not always an accurate reflection of the quality of education provided in each individual course. In addition, we know that students require a wider range of information. Course quality, teaching intensity and contact hours are all examples of information that are relevant to students. Information from the National Student Survey (NSS) (involving around 300,000 final-year undergraduates each year since 2004) and the annual, Higher Education Policy Institute surveys (undertaken with Higher Education Academy in 2015), gives some insight.

The question here is not about students finding the reputation important particularly but why. Degrees, if comparable, should mean students of the same degree

classification are equal on exiting their degree award; however, Moore (1995) asks the question “how often have you heard comments like:

- of course, a 2.1 honours from this university is the equivalent to a first (or a 2.2) from… university.
- Mind you, it is far more difficult to get a first in physics than in history.
- You know, standards are slipping. We must do something about it.”

The comments above suggest that comparability does not exist and that there is an also an issue with prestige (Blackmore, P.). The concept that achieving comparability becomes problematic is supported by Elton (1998). He acknowledges that degree programmes have diversified in size (this statement is increasingly valid for today’s education system) and that the common standard – although he argues with the exception of the threshold level – is even harder to achieve with so much diversity. He suggests the possibility that the standard of degrees are going sideways as opposed to up-and-down. Elton also predicts that the future could be that ‘anything goes’. The reality twenty-one years later is that degrees are falling in standard even though the research (academically) is lacking to support this statement. Yet, the question still remains: if the evidence surrounding league tables demonstrates that more information is required, and that teaching quality and student satisfaction can be found elsewhere, why do the previous statements still exist at all? It is widely believed that professional standards are slipping. Though, once again, there is little research to support that this is the case there are numerous press articles that challenge this notion. Journalists Helen Crane and Harry Lambert write about the “slipping standards” (Crane, 2012) and

“how the British degree lost its value” (Lambert, 2019). Whilst these articles, to a degree, are written to incentivise the public and to encourage debate, the sentiments of the articles are worth discussing. Whilst both these articles are written seven years apart, they argue that higher education values and standards are diminishing. Crane states that:

> “in arts subjects, standards do seem to vary widely between different courses and institutions. Perhaps it is time that a system of UK-wide regulation on marking was introduced. The sharp divide in worth between the 2:1 and 2:2 degree classes is also one that needs to be addressed”.
>
> (2012)

Whilst Crane speaks from personal experience, these issues may still be relevant. Lambert too speaks freely on the subject and many of his sources are statements made by other lecturers/professor’s own experiences – the fact remains that, albeit anecdotal, these challenges still exist. Lambert surmises that Universities are not incentivised in the moral academic high-ground that came before. He alludes to grade inflation possibly contributing to the decline of professional standards, and he quotes that Professor Fenton (Goldsmiths) states “students have been shackled in the way they learn” and expands on this further by saying that “rote learning has been a persistent problem in schools, but, it had, until, recently, been largely absent in Universities” returning back to the idea that students’ mindset and learning type have significantly changed and therefore has impacted on standards in general across the UK. The data Lambert acquires suggests that “Firsts have quadrupled from 7 per cent in 1994 to 29 per cent in 2019” which, indeed, is an alarming figure

and is not, unfortunately, testament to the growing excellence of our students, but more the steady decline of standards and inflation of grades across the sector. I say this as, from personal experience, students have spoken openly about how they were surprised with their marks and that they did not feel they deserved the First-class degree they received. With this in mind, I look now to research how this affects the way in which we assess, and if it is necessary to, as Crane says, “regulate the system” in how we assess and grade our students work.

When discussing comparability and standardisation, two types of assessment referencing should be discussed: norm-referencing, and criterion-referencing. Norm-referencing defines performance relative to another performance, for example, marking the students’ ability against another student regardless of personality and characteristics. Criterion-referencing describes assessment of students against “pre-defined criteria”. There is, however, an additional type of assessment based on criterion-referencing: ipsative assessment. This type of assessment means that the student is compared to past performances. The best way to understand this concept is to use the sports analogy where an individual competes against their ‘personal-best’. Another concept discusses consistency and comparability regarding summative assessment (Bloxham & Boyd, 2007). They argue that curriculum and delivery does not need to be similar in approach, however they do state that “the level of learning to be demonstrated and the notional hours required to complete the assessed tasks” is what is important. Their perception allows comparability and consistency to coincide with summative assessment, and that through the regulations of course approval (as well as second marking and external monitoring) this can be achieved. Most popular music performance degrees assess against pre-defined grading criteria. Interestingly, whilst we do not outwardly apply norm-

referencing in performance as a method of assessment, it is worth asking whether or not as assessors we subconsciously mark group assessments against the quality of the work we see during evening of performances. The way we organise band line-ups is to have the strongest act last (to replicate the headline slot) and students are very much aware of this practice. If we were to assess performance through ipsative means (Hughes, Okumoto & Wood, 2011) where you assess comparatively to a learner's previous assessment we would be able to measure progress and development of our musicians as opposed to grading them against pre-defined learning outcomes that may not be relevant to the individual.

Quality of assessment is a growing issue. Johnston (1994) reviews six reasons as to why degree awards are increasing for higher degree classifications within the sector. The most valid, by his own admission, states that the biggest causality is the "changed modes of assessment [which] are leading to better evaluations of student quality resulting in increased achievement levels." Of course, Johnston is referring to traditional methods of assessment. Can the same be said for the creative arts in HE – and specifically music? Fullan (1991) suggests that the change in quality assessment, or the lack of change, is about people and not the assessment itself. He goes on to say that change is personal. Academics feel like their professional identity is in question (Archer, 2008): preventing change from happening. In popular music we value our practice and we value the fact that our staff are continuing their profession outside of the education system. With this additional layer of professional identity is it any wonder that the academics within our subject area take criticism even more personally that their own practice is being challenged? Controversially, is the problem that these members of staff should not actually be in the privileged position of teaching in Higher Education – let alone

assessing the musicians of our futures? Or that those who are privileged to be in these roles should feel more confident within these roles. The characteristics of professional identity could be extended to the institution itself where the pride of reputation is under scrutiny and therefore do not move forward in their practice. Agencies such as the QAA require statements regarding subject benchmarks and threshold levels, and institutions are required to give information to the Higher Education Statistics Agency (HESA) with an annual audit of staff qualifications which are published.

The Role of the External Examiner

The External Examiner shares some responsibility regarding the quality of assessment. The EE's role, when referring to quality, has these three functions: the ability to assist in the decision of academic standard (awards and the appropriation of student performance levels), assess process measures against the intended outcomes the student achievement, and whether this is fairly operated, and that the institution compares it standards against other institutions (HE) (Agency, 2000). These principles are straightforward. Nevertheless, the understanding of the role of the external examiner is not universal. Harmless conversations between colleagues can lead to misguidance, advising that externals have little responsibility when in fact – if the job is carried out correctly – they have great responsibility towards quality assurance in assessment. Those newly qualified though can find themselves in situations where they are misled into believing that they have more responsibility than they actually have. In a personal predicament surrounding quality assurance I found myself in a situation where the one giving advice had not been an external examiner themselves and gave incorrect information. I sought the advice of

a colleague (an experienced external) who agreed that in this particular scenario I was correct. As a young academic my judgement was overridden, and a student's grade was inflated incorrectly. Whilst these issues exist anecdotally, they do not exist in isolation. Research into new lecturers' professional developments and their experiences are discussed by Barlow & Antoniou, 2007; Nicholls, 2007; Warhust, 2008.

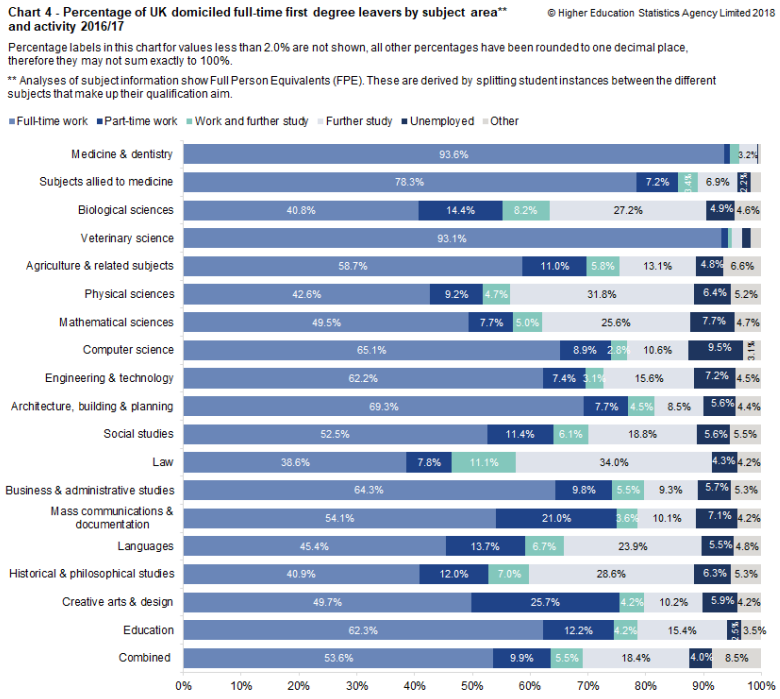
External Bodies

The pressures from the public sector (widening participation, industry professionals, governing bodies) have caused universities to extend their admission intake. Additionally, it has meant that retention is of utmost importance, but as Clark (1960) argues, what happens to the student who is destined for "early termination"? Purposefully choosing this aged quote demonstrates that this is not a new issue, and that in fact retention is harder now with even more students in the sector. With institutions lowering their entry standards there is an argument that failing students is inevitable but is it any wonder that this occurs when the sheer quantity of students being accepted (potentially for monetary value as an incentive) creates an inflation in problems relating to retention. However, Heywood (2000) discusses that "a high drop-out rate may not mean inefficiency; it may mean that a subject is particularly good at… sifting out potential failures: that it does so early is good for both student and the subject". Clark terms this as "cooling-out". The "cooling-out" period should indicate less failing overall, meaning results would not adhere to the distribution curve so easily. The negative impact of increasing numbers and greater drop-outs for retention however could be interpreted as failure and reflects badly on the

university's reputation, as well as the implication that HE is not caring or nurturing its young learners.

Other external factors that impact the assessment process involve employers and the professional statutory and regulatory bodies insisting that graduates must depart education with a specific set of skills. Heywood (2000) acknowledges that educators have the knowledge and skillset to provide in their specialism, yet there is now additional emphasis on the importance of transferable skills. Transferable skills (also known as *graduateness* (Glover, Law & Youngman, 2006)) within HE is not a new concept as Barnett (1992) considers "what's transferable about transferable skills", and his argument applies today where he questions the feasibility of transferable skills when courses are still narrow in discipline. Theorists have since developed a new term to represent these types of skills after education namely lifelong learning. Lifelong learning, as explained by Boud (2000), is to acquire and improve "knowledge skills and predispositions to underpin lifelong learning". Boud (1995), had previously argued that assessment undermined learning; however, he later questioned how assessment could contribute to lifelong learning? One of these questions asked if assessment tasks would equip the student to "engage in their own self-assessment now and in the future?"

Looking to the future, a student should feel prepared for employment if HE is now considered a career stepping-stone i.e. a direct route into the field that one studies. Yet, in a discipline such as music the likelihood of our students graduating and progressing into music careers straight away is unlikely. Data from the Destination of Leavers from Higher Education (DLHE) 2016/17 report is presented in the following table:




(DLHE, 2016)

With music falling under the creative arts and design category, 49.7 per cent of our graduates (those that filled out the DLHE) are in full-time employment, 25.7 per cent are in part-time employment, 4.2 per cent are in work and further study, 10.2 per cent are in further study, 5.9 per cent are unemployed, and finally 4.2 per cent fall under the category of other. Now, the data does not provide how many of those employed (full-time, part-time, or both work and study) are employed directly in the field, with this being potentially less for music graduates. The information presented here was the last and final academic year to fill out the DLHE survey. The 2017/18 graduates submitted their answers to the Graduate Outcome survey. The first survey was conducted in December 2018 and the results from the data collected in the first year will not be released, unfortunately, until Spring 2020. The gap between education and industry has decreased considerably since the major changes in 1992, thanks to the QAA and the subject benchmarks, but there is work to be done

still in this area. Race (1995) argued that funding bodies and quality auditors ""should align their policies so as to recognize, reward and promote best assessment practice". Whilst the quote is outdated, the fact remains that the bridge between university and employment could be made smaller.

Quality Assurance

The QAA stipulates that "information, guidance, rules and regulations on assessment should be clear, accurate, consistent and accessible to all staff, students, practice teachers and external examiners" (2006). There is a certain level of transparency required from higher education institutions. They are expected to publish documentation relating to programme specifications and marking schemes. It should be noted that the QAA advise only on the threshold and typical level. Boundaries other than the threshold level (pass-rate) are left to the universities' discretion. O'Donovan, Price, and Rust (2004) also address transparency but between student and staff. Teachers must find methods to help the student learn about assessment and the tacit knowledge associated with assessment.

Scrutiny of the marking process, however, is lacking in literature (Smith & Coombe, 2006; Yorke, Bridges, & Woolf, 2000). There is advice suggesting that we use criteria to assess (Swann & Ecclestone, 1999) but there is no literature advising how we should mark, or studies to elaborate on this point further. Research does however consider marking consistency and the increase of quality in procedures. These factors could impact the process of how we assess and, in turn, how we finally grade. Bloxham and Boyd (2007) discuss how tutors do not take the time, or have the energy, to spend on marking. Some members of staff can also view the work as laborious and low in status (beneath their pay grade). Another concept

that is unexplored is how tutors attach their own experiences and professional standing to their ability to mark.

Moderation

Nuttal (2007) evaluates moderation as a means of gaining greater consistency in marking. The following currently aid the moderation process: subject benchmarks statements; Framework for Higher Education Qualifications (FHEQ); Code of Practice; external examining and second/double marking. Second marking and double marking discussed in context have in recent years become interchangeable, but for clarification a definition is needed to describe what the difference is. Second marking allows the second marker to see the first assessors feedback but not the marks given. Double marking is essentially both markers assessing 'blind' to the respective examiner. In addition, Nuttal defines the resolution of marks. General agreement should occur when the second/double marker agrees with the first markers marks on the whole. Any discrepancies are ignored. Inconsistency between the first and second marker could result in further scrutiny of why the first marker came to their original marks. Adjustments may be made to all students affected, in order to be fair, so the marks are consistent overall. The adjustments applied could be either end of the spectrum or where the natural curve of distribution would lie. If neither assessor can agree, compromise, or come to a final decision a third marker is brought in to the process – to repeat the role of the second marker. Agreement is then made between the two (out of the three) which are closest in marks.

Degree Classification

Progression and level are undefined within higher education, and no agency advises what this may be. The terminology in grade descriptors vary (outstanding, excellent, very good and so on), in some institutions very good could mean 70, whereas in another university it may equate to the 60 boundary. However, Rhodes and Tallantyre (1999) discuss the attachment of level to skill (though not specific to higher education). The argument here is about attempting to distinguish not only grade boundaries but the difference in levels (4, 5 and 6). The lack of advice around the difference in degree levels impacts on decisions around repertoire for performance too. There is an issue that the discipline does not have a standard canon of works which is something that exists in both the classical and jazz domains. The collation of a list of works would be one means of representing the standard required at each level of the degree. For example, the repertoire in some institutions allow the students to perform songs consisting merely of four chords and performing these excellently in order to achieve firsts whereas other places that play complex repertoire (which is more aligned with the standard required) who do not perform as well suffer from their marks being lower as it impacts the performance aspect of the grade. Agencies like the FHEQ advise ‘threshold’ and ‘typical’ level at all undergraduate levels (and master levels too), but similar to the QAA does not advise the boundaries of each classification level. Could any of these individually or collectively be what is causing comparability issues? Most universities are accustomed to seeing a common assessment grid or rubric. These are sometimes subject-specific, but increasingly the model is being adapted to cover faculty areas of discipline and in specialised HEIs it encompasses the entire institution. Would it be possible to apply the model to one subject across the entire sector? Keeping the

grading criteria common (and more uniformed for parity) whilst cross-referencing against different learning outcomes against the differing courses (keeping value and ethos)? Price and Rusts' studies using SOLO (Structure of the Observed Learning Outcome) taxonomy was positive. Staff said that students "appreciated the grid especially where the staff drew the students' attention to it. The student feedback they had received also indicated a positive response to the grid." What was surprising from the results however was that in almost every case the students' performance did not change even when the students had a deeper level of understanding of the grading criteria. The study did not indicate when the student was shown the grid in relation to their assessment nor did it consider if this could have been a factor affecting the measure's impact. The implication from the feedback is that the students' standard of performance did not improve at their next assessment, which suggests that the feedback sheet did not fulfil its purpose. Staff using the grid reported that "sometimes it's easy to arrive at [a given] conclusion and the grid confirms it. Sometimes there are borderlines so it's good in that respect or poor in that respect ... it makes clear what one's thinking is, either to the student or another person marking, or the external."

Grade Boundaries

Borderline marking (referred to in Chapter Four - Assessment, Feedback, Teaching, and Learning) is a growing issue and one where assessors shy away from wanting to mark on the borderline: usually to avoid complaint and the avoid justifying and validating their mark. It is the general consensus though that the use of SOLO taxonomy is that the grid helped to clarify in several ways. The project proved that

the grid had potential and real benefits to staff and students, but that it “raised serious questions” about how we set, and inevitably define, standards in HE.

Similarly, the SOLO taxonomy works alongside Biggs’ newer model observing qualitative approaches to grading students (Biggs, 1992). The model looks at terminology and what the student should demonstrate at that level:

a. **Most desirable** (extended abstract): metacognitive understanding, students able to use the taught content in order to *reflect* on their own teaching, *evaluate* their decisions made in the classroom in terms of theory, and thereby *improve* their decision-making and practice. Other outcomes: *formulating* a personal theory of teaching that demonstrably drives decision-making and practice, *generating* new approaches to teaching on the basis of taught principles and content.

b. **Very satisfactory** (relational): students can *apply* course content, and *recognise* good and poor applications of principles. They "understand" in that course content is used as a theory of teaching that drives action.

c. **Moderately satisfactory** (multistructural): students understand declaratively, in that they can *discuss* content meaningfully, they *know* about a reasonable amount of content, but don't transfer or apply it easily.

d. **Barely satisfactory** (unistructural): sparse understandings, evidence of some effort in the acquisition of terminology; higher level understanding offset by some misunderstandings.

e. **Unsatisfactory outcomes**: fundamental misunderstandings, lack of effort/involvement in the unit.

Biggs refers to (a) – (e) as the American grading system of “A”, “B”, “C”, “D”, and “F” respectively. Yet, the conversions can be roughly translated to the UK system “1st”, “2:1”, “2:2”, “3rd”, and “Fail”. He also goes on to state that the criteria above can be used for both teaching and assessment outlines. Biggs’ models is measured through validity, but degree value is equally as important. Certification is integral to higher education and holds great value yet seemingly places less emphasis on reliability in this country. Questions around key skills and whether they should be integrated in degree classification or if it should be included as a “complementary *qualitative profile*” is discussed by Rhodes and Tallantyre (1999). Bloxham and Boyd (2007) takes this one step further and reasons that certification is one of the four purposes of assessment. Certification establishes “different levels of achievement” between students but also allows students to practice in the profession (providing the degree is a professional programme, such as medicine, law, engineering and so forth). The process indicates that certification is the result of assessment of learning.

The Subject Benchmark statement also outlines the following:

5.1 This Subject Benchmark Statement does not define or imply a common curriculum for Music. Indeed, the diversity of provision means that standards can only be measured against the learning outcomes of individual programmes. Given the different specialisations of degree programmes,

standards within the subcomponents of the discipline will not be common across the sector.

5.2 Given the richness and diversity of Music studies at degree level, the following outcomes do not represent a checklist that all programmes should follow. However, some broad general criteria can be put forward as indicators of both threshold standards (the minimally acceptable standard achieved by a bachelor's degree with honours graduate) and typical standards (the level of achievement which describes student performance around the median, which is where the performance of the majority of students currently lies, that could act as goals or targets).

5.3 While explicit standards are necessary for the development and review of programmes, these are not intended to lead to standardisation of the study of Music at the honours level. On the contrary, diversity of approach within the same discipline or sub-discipline can have positive value in questioning received wisdom and in developing good practice, thereby moving the study of the subject forward in innovative ways. The plurality of approach is a recognised strength of higher education in the United Kingdom.

5.4 The benchmark standards expressed in these tables are for programmes leading to Bachelor's degrees with honours in Music. They should be used in conjunction with the generic descriptors for qualifications at honours level as referenced in the Frameworks for Higher Education Qualifications of UK Degree-Awarding Bodies (Qualifications Frameworks), formerly published

> separately as The Framework for Higher Education Qualifications in England, Wales and Northern Ireland (FHEQ), and The Framework for Qualifications of Higher Education Institutions in Scotland (FQHEIS).
>
> (2016)

5.1 of the statement outlines that due to the diversity of programmes standards can only be measured against learning outcomes. It also states that that standards will not be “common across the sector”. Is this not a concerning factor that there is no common standard across the sector? Why is it that standards are compromised in order to deliver a diversity of programmes, and specifically if performance is deemed to be similar across the sector this is a contradiction? 5.2 outlines that there is, however, a threshold level for the minimum a degree should provide and a typical level that a student should achieve. If there is a common threshold for achieving essentially a pass at degree level, then again why is there not, on a general basis of grading criteria, an indication of threshold level for each grade boundary? 5.3 discusses why we do not standardise programmes. It is worth repeating here that the answer to the issues and challenges in popular music is not standardisation and this point is extremely valid; however, it still should be argued that comparability of professional standards should. This would still bring “positive value” and develop “good practice”.

## CHAPTER SIX – TRAINING (GAPS IN CURRENT PRACTICE)

One of the factors around assessment is the perceived lack of training. This chapter looks to assess if this perception is correct and if training is required then there is an area in which we can better improve our practice for the student experience. According to Brown, Race, and Rust (1995) “many of us working in higher education regard assessment as being a crucial element of the learning process, yet training is rarely given to lecturers new to the profession or wishing to develop their assessing abilities further”. Whilst the quote dates back twenty-four years, it is still applicable today. The training expected of aspiring teachers at secondary level and below is to achieve a relevant teaching qualification such as the Postgraduate Certificate in Education (PGCE), and most tutors at further education level are required to have a Postgraduate Certificate in Education (PGCert). Many full-time members of HE staff are encouraged to also undertake a PGCert in Higher Education, to apply for Associate Fellowship of the Higher Education Academy (AFHEA) and/or Fellowship of the Higher Education Academy (FHEA) status. Senior members of academia are encouraged to continue this to Senior Fellow level (SFHEA). The latter qualification is recognition only for teaching at degree level, from the formerly named Higher Education Academy, more commonly known as AdvanceHE. There is also the additional expectation that prospective candidates for lecturing posts is that they should hold a higher degree for example taught Masters, and/ or research degrees such as Masters in Philosophy (MPhil), Doctor of Philosophy (PhD), and Doctorate of Musical Arts (DMA).

Whilst a path of teaching and learning is evident and pursuable, there is not a traditional or formalised route into teaching at higher education level. The

aforementioned qualifications' emphasis on Teaching and Learning elements seem to suggest that these aspects are deemed to be more important than those of its counterparts, i.e. the areas of Assessment and Feedback. The varying experience of tutors going into their relevant fields is vast and leads to disparity across the sector and this chapter looks to investigate what, if anything, we can do to develop the training provided if this is not expected to be received at qualification level.

There is definitely encouragement to study teaching and/ or learning at postgraduate level in HE (Butcher & Stoncel, 2011) but there is a difference on how effective this is. The Postgraduate Certificate in Higher Education (PGCHE) is also branded as Postgraduate Certificate in Teaching in Higher Education (PGCTHE); Postgraduate Certificate in Learning and Teaching in Higher Education (PGCLTHE); Postgraduate Certificate in Teaching and Learning in Higher Education (PGCTLHE); or Postgraduate Certificate in Academic Practice (PGCAP). Whilst all of these titles allude to the practice of studying a combination of teaching and learning they fail to acknowledge the practice of assessment (and indeed feedback). The original PGCHE at least indicates that everything studied belongs to the umbrella of all HE practices. What is surprising is that these qualifications are not compulsory for those teaching at this level which, considering the level of education delivered (and the cost alongside it) it is not compulsory. Again, the emphasis has always been on Teaching and Learning (and to a degree rightly so), but it should not mean that all other practices are seemingly ignored. Interestingly, some committees in universities revolve around the terminology of Learning, Teaching, and Assessment which would be a better way of encompassing the full cycle.

One of the major factors within the way in which HE is staffed concerns the difference between the part-time (PT) member of staff and the full-time (FT) member of staff. There are a few challenges in regard to both roles, but it is important to acknowledge that the function of these members of staff can differ greatly in some aspects of the role, for example the administrational part of the role differs in expectation. We should also acknowledge that there is a definite increase, and desire, in hiring of PT members of staff. One definition describes a member of PT staff as someone who comes in to teach for one to three one-off seminars per year, or someone who takes tutorials over one or two days a week (Day & Hounsell, 1995). The responsibility of the PT role has evolved since 1995 as the explanation that they use to define the PT role would probably be better suited to describing the role of a visiting/guest lecturer of higher education now. Typically, the role has evolved so that the PT member of staff now takes on similar responsibilities as those of a FT member of staff but simply works fewer days a week. A PT lecturer can either be paid pro rata of a full-time equivalent (FTE) or be paid hourly (HPL). PT members of staff make up the majority of the work force in most higher education institutions across the UK. Traditionally speaking, universities have preferred hiring dedicated, committed members of staff who for all intents and purpose would not have time to dedicate their time to their practice. HE typically follows a master and apprentice model where entry into teaching would occur later in life given that extensive professional experience is seen to be a pre-requisite for an individual to qualify to be an expert in a particular area of study. Tutors are expected to be current practitioners of their field which understandably means their ability to commit is increasingly more difficult. The tension between practice and scholarly activity in turn creates challenges as training PT of members of faculty may be problematic given

that they may have limited time to begin with. The impact of the lack of training coincides with the professional standards in the credibility of our abilities as educators. In other corporate roles training is essential so why is it not essential, or seemingly so, at HE?

The need for training was recognised over twenty years ago. Day and Hounsell (1995) stated the following:

> Over the last decade in particular, universities have increasingly recognized the need for and the benefits of systematic training and development – at least in relation to their mainstream teaching staff… It is only much more recently that the spotlight has begun to pick out a less visible and more transient component of the university teaching workforce: the part-time tutors, language assistants and laboratory demonstrators who are being increasingly deployed – especially in universities with a high research profile.

In response to the recognition of FT teaching staff needing to benefit from training and development, universities and colleges do in fact have their own training procedures in place, yet the requirement that PT and FT staff engage in these is not always actively enforced. PT staff may be less ‘visible’ so there is a greater chance that they could be overlooked. The “teaching support roles” in the UK have “no well-established training and development strategies” in comparison to North America where training and development is “commonplace” (Day and Hounsell, 1995). With the addition of more universities joining the sector than ever before and universities also employing postgraduates in graduate ambassador/lecturer roles, there is an

assumption that postgraduates are familiar with the system and therefore do not need training. In popular music educators are increasingly younger in age when they begin their academic careers, in contrast to the traditional master and apprentice model. This is partly to do with universities/colleges/HE providers wanting to employ practitioners who are relevant and on trend. This differs from the traditional occupation of being a lecturer and challenges the notion that being a master in a given area relates directly to the extent of one's studies and experience gained over many years. With this significant development in the way lecturers are hired, the way they are trained therefore should be adapted and modernised. Employed postgraduates should in theory have the relevant skillset having trained as a student, having learned the curriculum being taught; but, do they have the relevant skills to teach and assess? Not necessarily. So, how do we go about training tutors in assessment when the courses across the institutions vary greatly, in addition to the differences in FT and PT staff? Visiting lecturers (and some PT staff) may not have the time to undertake any additional formal training that may be provided by their institution if you take into consideration that they may in fact be teaching across several institutions or practicing as professionals in their field. So, with faculties recruiting PT members and this largely making up the majority of faculty, how do we maintain standards? How do you make training mandatory to staff who have responsibilities and priorities elsewhere? It is a challenge that we face in HE and has not been addressed simply because there is no tangible solution – but just because it has always been that way does not mean that is how it should always continue. There are of course written resources available to educate postgraduate and PT staff. The Centre for Teaching, Learning and Assessment (TLA Centre) was created to advise tutors new to the profession; however, these are external groups to the

higher education system. Similar groups exist but mainly in the US. Is training something that should be taken to policy level? It is commonplace for PT staff to be subject to being double-marked, but is this mandatory? As previously mentioned, part of the argument that is highlighted again here is that in popular music performance there are many in the profession selected for being current practitioners of their field. This aspect is greatly admired in the hiring process due to students valuing those guiding them into their performative careers to be navigating a current industry that is relevant and not outdated. The difficulty here is that most professionals do not have the expertise in the academic field parallel to their practice. In a system that values validity over reliability we can see an ongoing cycle of how in order to push industry, value, and current practitioners we compromise on academic standards as there is no training available where there should be. Furthermore, as previously noted, with competing pressures on their time, PT members of staff may not always be able to commit to their continual professional development within their academic roles. It worth considering that this is not always the reason behind their lack of drive from professional development and that the other factor is the implication of cost and the funding (or lack of) to support the part-time members of staff to take time off and not lose money in the process of their training. We should be reassured that our assessors are skilful and reliable – this much we owe to our students. There is an assumption that a tutor working in HE will know how to assess correctly especially in areas such as performance where grades are awarded against aspects that the practising professional will have experienced time and time again. A more general problem concerns the application of objectivity within an area that relies on judgements that are subjective. There is an element of

training required to grade against learning outcomes and balancing that fine line between technique and performance executed in an academic context.

Inductions for staff is potentially an area for improvement. On the basis of my seven years' experience as an early-career academic so far it is interesting to note that I have not once had a thorough induction to a workplace, or an induction into my job role enabling me to better understand what I am required to do. It is only having recently transferred to a new academic role that I have been through a thorough induction with my new institution. The company is privatised and therefore almost corporate in its approach, but it maintains university policies and procedures towards education. I have been made to feel that I am secure in my new job; however, in most jobs outside of the HE sector, a thorough induction is mandatory in order to train/develop the necessary approach and skillset. Evidently this is, from first-hand experience, a welcoming change and what I believe to be excellent practice – surely this should be shared? It demonstrates once again the importance, and the familiar problem, of being transparent and open throughout the sector and encouraging institutions to share best practice, removing the need to be elite, or secretive in our approaches in success. Relying on internal procedures is simply not good enough; without regulating standards we are allowing good practice such as this training to be overlooked. The Degree Standards Project (2018), led by Advance HE, is a five-year project that looks to do the following:

- design, pilot and deliver different approaches to the professional development of external examiners

- propose evidence-based and cost-effective longer-term approaches to the professional development of external examiners operating across the higher education system in England, Northern Ireland and Wales

- explore approaches to the calibration of standards, presenting recommendations for future work in this area.

It highlights that there is an issue in the training gap for External Examiner's (EE).

The need to support staff is important and this is something that is recognised by Rhodes and Tallantyre (1999). Their argument relates to the challenges that arise from assessing key skills. Academics might be faced with "unfamiliar and non-traditional assessment tools", or they may be assessing within parameters that are not defined by their specialism. In either case, they will need to "cooperate and collaborate with other academic staff outside of their subject areas". They do however argue that students will also need support from the members of staff to guide them and help their understanding of identifying and developing their core skills. Whilst we will need to do this, I am surprised at the number of colleagues that are not comfortable in the idea of assessing another instrument other than their own. Other than technique, which we have yet to resolve how much we assess this by the time we see the end result in the performance, the rest of the assessment should allow for anyone with a good sense of musical knowledge to assess and be well within their rights to do so. Do we as assessors need more training in assessing other disciplines as well as the obvious development of assessing our individual specialisms? In recent discussions with students, it was clear that students wanted additional advice and feedback from other tutors, specifically those that were not necessarily specialists in the students' principal instrument. This is of course also

reflective of the music industry where you would be critiqued by others who may not have expertise in your relative field.

Another area of training and support to discuss is the role of the External Examiner. Elton (1998) summarises that there is “one essential condition which future external examiners must satisfy. They must become knowledgeable and even expert in the theory and practice of assessment”. He goes on to elaborate that this is recognised by the HEQC (Higher Education Quality Council) in 1997 when they “refer to ‘the general absence of assessment training for staff’”. Surely being trained in assessment is vital to the process and practice and becoming “expert” in the field that they advise upon? To understand the role of the external examiner in more detail, the main document to analyse is the *Code of Practice for the Assurance Agency of Academic Quality and Standards in Higher Education Section 4: External Examining*. The code of practice allows institutions the flexibility in choosing what is relevant to the course. It is diverse but informative. There are four sections that relate to assessment, they are the following:

- An institution should require its external examiners, in their expert judgement, to report on [standards].

- Institutions should state clearly the various roles, powers and responsibilities assigned to their external examiners.

- Prior to the publication of mark lists, pass lists or similar documents, institutions should require external examinations to endorse the outcomes of the assessment(s) they have been appointed to scrutinise.

- Institutions should ensure that their external examiners are competent to undertake the responsibilities defined in their contract.

(2000)

The code of practice clearly outlines that the external should report on standards, the standard of student performance, and the comparability of standards with similar programmes across the UK. The document frames the “functions of external examining” where the external could be responsible for setting their own standards, maintaining and verifying these standards, design of courses and the modules/ units within them, the evaluation of assessment policies and procedures (the development and soundness), possible moderator for assessment (or on exception, act as an additional marker for the internal examiners), to adjust marks (collectively and not individually) or alter decisions, and to evaluate the standards of achievement. It also states that the external should have the “appropriate levels of academic and / or professional expertise and experience in relation to the relevant subject area and assessment”, “the ability to command respect of colleagues”, and “the need not to exclude otherwise well-qualified candidates on the grounds that they have no previous experience as external examiners”. Whilst the booklet instructs that previous experience is not necessary, if the candidate is well-qualified, it does not specify who is qualified to start this role if not, necessarily, based on experience. External Examiners vary across the sector and there is a continual flux of contradiction across institutions and through the assessment period. For example, at one institution an EE may advise that adjustments are made where at another institution the opposite could be advised. The process as it currently stands is when

an EE is appointed they go through an induction with the validating institution, but, again, there does not appear to be a rigorous process in becoming an EE which is disconcerting, and contradictory, when these roles are meant to be fulfilled by experts of the field, who understand process from beginning to end, and bring parity across the sector.

One of the potential solutions is to regulate the minimum standard to assess, but what could this look like? In order to regulate, there would need to be compulsory training and we find ourselves returning to the argument that this should be changed at policy level. The assumptions made about PT staff and their level to commit and develop further is unfounded and most in the system wish to further their knowledge and ability. There is a fundamental lack of resource here in many parts of the sector and this is an issue that should be addressed – it directly impacts on those who work freelance in the creative industries who find themselves giving up potentially paid work in order to develop in their practice in academia.

It is evident from the research and ongoing discussions with academics and theorists that there is a lack of training all round for young early-career academics and lecturers as well as the added layer of difficulty in training academics who are resistant to change due to the notion of them feeling like their expertise is being challenged (previously mentioned in Chapter 6). More could be offered at provider-level to support academics, especially those in music who do not necessarily have the academic experience but have the musical knowledge and expertise to teach students.

## CONCLUSION

In an attempt to look at the issues and challenges that arise from the assessment of popular music performance, the thesis has drawn as much information as possible in the related investigation that is lacking from the research pool. Each chapter addresses the literature that exists, what the gaps of research are, and potential solutions that would help to resolve these: be these immediate changes within faculty through to suggesting making changes at policy level. The solutions presented are of course hypothetical, and the problems are not expressed by all practitioners – that does not mean these issues do not exist. They are unrefined in nature, but it attempts to breakdown the barriers that cause assessment processes and outcomes to suffer, whilst creating a platform for discussion and a foundation for moving the research on. These discussions are taking place but through verbatim and not through journal articles or doctorate submissions. I do not know if this is because the paper is too broad in scope or whether or not I am genuinely in the minority. With so much research on other traditional forms of assessment it would seem a natural, and logical, step to look at performance as assessment in more detail – or at least to document the different ways in which this is presented at HE level. Sharing good practice is vital in order to improve and learn from what we already do.

### Summary of Findings

- The research surrounding learning outcomes is thorough, and the concept of aligning learning outcomes with assessment is not new. However, aligning learning outcomes with performance – although implied is not something that

is accurately executed. The revalidation process looks to scrutinise and regulate with advisory panels and external perspectives to develop and enhance the way in which degrees are shaped. The Quality Assurance Agency allows us to work with objectivity in mind when assessing a subject that is naturally subjective. To what extent these being aligned by combining the language of the QAA and the learning outcome statements across the sector would alleviate the disparity is not an impossible feat at policy level.

- Arguments imply that there is always compensation when dealing with reliability and validity and that the two cannot work cohesively together. However, in an ideological world there are solutions to most of the problems that arise – it just has not been found yet or has not been logistically considered to implement. In this case, the changes that could be made would need to be made at policy level to affect a cross-sector change. Reliability is something that can be made comparable through aligning learning outcome statements against the QAA. The UK is recognised as a predominantly validating country where their degrees are concerned. We have worked towards making the students' experience and learning the core of our delivery. Again, ideologically speaking, potentially the answer is to remove one entirely. In this case, eliminating the reliability factor would mean that assessments would not need to be classified but more a pass/ fail mentality.

- Students are our biggest concern – no matter what their learning type. Their lack of understanding of what is expected of them is something the academic sector can change and improve upon. Whilst we can, and rightfully so, blame the government, our student's previous education pathways, there is an

element of us using this as a scapegoat. While we cannot make these changes prior to them coming to us we can work on how we can change their perceptions when they come to us and learn. We need to be able to manage their expectations – especially with consumerism behaviour at the fore.

- Whilst there are benchmarks in the QAA that dictate the lower threshold and the difference between levels (4, 5, and 6) the lack of benchmark for 1st classification is an issue. Musicians operate at different professional levels so again not having guidelines for what this should look like is problematic.

- An assessment design template would be worthy to trial, especially as the Europe funded project (discussed in Chapter 2) created its own assessment template.

- Possibly one of the most important points of discussion and reinforcing factor in concluding is the value in parity and transparency across the sector. The development of popular music would benefit greatly for more sharing of good practice to take place. Organisations such as the National Association of Music in Higher Education, soon to be branded MusicHE to align with other associations such as DanceHE and AdvanceHE, work on bringing these conversations together and moving them forward. My last institution, and current institution, did not (before me working their) know of NAMHE or what they were important for. If these types of institutions are not part of the conversation how do we expect to be part of the bigger conversations and development? Memberships to these types of societies should be made compulsory (considering the membership fee for small corporate departments is only £100).

- Popular music language is one of the biggest challenges that students face in interpreting and understanding. Higher education seems so far removed from further education that students struggle with the change, especially when we then take into considerations that they are also then adapting to the new environment they are studying in. The duty of care we hold for our students is so much so that confusing them even further is neither conducive nor beneficial to their studies. Whilst we do not want to spoon-feed them, we have to accept the variety of backgrounds they come from and modernise where we can without compensating our standards.

- Another change that could be made at policy level is the terminology used within grading criteria. Again, whilst this is a wider pedagogical sector issue, in popular music the criteria in performance is very specific and the language is the biggest part of the confusion.

- The reason for assessing and why we assess does indeed cause a divide in music. There is a school of thought that suggests that we do not assess at all, and merely offer an education that works solely on the teaching and learning aspect. This would work with the theory that we rid reliability altogether.

- What the teaching excellence framework has taught us is that there are a variety of institutions operating at different levels bringing valuable education to the system but do not meet the metrics created due to attempting to uniform excellent teaching with a University model. Subject level TEF could improve this as faculties vary within departments.

- Feedback design could definitely improve for performance. Whilst logistically the time and funding that would go into this for larger departments would be difficult, there is no reason why we should not modernise the way in which we conduct our feedback (or officially recognise feedback sessions). Written feedback for performance feels outdated and irrelevant if the assessment is not tangible and video evidence is not provided. Alternative solutions, such as screencast, and verbal commentary could work, but this is extremely time-consuming if each performance is an hour-long, and how do you monitor how much feedback you give without giving live commentary throughout? The ability to breakdown performance whilst it is fresh in the performer's mind is valuable, and, more importantly, beneficial to the development process of the students learning. If they have performed as a group this too would be better suited to receiving the feedback as a group to understand how they integrate within the ensemble.

- One of the fundamental issues and challenges that arise from the assessment of popular music performance is the notion of professional standards. Currently, this is undefinable. Whilst we know what these are, the differing levels of what this, and the way in which we then relate this to the grading criteria is disparate and extremely difficult to manage. Without a regulatory body that does not oversee this element of our practice, there is no way of maintaining standards across the sector, or, at the end of the day, equality. Unlike traditional subjects, popular music performance does not rely on the hierarchy of league tables as much as other subjects. There may indeed be a difference in perception of conservatoires against other Higher Education Institutes, but on a whole, the hierarchy is not comparable.

- Though this is a wider pedagogical issue (much broader in discipline than in popular music performance) degree certification/ classification begs the question as to why we even grade in the first place. With marking being incomparable between institutions, and then incomparable between subjects the faith in hierarchy systems is frustrating. The diverse nature of education now seems redundant in having competing schools and is potentially what is preventing the transparency of practice between institutions.

- Training throughout the sector is amiss, the research demonstrates that training is a desire of many, which makes sense if this is the quickest and fastest way that a tutor can develop in their position. The difficulty specific to performance is the need to hire those who are current practitioners. They present departments with several difficulties: their availability (gigs, tours etc.), working in other institutions as visiting lecturers, a non-academic background in terms of their own training which can cause a divide between tutor and students.

- Several issues remain in HE assessment that are problematic for popular music. There is an evident gap in the research, specifically with regards to the popular music sector in HE. Performance although acknowledged (briefly) is the one area not discussed in-depth within the wider reading. Other methods such as written assessment, examinations, portfolio, and coursework are discussed at length and improvements are advised within these areas. The discussion in music is recent and there seems to be an increase of importance regarding assessment too, much like the increase in research of general assessment thirty years ago. Study of the processes of

assessment in popular music performance appears to be in its infancy: hence relevant research has been found lacking within the literature review and this supports the originality of the thesis in its intention.

Scope for Further Research

There are a few suggestions from this thesis that are listed below where research could be taken further:

- To further investigate the potential of using one assessment grid for popular music performance that institutions use to cross-reference with their own designed learning outcomes.

- To trial a variety of assessment design for performance to see which is more accurate and efficient in design

- A questionnaire designed to recognise the students understanding of assessment and to improve the language that we use that revolves around assessment.

- What is the consequence of having a pass/ fail design for undergraduate studies?

- What benchmark statements for Level 4, 5, and 6 would look like and how do we define what first-class classification looks like in popular music performance.

- Guidelines of professional standards.
- The importance and value in parity and transparency across the sector.
- What does best practice look like in Popular Music Performance and to publish these findings?
- Popular music language.
- Grading criteria terminology.
- Subject level TEF.
- Feedback design specific to performance.
- Training for part-time lecturers.

## BILBLIOGRAPHY


Academic Standards in Universities (The Reynolds Report) CVCP, 1986

Academic Standards in Universities: A Third Report (The Sutherland Report) CVCP, 1989

Aerts, A. (2014) Well Aren't You Cue-Conscious? Cue-Profiles: Old Learning Approaches Revisited https://dspace.library.uu.nl/handle/1874/289893

Anon. (2017) AEC LEARNING OUTCOMES 2017. Association Européenne des Conservatoires, Académies de Musique et Musikhochschulen - https://www.aec-music.eu/publications/aec-learning-outcomes-2017-en

Archer, L. (2008) Younger academics' constructions of 'authenticity', 'success' and professional identity, Studies in Higher Education, 33:4, 385-403, DOI: 10.1080/03075070802211729

Archer, J. and McCarthy, B. (1988) 'Personal biases in student assessment.' Education Research, 30(2), 142-5

Barlow, J. and Antoniou, M. (2007) Room for improvement: the experiences of new lecturers in higher education, Innovations in Education and Teaching International, 44:1, 67-77, DOI: 10.1080/14703290601081365

Barnett, R. (1992) Improving Higher Education: Total Quality Care. Buckingham: Society for Research into Higher Education/Open University Press.

Beaumont, C., O'Doherty, M. and Shannon, L. (2011) Reconceptualising assessment feedback: a key to improving student learning? Studies in Higher Education, 36:6, 671-687, DOI: 10.1080/03075071003731135

Bennett, R. and Kane, S. (2014) Students' interpretations of the meanings of questionnaire items in the National Student Survey, Quality in Higher Education, 20:2, 129-164, DOI: 10.1080/13538322.2014.924786

Berger, J. L., Girardet, C., Vaudroz, C., & Crahay, M. (2018). Teaching Experience, Teachers' Beliefs, and Self-Reported Classroom Management Practices: A Coherent Network. SAGE Open. https://doi.org/10.1177/2158244017754119

Biggs, J, B. (1996) Enhancing teaching through constructive alignment, Higher Education, 32(3): 347-64

Biggs, J. (n.d.) Aligning teaching for constructing learning. Accessed: https://www.heacademy.ac.uk/sites/default/files/resources/id477_aligning_teaching_for_constructing_learning.pdf

Blackmore, P. (2016) https://www.lfhe.ac.uk/en/research-resources/research-hub/2016-research/the-role-of-prestige-in-uk-universities-vicechancellors-perspectives.cfm

Brown, S. (2018). 'Identifying student need' In R. Matheson, S. Tangney, M. Sutcliffe (eds) Transition In, Through and Out of Higher Education. London: Routledge
Brown, S. and Knight, P. (1994) Assessing Learners in Higher Education, London: Kogan Page.

Brown, S., Rust, C. and Gibbs, G. (1994) Strategies for Diversifying Assessment in Higher Education. Oxford: Oxford Centre for Staff Development.

Boud, D. (2000) Sustainable assessment: rethinking assessment for the learning society, Studies in Continuing Education, 22(2): 151-67.

Butcher, J. and Stoncel, D. (2012) The impact of a Postgraduate Certificate in Teaching in Higher Education on university lecturers appointed for their professional expertise at a teaching-led university: 'It's made me braver', International Journal for Academic Development, 17:2, 149-162, DOI: 10.1080/1360144X.2011.620107

Carless, D. and Boud, D. (2018) The development of student feedback literacy: enabling uptake of feedback, Assessment & Evaluation in Higher Education, 43:8, 1315-1325, DOI: 10.1080/02602938.2018.1463354

Carnegie Mellon University (2020) University Policies. Accessed: https://www.cmu.edu/policies/university-policy-development/index.html

Crook, C., Gross, H. and Dymott, R. (2006) Assessment relationships in higher education: the tension of process and practice, British Educational Research Journal, 32(1): 95-114

Clark, M. B. (1993) Towards the Reduction of Bias when Grading Examination Scripts. (Unpublished paper, available from the author: Accountancy Department, Lincoln University, Canterbury, New Zealand)

Clarke, B. R. (1960) 'The cooling-out function of higher education.' American Journal of Sociology 65, 569.

Cox, J. (2017) 'Future Musicianship and Present Educational Practices: A Response to Eight Questions on the Future of the Conservatoire as an Institution.3 - https://www.musicandpractice.org/volume-3/jeremy-cox-survey/

Daly, J. A. and Dickinson-Markman, F. (1982) 'Contrast effects in evaluating essays'. Journal of Educational Measurement, 19(4), 309-16.

Department for Business Innovation and Skills (2015) Teaching Excellence, Social Mobility and Student Choice. Accessed: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/474227/BIS-15-623-fulfilling-our-potential-teaching-excellence-social-mobility-and-student-choice.pdf

Duke, R. A., and Henninger, J. C. (1998). Effects of Verbal Corrections on Student Attitude and Performance. Journal of Research in Music Education, 46(4), 482–495. https://doi.org/10.2307/3345345

Duke, R. A., and Henninger, J. C. (2002). Teachers' Verbal Corrections and Observers' Perceptions of Teaching and Learning. Journal of Research in Music Education, 50(1), 75–87. https://doi.org/10.2307/3345694

Elton, L. (1998) 'Are UK degrees going up, down or sideways?' Studies in Higher Education 23, 1, 35-42.

Elton, L. and Johnston, B. (2002) Assessment in Universities: A Critical Review of Research, York: Higher Education Academy

Erwin, D. (1997) 'Developing Strategies and policies for changing universities.' In S. Armstrong, G. Thompson and S. Brown (eds) Facing up to Radical Change in Universities and Colleges. London: Kogan Page.

Entwistle, N (1993) Recent Research on Student Learning and the Learning Environment, Birmingham: paper presented to the conference of the Standing Conference on Educational Development.

Falchikov, N. (2004) Improving Assessment through Student Involvement: practical solutions for aiding learning in higher and further education, London: Routledge Farmer.

FHEQ, (2014) UK Quality Code for Higher Education. Accessed: https://www.qaa.ac.uk/docs/qaa/quality-code/qualifications-frameworks.pdf

Foster, F and Hounsell, D (1994) Tutoring. A handbook for postgraduate and other part-time tutors, Edinburgh: Centre for Teaching, Learning and Assessment.

Fullan, M. (1991) The New Meaning of Educational Change, London: Cassell.

Geisinger, K. F. (1980) 'Who are giving all these 'A's' Journal of Teacher Education 3, 2, 11-15.

Gibbons, C. (2012) Stress, Positive Psychology and the National Student Survey Psychology Teaching Review, 18, 2 22-30

Gibbs, G. (1999) Using assessment strategically to change the way students learn, in S. Brown and A. Glasner (eds), Assessment Matters in Higher Education: Choosing and Using Diverse Approaches. Buckinghamshire: SRHE/Open University Press.

Gibbs, G. and Coffey, M. (2004) The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. Accessed: https://journals.sagepub.com/doi/pdf/10.1177/1469787404040463

Glover, D., Law, S., and Youngman, A. (2002) Graduateness and Employability: student perceptions of the personal outcomes of university education, Research in Post-Compulsory Education, 7:3, 293-306, DOI:10.1080/13596740200200132

Groothuijsen, S. E. A., Bronkhorst, L. H., Prins G. T. & Kuiper W. (2019): Teacher-researchers' quality concerns for practice-oriented educational research, Research Papers in Education, DOI: 10.1080/02671522.2019.1633558

GSP (1997a) Final Report of the Graduate Standards Programme. Vol 1. London: Higher Education Quality Council.

GSP (1997b) Final Report of the Graduate Standards Programme. Vol 2. London: Higher Education Quality Council.

Hanrahan, S. J., and Isaacs, G. (2001) Assessing Self- and Peer-assessment: The students' views, Higher Education Research & Development, 20:1, 53-70, DOI: 10.1080/07294360123776

Helson, H. (1964) Adaption-Level Theory. New York: Harper and Row

Henley, J. (2019) Power Balance and CHAT Theory. Learning, Teaching, and Assessment Conference. 17th September 2019. ICMP: London

Heywood, J. (2000) Assessment in Higher Education: Student Learning, Teaching, Programmes and Institutions. Jessica Kingsley Publishers

Higher Education Quality Council (1996) Strengthening external examining. HEQC, London: - https://www.heacademy.ac.uk/project/review-external-examining-arrangements

Higher Education Quality Council (1997) Assessment in Higher Education and the Role of Graduateness, London: HEQC.

Hounsell, D and Murray, R (1992) Essay Writing for Active Learning, Sheffield: CVCP Universities' Staff Development Unit.

House of Commons (2015) The Teaching Excellence Framework: Assessing quality in Higher Education. Accessed: https://publications.parliament.uk/pa/cm201516/cmselect/cmbis/572/572.pdf

Incorporated Society for Musicians (2020) Website: https://www.baccforthefuture.com/

Johnes, J. and Taylor, J. (1990) Performance Indicators in Higher Education. Buckingham: Open University Press/Society for Research into Higher Education.

Johnson, C and Blinkhorn, S (1992) 'Validating NVQ assessment', Competence and Assessment, 20, 10-14.

Johnston, R. J. (1994) Resources, student: staff ratios and teaching quality in British higher education: some speculations aroused by Jenkins and Smith, Transactions, Institute of British Geographers, New Series 19, pp. 359-365.

Kent-Waters, J., Seago, O., and Smith, L. (2018) A Compendium of Assessment Techniques in Higher Education: From Students' Perspectives. LITE

Kift, S. (2009) Articulating a transition pedagogy to scaffold and to enhance the first year student learning experience in Australian higher education. Accessed: http://www.uws.edu.au/__data/assets/pdf_file/0009/709749/Kift_09.pdf

Knight, P. (1991) '"Value-added" and history in public sector higher education.' Public Sector History Newsletter 3, 1, 23-31.

Lortie, D. (1975). Schoolteacher: A Sociological Study. London: University of Chicago Press.

Marton, F., and Saljo, R. (1997). Approaches to learning. In F. Marton, D.J. Hounsell and N.J. Entwistle (Eds.), The experience of learning (2nd. Ed.) (pp. 39-58). Edinburgh: Scottish Academic Press (in press).

Messick, S. (1989) 'Validity', in Linn, RL (ed.) Educational Measurement, 3rd edn, New York: Macmillan.

Miller, C., and Parlett, M. (1974) Up to the Mark: a study of the examination game. Guildford: Society for Research into Higher Education.

National Center for Education Statistics (1992) National Assessment of College Student Learning: Issues and concerns, Washington, DC: US Government Printing Office

Nicholls, G. (2005) New lecturers' constructions of learning, teaching and research in higher education, Studies in Higher Education, 30:5, 611-625, DOI: 10.1080/03075070500249328

Nørsett, B. (2015). Assessment in vocal popular music in higher music education.

O'Donovan, B., Price, M. and Rust, C. (2004) Know what I mean? Enhancing student understanding of assessment standards and criteria, Teaching in Higher Education, 9(3): 325-35

Oppenheim, A. N., Jahoda, A. M. & James, R. L. (1967) Assumptions underlying the use of university examinations, Universities Quarterly, 21, pp.341-351.

Orlando, R. and Speelman, C. (2012). The Effect of Music Teaching Method on Music Reading Skills and Music Participation: An Online Study.

Pelligrino, J. W., Chudowsky, N. and Glaser, R. (eds) (2001) Knowing what Students Know – The Science and Design of Educational Assessment, Washington, DC: National Academic Press.

Price, M. and O'Donovan, B. (2006) Improving performance through enhancing student understanding of criteria and feedback, in C. Bryan and K. Clegg (eds) Innovative Assessment in Higher Education. London: Routledge.

Price, M. and Rust, C. (1999) The experience of introducing a common criteria assessment grid across an academic department, Quality in Higher Education, 5(2): 133-44

Quality Assurance Agency (2006) Code of Practice for the Assurance Agency of Academic Quality and Standards in Higher Education. Section 6: Assessment of students, 2nd edn. Gloucester: QAA

Quality Assurance Agency (2016) Subject benchmark statements: Music. Accessed: https://www.qaa.ac.uk/docs/qaa/subject-benchmark-statements/subject-benchmark-statement-music.pdf?sfvrsn=61e2cb81_4

Reynolds, P. A. (1990) 'Is the external system an adequate guarantee of academic standards?' In C. Loder (ed.) Quality Assurance and Accountability in Higher Education. London: Kogan Page.

Rink, J., Gaunt, H., and Willamon, A. (2017) Musicians in the Making: Pathways to Creative Performance. Oxford University Press

Rogers, G. (1994) 'Measurement and judgement in curriculum assessment systems.' Assessment Update 6, 1, 6-7.

Rorty, R. (1989) Contingency, Irony and Solidarity, Cambridge: Cambridge University Press.

Rowntree, D. (1987) Assessing Students: How Shall We Know Them? Taylor & Francis

Scott, S. J. (2012). Rethinking the Roles of Assessment in Music Education. Music Educators Journal, 98(3), 31–35. https://doi.org/10.1177/0027432111434742

Smith, E. and Coombe, K. (2006) Quality and qualms in the marking of university assignments by sessional staff: an exploratory study, Higher Education, 51(1): 45-69

Smith, G. D., & Powell, B. (2017). Welcome to the journal. Journal of Popular Music Education, 1(1).

South East England Consortium for Credit Accumulation and Transfer (SEEC), Credit guidelines, models and protocols. DfEE, 1996

Stowell, M. and Woolf, H. (2000) Mark distributions and marketing practices in UK higher education, Active Learning in Higher Education, 1(1): 7-27

Surridge, P. (2008) The National Student Survey 2005-2007: Findings and Trends. Accessed:

https://www.hefcw.ac.uk/documents/council_and_committees/council_papers_and_minutes/2008/TQI%2008%2079%20Annex%20A.pdf

Swann, J. and Ecclestone, K. (1999a). Litigation and learning: tensions in improving university lecturers' assessment practice, Assessment in Higher Education, 6(3): 357-75

Tarsh, J. (1990) 'Graduate employment and degree class.' Employment Gazette 98, 10, 489-500

The Degree Standards project (2018) Accessed: https://www.heacademy.ac.uk/degree-standards

Thorndike, R. L. and Hagen, E. R. (1977) Measurement and Evaluation in Psychology and Education. New York: Wiley.

Warhurst, R. P. (2008) 'Cigars on the flight-deck': new lecturers' participatory learning within workplace communities of practice, Studies in Higher Education, 33:4, 453-467, DOI: 10.1080/03075070802211828

Wesolowski, B. C. (2012). Understanding and Developing Rubrics for Music Performance Assessment. Music Educators Journal, 98(3), 36–42. https://doi.org/10.1177/0027432111432524

Wiliam, D. (2010). "The role of formative assessment in effective learning environments", in Hanna Dumont, David Istance and Francisco Benvides (eds). *The Nature of Learning: Using Research to Inspire Practice*, OECD Publishing. http://dx.doi.org/10.1787/9789264086487-8-en

Williams, J. and Cappuccini-Ansfield, G. (2007) Fitness for Purpose? National and Institutional Approaches to Publicising the Student Voice, Quality in Higher Education, 13:2, 159-172, DOI: 10.1080/13538320701629186

William, J. and Kane, D. (2008) Exploring the National Student Survey: Assessment and feedback issues. Accessed: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.547.4159&rep=rep1&type=pdf

Wigdor, A. K. and Garner, W. R. (eds) (1982) Ability Testing: Uses, Consequences and Controversies (2 vols). Washington, DC: National Academic Press.

Yorke, M., Barnett, G., Evanson, P., Haines, C., Jenkins, D., Knight, P., Scurry, D., Stowell, M. and Woolf, H. (2000) Mark distributions and marketing practices in UK higher education, Active Learning in Higher Education, 1(1): 7-27